

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
2 March 2006 (02.03.2006)

PCT

(10) International Publication Number
WO 2006/023973 A2

(51) International Patent Classification:
A01H 1/00 (2006.01)

(74) Agent: MCCALLUM, Jennifer, M.; The McCallum Law Firm, LLC, 132 Kolar Ct., Erie, CO 80516 (US).

(21) International Application Number:
PCT/US2005/030100

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US (patent), UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 22 August 2005 (22.08.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/603,368 20 August 2004 (20.08.2004) US

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/603,368 (CON)
Filed on 20 August 2004 (20.08.2004)

Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv)) for US only*

(71) Applicant (*for all designated States except US*): UNIVERSITY OF UTAH RESEARCH FOUNDATION
[US/US]; 615 Arapleen Drive, Suite 310, Salt Lake City, UT 84108 (US).

Published:

— *without international search report and to be republished upon receipt of that report*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): DREWS, Gary, N. [US/US]; 62 North Pioneer Fork Road, Salt Lake City, UT 84112 (US). KASHAHARA, Ryushiro [JP/US]; 2146 Heritage Center, Salt Lake City, UT 84112 (US). PORTEREIKO, Michael, F. [US/US]; 1620 East 7080 South, Salt Lake City, UT 84112 (US).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CONTROL OF GENE EXPRESSION IN SYNERGID CELLS

(57) Abstract: The present invention provides for compositions and methods of controlling and mediating sexual reproduction in a plant host cell or organism. The method of the present invention provides organisms with specific gene modifications that induce, inhibit or prevent expression of a synergid gene in order to promote, prevent or inhibit fertility.



WO 2006/023973 A2

CONTROL OF GENE EXPRESSION IN SYNERGID CELLS

FIELD OF THE INVENTION

The present invention relates broadly to control and mediation of sexual reproduction in plants; and, specifically to the newly discovered *MYB98* gene of *Arabidopsis thaliana* expressed in synergid cells, and homologues in other species.

BACKGROUND OF THE INVENTION

Angiosperms, or flowering plants, are the most abundant plants on Earth. They reproduce sexually when pollen (the male gamete) is transferred from anther to stigma. A pollen tube grows into the style, turns to enter the micropyle of the ovule, penetrates one of the two synergid cells and ruptures to deliver its contents, which include two sperm cells. The two sperm cells affect "double fertilization" of the egg cell and the central cell located in the female gametophyte of the ovule. The fertilized egg cell then grows into an embryo and the fertilized central cell becomes the endosperm, the nutrient supply for the growing embryo. The synergid cell that is not penetrated by the pollen tube undergoes cell death or apoptosis.

Hence, the female gametophyte plays an important part in the plant life cycle and mediates several reproductive processes including pollen tube guidance, fertilization and induction and maternal control of seed development. A method of controlling one or more of these reproductive processes would greatly facilitate the agricultural market sector enhancement, inhibition or prevention of fertility in plants.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods for controlling and mediating sexual plant reproduction through the manipulation of gene expression. Plant reproduction may be selectively enhanced, inhibited or prevented through the methods and compositions of the present invention.

In a particular embodiment, methods of the present invention comprise deletion or mutation of the synergid gene or synergid gene promoter to create a female sterile phenotype that is incapable of expressing the synergid gene necessary for attracting the pollen tube to the female gametophyte.

In another particular embodiment, the present invention can be utilized to inhibit a synergid gene or synergid gene promoter by introduction of antisense DNA or an antibody, for example, into a plant host cell or organism. Inhibition of the synergid

gene or synergid gene promoter would inhibit or prevent sexual reproduction of the host plant. In a certain embodiment, the present invention could be used to reduce or eliminate weeds. In another embodiment, the present invention could be used to eliminate or reduce the chance of genetically modified plants escaping into nature by eliminating or reducing the possibility of fertilization.

In another embodiment, the present invention can be utilized to enhance expression of the synergid gene by the insertion of a nucleic acid construct that expresses the synergid gene into a cell of the female gametophyte. In a certain embodiment, the nucleic acid construct may be incorporated into a vector using techniques that are well known in the art. In another particular embodiment, the expression of the synergid gene may be regulated by a control element. Such a control element may be selected from the group consisting of, but not limited to, constitutively active, inducible, ectopic, tissue-specific and development stage-specific promoters. For example, up-regulation or over expression of the synergid gene in the synergid cell may attract the pollen tube to the ovule in a shorter amount of time than observed under control conditions. Such an enhancement in the rate of fertilization may help to reproduce rare or endangered plants in a shorter amount of time than would be required without assistance.

In an alternate embodiment, ectopic expression of the synergid gene would mislead or confuse the pollen tube. The pollen tube would then be unable to properly locate and/or enter the micropyle of the ovule, thus resulting in reduced fertilization or sterilization.

In yet another embodiment, expression of the synergid gene in a cell other than the synergid cell may also mislead or confuse the pollen tube. The pollen tube would then be unable to properly locate and/or enter the micropyle of the ovule, thus resulting in reduced fertilization or sterilization.

In all of the above embodiments, the gene expressed in the synergid cell may be MYB98. In this particular embodiment, fertility is enhanced by the gene product of MYB98 and inhibited by mutations to it, for example by mutations such as *myb98-1* or *myb98-2*. In another particular embodiment, the nucleic acid sequence of MYB98 is SEQ ID No 1. In another particular embodiment, the protein sequence of MYB98 is SEQ ID No 2.

In an alternate embodiment, expression of a gene not normally expressed in the synergid cell may be achieved through fusion with the MYB98 promoter.

Methods of the present invention are applicable to a wide range of plants, including monocotyledons and dicotyledons.

In certain embodiments, the compositions and methods comprise the transformation of a plant host organism by introducing a nucleic acid construct into a plant cell or organism and identifying a resulting plant cell or organism in which a desired phenotype is present.

Many alterations and variations of the invention exist as described herein. The invention is exemplified for the plant, *Arabidopsis*. The elements necessary to carry out the methods of the present invention as herein disclosed can be adapted for application in numerous plant species. The invention therefore provides a general method for controlling or mediating sexual reproduction in plants.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to methods and compositions for controlling and mediating sexual reproduction in plants and apoptosis through manipulation of the expression of the synergid gene. Plant reproduction may be selectively enhanced, inhibited or prevented through the methods and compositions of the present invention.

In a particular embodiment, methods of the present invention comprise deletion or mutation of a synergid gene or synergid gene promoter to create a female sterile phenotype that is incapable of expressing the synergid gene necessary for attracting the pollen tube to the female gametophyte. In another particular embodiment, the synergid gene participates in a function or mediates a function selected from the group consisting of seed development, delivery of the contents of the pollen tube, and synergid cell apoptosis. In a certain embodiment, the synergid gene participates in these functions through a physiochemical or biochemical mechanism.

