Title: PROCESS OF PRODUCING MALE STERILE MONOCOTYLEDONOUS PLANTS

Abstract: The present invention relates to a process of producing male sterile monocotyledonous plants by introducing into said plants fragments of a nucleotide sequence coding for a protein which provides for male sterility and reconstituting the complete protein which provides for male sterility by intein-mediated trans-splicing.
The present invention relates to a process of producing male sterile monocotyledonous plants by introducing into said plants fragments of a nucleotide sequence coding for a protein which provides for male sterility and reconstituting the complete protein which provides for male sterility by intein-mediated trans-splicing.

In agriculture, the demand for male sterile plants is increasing for several reasons. First, the use of male sterile plants leads to an enhanced safety of transgenic plants, as these male sterile plants allow to control transgene flow and avoid the undesired and random transfer of foreign genes to related plant species through pollen.

Second, it allows the production of hybrid seeds which are produced by cross-pollination of genetically different parental lines. The hybrid progeny shows the so-called "heterosis effect" which means that they display superior plant growth, seed yield and a pronounced stress tolerance in comparison to both parental lines.

For naturally self-pollinating plants, the plant must be male sterilized during the crossing process in order to avoid self-fertilization. In corn, castration of the female crossing partner is easily achieved by mechanical detasselling (removal of anthers), and is followed by pollination with pollen from another line. However, this technique is not commercially feasible in the case of crops with small flowers, i.e. all economically important crops other than maize.

To induce male sterility on a cellular basis, both cytoplasmic male sterility (CMS) and nuclear male sterility (NMS) systems have been developed. However, CMS-based hybridization technology is a multi-component genetic system that is difficult to discover, introgress and maintain. First, a genetic source of male sterility must be identified. Second, for the propagation of male sterile phenotypes, so-called "maintainer lines" are needed. Third, in crops in which the seed or fruit is the
harvested product, fertility in the hybrid must be restored, thus fertility restorer lines are required. Historically, these genetic components have had to be discovered and brought together for each crop species separately. As a result, the existing hybridization systems introduced by classical breeding methods are difficult to develop, complex to maintain, often marginally reliable and species-specific.

Hence, nowadays it is an object to develop nuclear male sterility systems which are characterized by mutations in the genome of the plant. One transgenic system which was first developed in rapeseed and which is used commercially for hybrid production is based on male sterility conferred by the tapetum-specific expression of the toxic enzyme barnase which is a ribonuclease from Bacillus amyloliquefaciens (Mariani et al. (1990) Nature 347: 737 - 741; Mariani et al. (1992) Nature 357: 384 - 387). This barnase system for inducing male sterility has also been successfully used in monocotyledonous plants (De Block et al. (1997) Theor. Appl. Genet. 95: 125 - 131).

However, such a one-component system using continuous coding sequences for the genes of interest could suffer from a lack of specificity and therefore secondary, undesirable effects are possible if, for example, gene flow transfers the functional transcriptional unit to a recipient. Furthermore, it cannot be guaranteed that the male fertility is completely restored.

Hence, in a first approach the barnase enzyme was split into two inactive peptides wherein each partial peptide fragment carries at least one active site residue, so that neither peptide has enzyme activity (Burgess et al. (2002) The Plant Journal 31(1): 113 - 125). The ribonuclease activity is then reconstituted by crossing parents wherein one parent has the first fragment of the barnase enzyme and the other parent has the second fragment of the barnase enzyme. However, the two fragments present
in the progeny of these crosses are not linked by peptide bonds. Hence, the barnase activity is not stable at high temperatures and the progeny plants are not suitable for growing in the field.

Further approaches used intein-mediated trans-splicing of proteins to assemble the two fragments of a protein which confers male sterility into one complete polypeptide, wherein both fragments are linked by peptide bonds.

WO 03/102197 A1 describes a process for producing a transgenic multi-cellular plant expressing a trait of interest, e.g. male sterility, with said trait having a controlled distribution of said trait to progeny, wherein a first fragment of a nucleotide sequence encoding said trait is located on a first locus of a nuclear chromosome and a second fragment of a nucleotide sequence encoding said trait is located on a second locus of a nuclear chromosome. Upon crossing the first and the second plant, progeny is created which has the functional trait of interest due to the binding between the protein encoded by said first heterologous nucleotide sequence and the protein encoded by said second heterologous nucleotide sequence. The polypeptides are assembled to a functional protein by intein-mediated trans-splicing of the peptides.