In an embodiment, the gene expressed in the synergid cell may be MYB98. In this particular embodiment, fertility is enhanced by the gene product of MYB98 and inhibited by mutations to it, for example by mutations such as *myb98-1* or *myb98-2*. In another particular embodiment, the nucleic acid sequence of MYB98 is SEQ ID No 1. In another particular embodiment, the protein sequence of MYB98 is SEQ ID No 2.

In another particular embodiment, the present invention can be utilized to inhibit the synergid gene or synergid gene promoter by introduction of antisense DNA

or an antibody, for example, into a plant host cell or organism. Inhibition of the synergid gene or synergid gene promoter would inhibit or prevent sexual reproduction of the host plant. In a certain embodiment, the present invention could be used to reduce or eliminate weeds. In another particular embodiment, the present invention could be used to eliminate or reduce the chance of genetically modified plants escaping into nature by eliminating or reducing the possibility of fertilization.

A method of the present invention in one embodiment provides for increased expression of a synergid gene. The method involves transformation of a cell with a nucleic acid construct minimally including DNA encoding a synergid gene or a fragment thereof. Other schemes based on these general concepts are within the scope and spirit of the invention, and are readily apparent to those skilled in the art.

As used herein, the cells in which genetic manipulation occurs through an exogenous DNA segment or gene that has been introduced through the hand of man are called recombinant cells. Therefore, recombinant cells are distinguishable from naturally occurring cells which do not contain a recombinantly introduced exogenous DNA segment or gene. Recombinant cells include those having an introduced cDNA or genomic gene, those missing an endogenous DNA segment or genomic gene, and also include genes positioned adjacent to a heterologous promoter not naturally associated with the particular introduced gene.

The recombinant cells of the invention may also contain a nucleotide sequence which can be used as a marker of the said recombinant sequences, in particular for differentiation (and thus selection) of those of the plant cells which are transformed by the said recombinant sequences from those which are not. The nucleotide sequence which can be used as a marker of the recombinant sequences may be genes of resistance to antibiotics and the like.

To express a recombinant encoded protein or peptide, whether mutant or wild-type, in accordance with the present invention one would prepare an expression vector that comprises isolated nucleic acids under the control of, or operatively linked to, one or more promoters, which may be inducible, constitutively active or tissue specific, for example. To bring a coding sequence "under the control of" a promoter, one positions the 5' end of the transcription initiation site of the transcriptional reading frame generally between about 1 and about 50 nucleotides "downstream" (i.e., 3') of the chosen promoter. The "upstream" promoter stimulates transcription of the DNA and promotes expression of the encoded recombinant protein. This is the meaning of "recombinant expression" in this context.

Ways of effecting protein expression are well known in the art. One skilled in the art is capable of expressing a protein of his or her choice in accordance with the present invention.

The invention also relates to any genetically transformed plant cell containing one (or more) recombinant nucleotide sequence(s) as described above, according to the invention, integrated into its genome in a stable manner. In a particular embodiment, the sequence is selected from the group consisting of a synergid gene, and mutations of the same, for example.

The invention also relates to genetically transformed seeds containing one (or more) recombinant nucleotide sequence(s) as described above, according to the invention, integrated into their genome in a stable manner. In a particular embodiment, the sequence is selected from the group consisting of a synergid gene, and mutations of the same, for example.

One skilled in the art will also readily understand that progeny of plants manipulated by the compositions and methods of the present invention may be produced asexually, thus allowing sterile plants created by the methods and compositions of the present invention to be reproduced to form a multitude of sterile plants from a single genetically modified specimen.

It is contemplated that the compositions and methods of the present invention can be used in any variety of plant species, including monocots or dicots. In certain embodiments, the invention can be used in plants such as grasses, legumes, starchy staples, *Brassica* family members, herbs and spices, oil crops, ornamentals, woods and fibers, fruits, medicinal plants, poisonous plants, corn, cotton, castor bean and any other crop specie. In alternative embodiments, the invention can be used in plants such as sugar cane, wheat, rice, maize, potato, sugar beet, cassava, barley, soybean, sweet potato, oil palm fruit, tomato, sorghum, orange, grape, banana, apple, cabbage, watermelon, coconut, onion, cottonseed, rapeseed and yam. In some embodiments, the invention can be used in members of the *Solanaceae* specie, such as tobacco, tomato, potato and pepper. In other embodiments, the invention can be used in poisonous ornamentals, such as oleander, any yew specie and rhododendron.

Grasses include, but are not limited to, wheat, maize, rice, rye, triticale, oats, barley, sorghum, millets, sugar cane, lawn grasses and forage grasses. Forage grasses include, but are not limited to, Kentucky bluegrass, timothy grass, fescues, big bluestem, little bluestem and blue gamma. Legumes include, but are not limited

to, beans like soybean, broad or Windsor bean, kidney bean, lima bean, pinto bean, navy bean, wax bean, green bean, butter bean and mung bean; peas like green pea, split pea, black-eyed pea, chick-pea, lentils and snow pea; peanuts; other legumes like carob, fenugreek, kudzu, indigo, licorice, mesquite, copaifera, rosewood, rosary pea, senna pods, tamarind, and tuba-root; and forage crops like alfalfa. Starchy staples include, but are not limited to, potatoes of any species including white potato, sweet potato, cassava, and yams. Brassica, include, but are not limited to, cabbage, broccoli, cauliflower, brussel sprouts, turnips, collards, kale and radishes. In a particular embodiment, the Brassica specie is *Arabidopsis*. Oil crops include, but are not limited to, soybean, palm, rapeseed, sunflower, peanut, cottonseed, coconut, olive palm kernel. Woods and fibers include, but are not limited to, cotton, flax, and bamboo. Other crops include, but are not limited to, quinoa, amaranth, tarwi, tamarillo, oca, coffee, tea, and cacao.

The plant specie of the present invention may further include a transgenic plant containing a genetic modification in order to produce a pharmaceutical or chemical product. Such plants may be used for the production of chemicals or pharmaceuticals as a low-cost manufacturing method to traditional chemical synthetic routes. Plants of the present invention may further provide new fuels, also called biofuels, which have the advantage of being less polluting than fuels derived from petroleum and less expensive to produce in a transgenic plant than a manufacturing plant.

The invention relates to antibodies directed against the recombinant polypeptides of the invention, and more particularly those directed against the products of the synergid cell. In a particular embodiment, the sequence is polypeptide SEQ ID 2. In another embodiment, the antibodies are directed to the protein product of SEQ ID 1.

Such antibodies can be obtained by immunization of an animal with these polypeptides, followed by recovery of the antibodies formed. It goes without saying that this production is not limited to polyclonal antibodies.

It also applies to any monoclonal antibody produced by any hybridoma which can be formed by conventional methods from animal spleen cells, in particular from the mouse or rat, immunized against one of the purified polypeptides of the invention on the one hand, and cells of a suitable myeloma on the other hand, and which can be selected according to its capacity to produce monoclonal antibodies which

recognize the abovementioned polypeptide initially used for immunization of the animals.

Definitions:

For the purposes of the present invention, the following terms shall have the following meanings:

As used herein, the terms "host cell" or "host organism" or, simply, "target host", refer to a cell or an organism that has been selected to be genetically transformed to carry one or more genes for expression of a function used in the methods of the present invention. A host can further be an organism or cell that has been transformed by insertion, deletion or mutation methods of the present invention.

For the purposes of the present invention, the term "synergid gene" refers to an isolated polynucleotide sequence or a fragment or analog of an isolated polynucleotide sequence that is expressed in a synergid cell and which functions in attracting a pollen tube to an ovule, synergid cell death, inducing release of pollen tube contents and/or seed development. In a particular embodiment, a synergid gene expression is affected by SEQ ID NO:1 or a fragment thereof. In a particular embodiment, a synergid protein comprises SEQ ID NO:2 or a fragment thereof.

As used herein, the term "vector" refers to a vehicle that is capable of delivering a nucleic acid sequence into a host cell for replication purposes. Vectors are especially designed to provide an environment which allows the expression of the cloned gene after transformation into the host. One manner of providing such an environment is to include transcriptional and translational regulatory sequences on such vectors, such transcriptional and translational regulatory sequences capable of being operably linked to the cloned gene. In a vector, the gene to be cloned is usually operably-linked to certain control sequences such as promoter sequences. Expression control sequences may contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements, and/or translational initiation and termination sites. The terms expression vector and vector also encompass naked DNA which may be operably linked to a promoter.

For the purposes of the present invention, the term "sequence" means any series of nucleic acid bases or amino acid residues, and may or may not refer to a sequence that encodes or denotes a gene or a protein. Many of the genetic constructs used herein are described in terms of the relative positions of the various genetic elements to each other. For the purposes of the present invention, the term "adjacent" is

used to indicate two elements that are next to one another without implying actual fusion of the two elements. Additionally, for the purposes of the present invention, "flanking" is used to indicate that the same, similar, or related sequences exist on either side of a given sequence. Segments described as "flanking" are not necessarily directly fused to the segment they flank, as there can be intervening, non-specified DNA between a given sequence and its flanking sequences. These and other terms used to describe relative position are used according to normal accepted usage in the field of genetics.

For the purposes of the present invention, the term "controlling sexual reproduction" refers to direct interaction that affects the rate or occurrence of sexual reproduction of a plant. Control of sexual reproduction may arise, for example, from overexpression of a synergid gene in the synergid cell by introduction of a nucleic acid construct that constitutively expresses the gene. Conversely, underexpression or no expression may result in a decreased fertility and/or absolute sterility.

For the purposes of the present invention, the term "mediating sexual reproduction" refers to indirect interaction that affects the rate or occurrence of sexual reproduction of a plant. Introduction of an antibody into the synergid cell of a plant, for example, may mediate sexual reproduction.

As used herein, isolated proteins of the present invention can be full-length proteins or any homologue of such proteins. Examples of proteins include proteins in which amino acids have been deleted (e.g., a truncated version of the protein, such as a peptide), inserted, inverted, substituted and/or derivatized (e.g., by glycosylation, phosphorylation, acetylation, myristoylation, prenylation, palmitoylation, amidation and/or addition of glycerophosphatidyl inositol) such that the homologue includes at least one epitope capable of eliciting an immune response against a protein of the present invention, and/or of binding to an antibody directed against a protein of the present invention. That is, when the homologue is administered to an animal as an immunogen, using techniques known to those skilled in the art, the animal will produce an immune response against at least one epitope of a natural biomarker protein. The ability of a protein to effect an immune response can be measured using techniques known to those skilled in the art.

As used herein, the term "epitope" refers to the smallest portion of a protein or other antigen capable of selectively binding to the antigen binding site of an antibody or a T cell receptor. It is well accepted by those skilled in the art that the minimal size of a protein epitope is about four to six amino acids. As is appreciated by those skilled in the art, an epitope can include amino acids that naturally are contiguous to

each other as well as amino acids that, due to the tertiary structure of the natural protein, are in sufficiently close proximity to form an epitope.

Moreover, for the purposes of the present invention, the term "a" or "an" entity refers to one or more than one of that entity; for example, "a protein" or "an nucleic acid molecule" refers to one or more of those compounds, or at least one compound. As such, the terms "a" or "an", "one or more" and "at least one" can be used interchangeably herein. It is also to be noted that the terms "comprising," "including," and "having" can be used interchangeably. Furthermore, a compound "selected from the group consisting of" refers to one or more of the compounds in the list that follows, including mixtures (i.e. combinations) of two or more of the compounds. According to the present invention, an isolated or biologically pure compound is a compound that has been removed from its natural milieu. As such, "isolated" and "biologically pure" do not necessarily reflect the extent to which the compound has been purified. An isolated compound of the present invention can be obtained from its natural source, can be produced using molecular biology techniques or can be produced by chemical synthesis.

Nucleic Acid Delivery

Transformation can be carried out by a variety of known techniques which depend on the particular requirements of each cell or organism. Such techniques have been worked out for a number of organisms and cells, and can be adapted without undue experimentation to all other cells. Stable transformation involves DNA entry into cells and into the cell nucleus. For single-celled organisms and organisms that can be regenerated from single-cells (which includes all plants), transformation can be carried out in *in vitro* culture, followed by selection for transformants and regeneration of the transformants. Methods often used for transferring DNA or RNA into cells include forming DNA or RNA complexes with cationic lipids, liposomes or other carrier materials, micro-injection, particle gun bombardment, electroporation, use of *Agrobacterium* or other infectious organisms or parts of organisms capable of delivering DNA into a cell and incorporating transforming DNA or RNA into vectors. Other techniques are well known in the art.

Examples of some Delivery Systems useful in practicing the present invention

Transformation of Plants: Transformed plants are obtained by a process of transforming whole plants, or by transforming single cells or tissue samples in culture and regenerating whole plants from the transformed cells. When germ cells or seeds are transformed there is no need to regenerate whole plants, since the transformed

plants can be grown directly from seed. A transgenic plant can be produced by any means known in the art, including but not limited to *Agrobacterium tumefaciens*-mediated DNA transfer, preferably with a disarmed T-DNA vector (which contains no hormone genes to reduce the size of the vector), polyethylene glycol mediated transfer, liposome mediated transfer, laser, electroporation, microinjection, direct DNA transfer, and particle bombardment. Techniques are well known in the art for the introduction of DNA into monocots as well as dicots, as are the techniques for culturing such plant tissues and regenerating those tissues. Regeneration of whole transformed plants from transformed cells or tissue has been accomplished in most plant genera, both monocots and dicots, including all agronomically important crops.

Hybridization

In one embodiment of the present invention, nucleic acid molecules of the present invention hybridize under stringent hybridization conditions to genes or other nucleic acid molecules encoding them, respectively. The minimal size of such sequences of the present invention is a size sufficient to be encoded by a nucleic acid molecule capable of forming a stable hybrid (i.e., hybridizing under stringent hybridization conditions) with the complementary sequence of a nucleic acid molecule encoding the corresponding natural protein. The size of a nucleic acid molecule encoding such a protein is dependent on the nucleic acid composition and the percent homology between the nucleic acid molecule and the complementary nucleic acid sequence. It can easily be understood that the extent of homology required to form a stable hybrid under stringent conditions can vary depending on whether the homologous sequences are interspersed throughout a given nucleic acid molecule or are clustered (i.e., localized) in distinct regions on a given nucleic acid molecule.

The minimal size of a nucleic acid molecule capable of forming a stable hybrid with a gene encoding a of the present invention is at least about 12 to about 15 nucleotides in length if the nucleic acid molecule is GC-rich and at least about 15 to about 17 bases in length if it is AT-rich. The minimal size of a nucleic acid molecule used to encode a protein homologue of the present invention is from about 12 to about 18 nucleotides in length. Thus, the minimal size of homologues of the present invention is from about 4 to about 6 amino acids in length. There is no limit, other than a practical limit, on the maximal size of a nucleic acid molecule encoding a biomarker protein of the present invention because a nucleic acid molecule of the present invention can include a portion of a gene or cDNA or RNA, an entire gene or

cDNA or RNA, or multiple genes or cDNA or RNA. The optimal size of a protein encoded by a nucleic acid molecule of the present invention depends on whether a full-length, fusion, multivalent, or functional portion of such a protein is desired.

Stringent hybridization conditions are determined based on defined physical properties of the nucleic acid molecule to which the nucleic acid molecule is being hybridized, and can be defined mathematically. Stringent hybridization conditions are those experimental parameters that allow an individual skilled in the art to identify significant similarities between heterologous nucleic acid molecules. These conditions are well known to those skilled in the art. See, for example, Sambrook, *et al.*, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Labs Press, and Meinkoth, *et al.*, 1984, *Anal. Biochem.* 138, 267-284, each of which is incorporated by reference herein in its entirety. As explained in detail in the cited references, the determination of hybridization conditions involves the manipulation of a set of variables including the ionic strength (M, in moles/liter), the hybridization temperature (°C), the concentration of nucleic acid helix destabilizing agents (such as formamide), the average length of the shortest hybrid duplex (n), and the percent G + C composition of the fragment to which an unknown nucleic acid molecule is being hybridized. For nucleic acid molecules of at least about 150 nucleotides, these variables are inserted into a standard mathematical formula to calculate the melting temperature, or T_m , of a given nucleic acid molecule. As defined in the formula below, T_m is the temperature at which two complementary nucleic acid molecule strands will disassociate, assuming 100% complementarity between the two strands:

$$T_m = 81.5^\circ\text{C} + 16.6 \log M + 0.41(\%G + C) - 500/n - 0.61(\%\text{formamide}).$$

For nucleic acid molecules smaller than about 50 nucleotides, hybrid stability is defined by the dissociation temperature (T_d), which is defined as the temperature at which 50% of the duplexes dissociate. For these smaller molecules, the stability at a standard ionic strength is defined by the following equation:

$$T_d = 4(G + C) + 2(A + T).$$

A temperature of 5°C below T_d is used to detect hybridization between perfectly matched molecules.

Also well known to those skilled in the art is how base pair mismatch, i.e. differences between two nucleic acid molecules being compared, including non-complementarity of bases at a given location, and gaps due to insertion or deletion of one or more bases at a given location on either of the nucleic acid molecules being compared, will affect T_m or T_d for nucleic acid molecules of different sizes. For

example, T_m decreases about 1°C for each 1% of mismatched base pairs for hybrids greater than about 150 bp, and T_d decreases about 5°C for each mismatched base pair for hybrids below about 50 bp. Conditions for hybrids between about 50 and about 150 base pairs can be determined empirically and without undue experimentation using standard laboratory procedures well known to those skilled in the art. These simple procedures allow one skilled in the art to set the hybridization conditions (by altering, for example, the salt concentration, the formamide concentration or the temperature) so that only nucleic acid hybrids with greater than a specified % base pair mismatch will hybridize. Because one skilled in the art can easily determine whether a given nucleic acid molecule to be tested is less than or greater than about 50 nucleotides, and can therefore choose the appropriate formula for determining hybridization conditions, he or she can determine whether the nucleic acid molecule will hybridize with a given gene under conditions designed to allow a desired amount of base pair mismatch.

Hybridization reactions are often carried out by attaching the nucleic acid molecule to be hybridized to a solid support such as a membrane, and then hybridizing with a labeled nucleic acid molecule, typically referred to as a probe, suspended in a hybridization solution. Examples of common hybridization reaction techniques include, but are not limited to, the well-known southern and northern blotting procedures. Typically, the actual hybridization reaction is done under non-stringent conditions, i.e., at a lower temperature and/or a higher salt concentration, and then high stringency is achieved by washing the membrane in a solution with a higher temperature and/or lower salt concentration in order to achieve the desired stringency.

For example, if the skilled artisan wished to identify a nucleic acid molecule that hybridizes under conditions that would allow less than or equal to 30% pair mismatch with a biomarker nucleic acid molecule of the present invention of about 150 bp in length or greater, the following conditions could be used. For exemplary purposes only, assume that the average G + C content of patient DNA is about 51%, as calculated from known patient nucleic acid sequences. The unknown nucleic acid molecules would be attached to a support membrane, and the 150 bp probe would be labeled, e.g. with a radioactive tag. The hybridization reaction could be carried out in a solution comprising 2X SSC in the absence of nucleic acid helix destabilizing compounds, at a temperature of about 37°C (low stringency conditions). Solutions of differing concentrations of SSC can be made by one of skill in the art by diluting a

stock solution of 20X SSC (175.3 gram NaCl and about 88.2 gram sodium citrate in 1 liter of water, pH 7) to obtain the desired concentration of SSC. The skilled artisan would calculate the washing conditions required to allow up to 20% base pair mismatch. For example, in a wash solution comprising 1X SSC in the absence of nucleic acid helix destabilizing compounds, the T_m of perfect hybrids would be about 85.4°C:

$$81.5^{\circ}\text{C} + 16.6 \log (.15\text{M}) + (0.41 \times 51) - (500/150) - (0.61 \times 0) = 85.4^{\circ}\text{C}.$$

Thus, to achieve hybridization with nucleic acid molecules having about 20% base pair mismatch, hybridization washes would be carried out at a temperature of less than or equal to 65.4°C. It is thus within the skill of one in the art to calculate additional hybridization temperatures based on the desired percentage base pair mismatch, formulae and G/C content disclosed herein. For example, it is appreciated by one skilled in the art that as the nucleic acid molecule to be tested for hybridization against nucleic acid molecules of the present invention having sequences specified herein becomes longer than 150 nucleotides, the T_m for a hybridization reaction allowing up to 20% base pair mismatch will not vary significantly from 65.4°C. Similarly, to achieve hybridization with nucleic acid molecules having about 10% base pair mismatch, hybridization washes would be carried out at a temperature of less than or equal to 75.4°C and to achieve hybridization with nucleic acid molecules having about 5% base pair mismatch, hybridization washes would be carried out at a temperature of less than or equal to 80.4°C.

Furthermore, it is known in the art that there are commercially available computer programs for determining the degree of similarity between two nucleic acid or protein sequences. These computer programs include various known methods to determine the percentage identity and the number and length of gaps between hybrid nucleic acid molecules or proteins. Methods to determine the percent identity among amino acid sequences and also among nucleic acid sequences include analysis using one or more of the commercially available computer programs designed to compare and analyze nucleic acid or amino acid sequences. These computer programs include, but are not limited to, the SeqLab® Wisconsin Package™ Version 10.0-UNIX sequence analysis software, available from Genetics Computer Group, Madison, WI (hereinafter "SeqLab"); and DNAsis® sequence analysis software, version 2.0, available from Hitachi Software, San Bruno, CA (hereinafter "DNAsis"). Such software programs represent a collection of algorithms paired with a graphical user interface for using the algorithms. The DNAsis and SeqLab software, for example,

employ a particular algorithm, the Needleman-Wunsch algorithm to perform pair-wise comparisons between two sequences to yield a percentage identity score, see Needleman, S.B. and Wunch, C.D., 1970, *J. Mol. Biol.*, 48, 443, which is incorporated herein by reference in its entirety. Such algorithms, including the Needleman-Wunsch algorithm, are commonly used by those skilled in the nucleic acid and amino acid sequencing art to compare sequences.

One embodiment of the present invention includes a protein encoded by a nucleic acid molecule of the present invention. This particular protein includes a protein encoded by a nucleic acid molecule that hybridizes under conditions that allow less than or equal to 30% base pair mismatch, under conditions that allow less than or equal to 20% base pair mismatch, under conditions that allow less than or equal to 10% base pair mismatch, under conditions that allow less than or equal to 8% base pair mismatch, under conditions that allow less than or equal to 5% base pair mismatch or under conditions that allow less than or equal to 2% base pair mismatch with a nucleic acid molecule of the present invention.

Another protein of the present invention includes a protein that is encoded by a nucleic acid molecule that is at least 70%, at least 80%, at least 90% identical, at least 92% identical, at least 95% identical or at least 98% identical to a nucleic acid molecule of the present invention; also contemplated are fragments (i.e. portions) of such proteins encoded by nucleic acid molecules that are at least 50 nucleotides.

EXAMPLES

The following examples are included to demonstrate particular embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventors to function well in the practice of the invention, and thus can be considered to constitute representative modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Experimental Parameters

All experiments were carried out using the model genetic organism *Arabidopsis thaliana*, because of a number of desirable features of this system including small size, small genome, and fast growth.

Example 1: Discovery of the *MYB98* Gene and Determination of its Function

The MYB transcription factor gene family was initially screened for expression in wild type, mutant and seedling tissue. It was determined that 17 genes belonging to the MYB transcription factor family were expressed in the wild type (MYB) and not in the mutant (*myb*); 13 of the 17 genes expressed in the wild type were also expressed in seedling tissue. Functional analysis was then performed on the MYB98 gene by linking the MYB gene promoter to Green Fluorescent Protein (GFP) and β -glucuronidase (GUS) genes to allow for expression of a reporter gene when the promoter was activated. It appeared that gene expression had occurred in the micropylar region of the ovule, and specifically in the synergid cells. Expression appeared to be detected in the mature female gametophyte and not during earlier developmental stages. Expression also appeared to be present in the persistent synergid cell for several days after fertilization. Expression was not detected in the embryo, endosperm or seed coat although expression was detected in pistils, which contain synergid cells; and young siliques, which correlates to persistent synergid cells.

A genetic cross experiment utilizing a *myb98* insertional mutant was then performed to observe the affect of the mutation in the male and female gametophyte.

Lines containing two independent T-DNA insertions were utilized, *myb98-1* and *myb98-2*. The T-DNA insertion in *myb98-1* was inserted into the first exon 42 nucleotides downstream of the start codon and 152 nucleotides upstream of the predicted R2 MYB repeat and appeared to be associated with a 23 bp deletion. There did not appear to be any detection of wild-type MYB98 transcripts by RT-PCR which indicated that *myb98-1* may be a null allele. The T-DNA insertion in *myb98-2* is inserted in the genomic sequence 195 nucleotides upstream of the start codon and appeared to be associated with a 19 bp deletion. With both alleles there appeared to be a 50% reduction in seed set upon examination of siliques from heterozygous plants indicating gametophytic lethality.

As shown in Table 1, a cross between heterozygous male and female plants produced a 1:1 ratio of wild type (+/+) to heterozygous (+/-) plants, instead of a Mendelian 1:2:1 (+/+ : +/- : -/-) ratio, indicating the presence of a gametophyte defect. Furthermore, the cross between wild type males and heterozygous females produced almost exclusively wild type progeny, indicating that the defect was present in the female gametophyte. Confirming this finding, the cross between heterozygous males and wild type females resulted in a 1:1 ratio of wild type to heterozygous plants suggesting that the male gametophyte contained no defect. Additionally, plants homozygous for the *myb98-1* or *myb98-2* mutations were present in the progeny of self-pollinated plants, although they had a reduced seed set due to the defect. For example, in *myb98-1* 17% of ovules became fertilized and gave rise to seeds following self-fertilization. Homozygous lines did not appear to exhibit sporophytic defects or trichome defects with regard to either density or branching. Both mutations appear to affect the female gametophyte specifically.

Table 1: Results of Genetic Cross Experiments

Cross		# of Progeny		
Male	Female	MYB98/MYB98	Myb98/MYB98	Myb98/myb98
<i>myb98-1</i> /MYB98	<i>myb98-1</i> /MYB98	109	113	5
MYB98/MYB98	<i>myb98-1</i> /MYB98	304	3	0
<i>myb98-1</i> /MYB98	MYB98/MYB98	160	172	0
<i>myb98-2</i> /MYB98	<i>myb98-2</i> /MYB98	151	157	5
MYB98/MYB98	<i>myb98-2</i> /MYB98	368	5	0
<i>myb98-2</i> /MYB98	MYB98/MYB98	187	217	0

A wild-type copy of the MYB98 gene was then introduced back into the *myb98-1* mutant to determine if the female gametophyte defect was due to disruption of the MYB98 gene. Female plants heterozygous for the *myb98-1* allele and hemizygous for the rescue construct were identified and crossed with wild-type males. In the F1 generation, *myb98-1*/MYB98 and MYB98/MYB98 progeny were present in approximately a 1:2 ratio indicating that disruption of the MYB98 gene is responsible for the female gametophyte defect in *myb98-1* mutants.

The *myb98* mutants were analyzed with confocal laser scanning microscopy to determine if female gametophyte development was affected. The terminal development stage of each (FG7 stage) was examined and the female gametophytes and *myb98* synergid cells appeared normal, indicating that female gametophyte development and the overall morphology of the synergid cells themselves did not appear affected by either the *myb98-1* or *myb98-2* mutations.

The mutants were then examined at the ultrastructural level via transmission electron microscopy. At this level, the cell wall at the micropylar region of *myb98-1* was irregular. Wild-type synergid cells demonstrated a cell wall at the micropylar pole that is extensively invaginated and functions to form a structure referred to as the filiform apparatus. The *myb98-1* mutant lacked the extensive invaginations characteristic of the filiform apparatus and instead appeared to have an abnormal structure that resembled a thickened cell wall. Additionally, the cell wall between the synergic cells was thicker in the mutant than in the wild type. Other aspects of the synergic cell structure at this ultrasonic level appeared the same between the wild-type and mutant.

To determine the functional role of the MYB98 gene in female gametophyte development and function (e.g., synergid death, pollen tube attraction, fertilization, and/or seed development), wild type and *myb98* mutant plants were pollinated. After 36 hours, the wild type synergid cell had been degraded and fertilization had occurred, while the synergid cell of the mutant was intact. The terminal phenotype of the *myb98* mutant was analyzed and no apparent morphological defect was found. Also, in pollination experiments, no evidence of seed, zygote or embryo development was observed, indicating that mutation of MYB98 may inhibit aspects of seed development.

Defects in pollen tube guidance was then analyzed by pollinating *mb98-1/myb98-1* and *myb98-1*/MYB98 pistils with wild-type pollen. Wild-type pollen grew abnormally on ovules containing the mutant gametophyte and appeared to grow from

the placenta to the funiculus but then appeared to fail to grow into the micropyle. Similar results were obtained with *myb98-2*/MYB98 pistils indicating that pollen tube guidance is affected in both mutant female gametophytes.

Pollen tube growth was also examined 24 hours after pollination. At this time point, multiple pollen tubes were present on each *myb98-1* ovule with the majority of them appearing to be present at random positions on the ovule. In approximately 20% (n=100 ovules), one of the pollen tubes appeared to have grown into a micropyle. These may have been the cause of the observed 17% seed set in *myb98-1/myb98-1 mutants*.

In another experiment, LAT52::GUS pollen was used to pollinate wild type and mutant *Arabidopsis*. As shown in Table 2, the majority of wild type plants displayed GUS activity in the ovule suggesting that the pollen had ruptured and delivered its contents. On the other hand, less than 2% of the mutants were GUS positive suggesting that the MYB98 gene affects delivery of the contents of the pollen tube, and possibly pollen tube attraction

Table 2: LAT52::GUS Activity in Wild Type and Mutant Arabidopsis

Genotype of Pollinated Plant	# of Ovules with Ruptured Pollen Tube	
	GUS ⁺	GUS ⁻
Wild Type	74	9
<i>myb98/myb98</i>	3	156

To determine whether the *myb98* mutant is capable of attracting a pollen tube, mutants were pollinated with wild type pollen and stained with Congo Red. Microscopy revealed that the pollen tube had not entered the micropyle of the ovule up to 24 hours after pollination.

Mutants were then examined for the ability to become fertilized and subsequently produce seeds if penetrated by a pollen tube. Thick plastic sections of 30 *myb98-1* female gametophytes were analyzed and all appeared to have defects in the filiform apparatus indicating that with the exception of pollen tube guidance; the synergic cells appeared to function normally in mutant female gametophytes.

Together this data indicates that MYB98 functions in synergid cells, and appears to be involved with pollen tube attraction and delivery of pollen tube contents.

Example 2: Gene Structure

To determine the structure of the MYB98 gene, a cDNA clone was isolated and compared to its genomic sequence. It appeared to have three exons and two introns and to encode a protein of 427 amino acids. It also appears to contain a MYB domain and a putative nuclear localization signal. The MYB domain is a DNA-binding domain found in the vertebrate proto-oncogene *c-myc*. From the structural analysis, it appears MYB98 binds DNA and functions as a transcriptional regulator.

Example 3: Inhibition of the MYB98 Gene

Antibodies of the *MYB98* gene may be generated by overexpressing MYB98 or fragments thereof in *E. coli* or another suitable organism and using the purified protein to immunize animals. Antibodies purified from the immune serum may be inserted into Agro vectors and the vectors expressing the antibody may be introduced into wild type *Arabidopsis*. Fertilization frequencies will be monitored versus a control group to determine whether the antibody is able to control or mediate the function of the synergid gene.

Example 4: Induction of the MYB98 Gene

The MYB98 gene may be incorporated into an Agro vector downstream from a constitutively active promoter. The vector may then be introduced into wild type *Arabidopsis* and the rate of fertilization will be monitored versus a control group.

All of the COMPOSITIONS, METHODS and APPARATUS disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the COMPOSITIONS, METHODS and APPARATUS and in the steps or in the sequence of steps of the methods described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents that are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

We Claim:

1. An isolated nucleic acid molecule, the nucleotide sequence of which consists of SEQ ID NO:1.
2. An isolated nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, or a degenerate variant of SEQ ID NO:1.
3. An isolated nucleic acid comprising a sequence that encodes a polypeptide with the amino acid sequence of SEQ ID NO:2.
4. An isolated nucleic acid comprising a sequence that hybridizes under highly stringent conditions to a hybridization probe the nucleotide sequence of which consists of SEQ ID NO:1, or the complement of SEQ ID NO:1.
5. An isolated nucleic acid comprising a sequence at least 80% identical to SEQ ID NO:1.
6. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:2 and variants at least 80% identical to SEQ ID NO:2.
7. An isolated nucleic acid comprising a sequence that encodes a polypeptide having the sequence of SEQ ID NO:2, or SEQ ID NO:2 with conservative amino acid substitutions.
8. An isolated nucleic acid comprising a sequence that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or of a fragment of SEQ ID NO:2 at least six residues in length.
9. A method of controlling or mediating sexual reproduction in a plant host cell comprising manipulating expression of a synergid gene of the plant host cell.

10. The method of claim 9, wherein the step of manipulating expression of the synergid gene comprises inducing expression, inhibiting expression, or preventing expression.
11. The method of claim 10, wherein said manipulation causes an effect selected from the group consisting of increased fertility, decreased fertility and infertility.
12. The method of claim 9, wherein said manipulation of expression of the synergid gene comprises alteration of SEQ ID 1.
13. The method of claim 9, wherein said manipulation of the synergid gene comprises alteration of SEQ ID 2.
14. The method of claim 10 further comprising the step of introducing a control element.
15. The method of claim 14, wherein the control element is selected from the group consisting of an inducible, constitutive, ectopic, tissue-specific or development stage-specific promoter.
16. The method of claim 9, wherein the plant is selected from the group consisting of a monocotyledon and a dicotyledon.
17. The method of claim 16, wherein the plant is selected from the group consisting of maize, rice, wheat, potato, soybean, tomato, tobacco, members of the *Brassica* family, and *Arabidopsis*.
18. A method of producing a transgenic plant comprising alteration of SEQ ID 1.
19. The process of claim 18, wherein female-sterile plants are produced.
20. A nucleic acid vector for transformation of a plant cell and including the polynucleotide according to claims 1, 2, 3, 4, 5, 7 and 8.

21. A plant host cell containing a heterologous polynucleotide or nucleic acid vector each comprising the polynucleotide according to claims 1, 2, 3, 4, 5, 7 and 8.
22. A plant comprising the plant host cell according to claim 21.
23. A method of producing a plant, the method including incorporating a polynucleotide according to claims 1, 2, 3, 4, 5, 7 and 8 into a plant cell and regenerating a plant from said plant cell.
24. A method according to claim 23 further including asexually propagating or growing off-spring or a descendant of the plant regenerated from said plant cell.
25. An expression vector comprising the nucleic acid of claims 1, 2, 3, 4, 5, 7 and 8 operably linked to a control element.
26. An expression vector of claim 25 wherein said control element is selected from the group consisting of an inducible, constitutive, ectopic, tissue-specific or development stage-specific promoter.
27. A cultured cell comprising the vector of claim 25.
28. A cultured cell transfected with the vector of claim 25, or a progeny of said cell, wherein the cell expresses the polypeptide.
29. A method of producing a protein, the method comprising culturing the cell of claim 28 under conditions permitting expression of the polypeptide.
30. A method of producing a protein, the method comprising culturing the cell of claim 28, under conditions permitting expression under the control of the control element and purifying the polypeptide from the cell or the medium of the cell.
31. The method of claim 29, wherein said polypeptide is applied to plants in order to prevent fertilization of such plants.

- 32. A plant cell comprising a protein of claim 6.
- 33. A DNA microarray comprising SEQ ID NO 1.
- 34. A DNA microarray comprising the nucleic acid of claims 1, 2, 3, 4, 5, 7 and 8.

FILED PCT 007180-52 Control of Gene Expression in Synergid Cell Patent In.ST25.txt
SEQUENCE LISTING

<110> The University of Utah Research Foundation
 <120> CONTROL OF GENE EXPRESSION IN SYNERGID CELLS
 <130> 007180-52
 <140> NOT ASSIGNED
 <141> 2005-08-22
 <150> US 60/603,368
 <151> 2004-08-20
 <160> 2
 <170> PatentIn version 3.3
 <210> 1
 <211> 3598
 <212> DNA
 <213> Arabidopsis thaliana

<400> 1
 ggtgaagaga gagagagaga gattgagaaa gagaaggata gtgtcagaga aagtaaaaga 60
 gagaaagaga agaagaagga agagaaagat tggatttttc aaagatgggt catgtgacag 120
 ttgggttttg tatactttta gccttctgtt tttggtttaa ttacgatttt gaccttattc 180
 taatcttctt cttttcttct atatgaactc ttattttgtt gcatttagta ttttaactatg 240
 tttttggctt tttagttcaa ctttggttgt tatggataga gattcttttg ttgttttttt 300
 gctatgtgtt agaaactttg acaaagttaa acaacctttt aaaactcatt attcgacatg 360
 taccaaccat aatcaattta tttgctttac attttctcta tgtcattgta accattcaaa 420
 actctataaa gaagtgtata aataaggat attaccaa gttatttcac caggatgatt 480
 ttatgctctg ccctggattt gatcgtgtaa tctaagactc cattgatttt gcttttgtaa 540
 atagtccatc gatattagca ttagtccagt ctttatgcat tgagggtggg acaagaagaa 600
 tatcatagga atcatatttt tttgtaaccc ccatacatga attaataagg gaggtctttt 660
 aaaaagtttt cggacagAAC tcagatttag gcagaaaaaa gaaatacatg catgaactaa 720
 agaactaaga agttgtacac tgcaagaaac accattagag tcacactgtt aacactttgt 780
 tataaacgat gcggtttttg agtttttttt ttgtttcact ttggtcagag actttttgtg 840
 agggtttact tctagaagac ggagtgaaca aaatgaaaat tagaccttag agctaatatg 900
 gtcaatacac tcattgtcct tcggcattta aaatggtaat tacatttttg ttgactaaat 960
 taatactaca ctttatgata aattaactaa agacaactaa atatacaata aatgagtcatt 1020
 ttacacatta aaaactagat ttaattcata ctagcgtttt cacaacatt caactatggt 1080

FILED PCT 007180-52 Control of Gene Expression in Synergid Cell Patent In.ST25.txt

taccccaa	tcaacaacac	ctaattgtga	acgtcaacaa	agattattat	ctcacatttt	1140
ctaacttgat	tttcaatcta	attatatatttc	tatcaaaata	tggtaaaaga	aataattaat	1200
ttgaactatt	tttagtaata	cctctaataa	tctaactagg	ataaaaaaaa	aaaaaatct	1260
cattgtaa	taacttcgaa	attaattttc	ttaattttac	ctcactttca	tgttgagata	1320
tgatttaaga	aaaacacaaa	aatggctaag	catttaagaa	ataattttaa	aaaataatta	1380
attaatttaa	tatatacaaa	atcacactaa	tgctaccaac	aaactacaaa	cgtttagattt	1440
caatccataa	caaatttatg	ggtatagtaa	aaacacacat	taaatgtctt	ttctttcttc	1500
tcctacctta	caaatacgag	aaagaacaaa	gctaacacaa	aatcggaaaa	agacttcact	1560
tttttttttt	tactaattaa	ttagtttttt	ttttctcctt	tccaaaacaa	tggaagaattt	1620
cgtcgacgag	aatgggtttt	cttctctaaa	ccaaaacatc	ttcacacgtg	atcaagaaca	1680
catgaaagaa	gaagattttc	cattcgaagt	cgtcgaccaa	tcaaaaccta	caagctttct	1740
tcaagatttt	caccatcttg	atcatgatca	tcagtttgat	catcatcatc	atcatggctc	1800
ctcatcttca	catcctttgc	tcagcgtcca	aactacgtct	tcttgatca	ataatgctcc	1860
tttcgagcat	tgctcttacc	aagaaaacat	ggtcgatttc	tatgaaacta	aaccaaattt	1920
gatgaatcat	catcatttcc	aagcagtggg	aaactcatac	ttcactcgta	atcatcatca	1980
tcatcaagag	atcaatttgg	tcgatgaaca	tgatgatcct	atggacttgg	agcaaaacaa	2040
catgatgatg	atgaggatga	tcccttttga	ttaccctcct	acagagactt	tcaaacctat	2100
gaacttcgta	atgccagatg	aaatttcatt	tgtttctgca	gataatgatt	gttatagagc	2160
aacaagtttc	aacaagacca	aaccatttct	tacacgaaag	ttgtcttctt	cttcttcatt	2220
atcatcatgg	aaagaaacca	aaaagtcaac	cttagtcaaa	ggacaatgga	ctgctgaaga	2280
agacagttcg	ttattaacac	aagaaacctc	aagaatcaat	taacataatc	aagacttttg	2340
atgattacac	tattctagag	gtttttacta	tagtcacccg	tttttgcttg	taaatgaagt	2400
tttcttggtt	tggaataat	aggggtactga	ttcaactcgt	ggagaagtat	ggattgcgta	2460
aatgggtcga	tatcgctcaa	gtgttaccgg	gaagaatcgg	gaaacaatgt	agagagaggt	2520
ggcataacca	tttgagacct	gacattaagg	tattttacta	gtagctataa	gatcatgtat	2580
tgatgattat	ggaacctaaa	agataaagat	tgttgcagaa	agaaacatgg	agtgaagaag	2640
aggacagagt	gttgatagaa	tttcacaaa	agattggaaa	caaattgggca	gagattgcga	2700
aaagactccc	gggaagaaca	gagaactcga	tcaagaacca	ttggaacgca	acaaaaagaa	2760
gacaattctc	taaaagaaaa	tgtagatcta	agtatccaag	accttctctg	ttgcaggatt	2820
acatcaagag	cttgaatatg	ggagctttga	tggtcttctt	tggtcctgca	agaggtagac	2880

FILED PCT 007180-52 Control of Gene Expression in Synergid Cell Patent In.ST25.txt
gcagagagag taataacaag aagaaggatg ttgttggtgc gggtgaggag aagaagaagg 2940

aagaggaggt gtatggacaa gacaggattg tgcctgaatg tgtgtttact gatgattttg 3000
gattcaatga gaagctgctt gaggaaggat gtagcattga ctctttgctt gatgacattc 3060
ctcagcctga cattgatgct tttgttcatg gactctgatt tgtatttttt attctgcttg 3120
tttcagtttt gttgtttttt gtttgtcttt ttatacgaga cagattccac caaacttcaa 3180
taatttgaaa agatataaaa tattttgctt tttaaaaaca tgttgatggt gaaacaatag 3240
aaacctaact tgcggaaccc gccaaagactt gttgctaaaa taagaacatt acttaacttg 3300
tggtcaagc aaccatatga aaatttataa aacaaccaa acacatatga tcaatctttc 3360
aaaatacaaa cataacgtat cttctcttag caatcgatca aaagacagtt cagagaaagg 3420
gaagtcgtat cggtttttga gacgaaagga ttgatacgaa cccaaacaag tgaaaacaca 3480
gaagcaagca agatagacca gagaataaca atagtaggtg ttctgttttg tcttcccata 3540
agacctttga ggaatggata aagatgaaga atcacccaaa aggcgaaaaa taccttcc 3598

<210> 2
<211> 427
<212> PRT
<213> Arabidopsis thaliana

<400> 2

Met Glu Asn Phe Val Asp Glu Asn Gly Phe Ala Ser Leu Asn Gln Asn
1 5 10 15

Ile Phe Thr Arg Asp Gln Glu His Met Lys Glu Glu Asp Phe Pro Phe
20 25 30

Glu Val Val Asp Gln Ser Lys Pro Thr Ser Phe Leu Gln Asp Phe His
35 40 45

His Leu Asp His Asp His Gln Phe Asp His His His His His Gly Ser
50 55 60

Ser Ser Ser His Pro Leu Leu Ser Val Gln Thr Thr Ser Ser Cys Ile
65 70 75 80

Asn Asn Ala Pro Phe Glu His Cys Ser Tyr Gln Glu Asn Met Val Asp
85 90 95

Phe Tyr Glu Thr Lys Pro Asn Leu Met Asn His His His Phe Gln Ala
100 105 110

FILED PCT 007180-52 Control of Gene Expression in Synergid Cell Patent In.ST25.txt

Val Glu Asn Ser Tyr Phe Thr Arg Asn His His His His Gln Glu Ile
 115 120 125

Asn Leu Val Asp Glu His Asp Asp Pro Met Asp Leu Glu Gln Asn Asn
 130 135 140

Met Met Met Met Arg Met Ile Pro Phe Asp Tyr Pro Pro Thr Glu Thr
 145 150 155 160

Phe Lys Pro Met Asn Phe Val Met Pro Asp Glu Ile Ser Cys Val Ser
 165 170 175

Ala Asp Asn Asp Cys Tyr Arg Ala Thr Ser Phe Asn Lys Thr Lys Pro
 180 185 190

Phe Leu Thr Arg Lys Leu Ser Ser Ser Ser Ser Ser Ser Trp Lys
 195 200 205

Glu Thr Lys Lys Ser Thr Leu Val Lys Gly Gln Trp Thr Ala Glu Glu
 210 215 220

Asp Arg Val Leu Ile Gln Leu Val Glu Lys Tyr Gly Leu Arg Lys Trp
 225 230 235 240

Ser His Ile Ala Gln Val Leu Pro Gly Arg Ile Gly Lys Gln Cys Arg
 245 250 255

Glu Arg Trp His Asn His Leu Arg Pro Asp Ile Lys Lys Glu Thr Trp
 260 265 270

Ser Glu Glu Glu Asp Arg Val Leu Ile Glu Phe His Lys Glu Ile Gly
 275 280 285

Asn Lys Trp Ala Glu Ile Ala Lys Arg Leu Pro Gly Arg Thr Glu Asn
 290 295 300

Ser Ile Lys Asn His Trp Asn Ala Thr Lys Arg Arg Gln Phe Ser Lys
 305 310 315 320

Arg Lys Cys Arg Ser Lys Tyr Pro Arg Pro Ser Leu Leu Gln Asp Tyr
 325 330 335

Ile Lys Ser Leu Asn Met Gly Ala Leu Met Ala Ser Ser Val Pro Ala
 340 345 350

FILED PCT 007180-52 Control of Gene Expression in Synergid Cell Patent In.ST25.txt
Arg Gly Arg Arg Arg Glu Ser Asn Asn Lys Lys Lys Asp Val Val Val
355 360 365

Ala Val Glu Glu Lys Lys Lys Glu Glu Glu Val Tyr Gly Gln Asp Arg
370 375 380

Ile Val Pro Glu Cys Val Phe Thr Asp Asp Phe Gly Phe Asn Glu Lys
385 390 395 400

Leu Leu Glu Glu Gly Cys Ser Ile Asp Ser Leu Leu Asp Asp Ile Pro
405 410 415

Gln Pro Asp Ile Asp Ala Phe Val His Gly Leu
420 425