A similar system was also utilized by Gils et al. (2008) Plant Biotechnol. J. 6(3): 226 - 235 which describe a two-component hybrid system for producing male sterile plants, in which system the coding information for both male sterility and herbicide resistance is divided at two separate loci. Hence, only progeny which inherit both loci will show male sterility and herbicide resistance. However, the operability of this system was only shown in dicotyledonous plants.
Therefore, there is still a need to develop a functional trans-splicing system which works for monocotyledonous plants.

Thus, it is an object of the present invention to provide a method for producing male sterile monocotyledonous plants which remain male sterile at higher temperatures and minimize the risk of transgene escape by strictly controlling the distribution of the transgene to the progeny. This enables the production of a biologically safe transgenic plant which minimizes the undesired transmission of the transgenic trait to other crops. Furthermore, these male sterile monocotyledonous plants may be used in crosses with male fertile plants to produce hybrid plants with superior characteristics.

Another object of the present invention is to provide a process of producing a male sterile transgenic plant, whereby distribution of the male sterility phenotype to the progeny is strictly controlled and occurs with low probability.

These and other objects of the invention are attained by the subject-matter of the independent claims. Advantages and embodiments are defined in the dependent claims.

Hence, the present invention provides a method of producing male sterile monocotyledonous plants, comprising the steps of:

a) introducing into a monocotyledonous plant or plant cell a first expression cassette comprising the following elements in 5’ to 3’ orientation:

- a tapetum-specific promoter functional in cells of a monocotyledonous plant;
- operatively linked thereto a nucleic acid sequence coding for an N-terminal part of a protein which provides for male sterility;
- a nucleic acid sequence coding for the N-terminal part of a first intein; and
- operatively linked thereto a terminator sequence functional in plant cells; and

b) introducing into said plant or plant cell a second expression cassette comprising the following elements in 5’ to 3’ orientation:

- a tapetum-specific promoter functional in cells of a monocotyledonous plant;
- operatively linked thereto a nucleic acid sequence coding for a C-terminal part of said first intein
- a nucleic acid sequence coding for at least one copy of a flexible linker sequence;
- a nucleic acid sequence coding for a C-terminal part of said protein which provides for male sterility; and
- operatively linked thereto a terminator sequence functional in plant cells.

A further aspect of the present invention provides a method of producing monocotyledonous hybrid plants, comprising the steps of:

a) producing a male sterile monocotyledonous plant by a method of the present invention and

b) crossing the male sterile monocotyledonous plant of step a) with a male fertile monocotyledonous plant.

In a further aspect, the present invention relates to a transgenic monocotyledonous plant produced by a method of the present invention.

Further, the present invention relates to a transgenic monocotyledonous plant comprising a nucleic acid sequence coding for an N-terminal part of barnase.

In a further aspect, the invention provides a transgenic monocotyledonous plant comprising a nucleic acid sequence coding for a C-terminal part of barnase.
Further, the present invention relates to a transgenic monocotyledonous plant comprising a nucleic acid sequence coding for an N-terminal part of acetolactate synthase.

In a further aspect, the invention provides a transgenic monocotyledonous plant comprising a nucleic acid sequence coding for a C-terminal part of acetolactate synthase.

Further, the present invention relates to a recombinant nucleic acid molecule comprising the following elements:

a) a first expression cassette comprising the following elements in 5' to 3' orientation:
   - a tapetum-specific promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for an N-terminal part of a protein which provides for male sterility;
   - a nucleic acid sequence coding for the N-terminal part of a first intein; and
   - operatively linked thereto a terminator sequence functional in said plant cells; and

b) a second expression cassette comprising the following elements in 5' to 3' orientation:
   - a tapetum-specific promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for a C-terminal part of said first intein
   - a nucleic acid sequence coding for at least one copy of a flexible linker sequence;
   - a nucleic acid sequence coding for a C-terminal part of said protein which provides for male sterility; and
   - operatively linked thereto a terminator sequence functional in said plant cells.
In still a further aspect, the present invention provides an isolated nucleic acid molecule comprising a nucleic acid sequence according to SEQ ID No. 51 or a fragment of said nucleic acid sequence coding for a functional acetolactate synthase fragment.

Further, the present invention provides an isolated nucleic acid molecule coding for the N-terminal part of an acetolactate synthase, selected from the group consisting of:

a) a nucleic acid sequence according to SEQ ID No. 33 or 53;
b) nucleic acid sequences coding for a protein according to SEQ ID No. 34 or 54 or a functional fragment thereof;
c) nucleic acid sequences hybridising to a complementary strand of the nucleic acid sequence according to SEQ ID No. 33 or 53 under stringent conditions; and
d) nucleic acid sequences which are at least 50% homologous to the nucleic acid sequence shown in SEQ ID No. 33 or 53.

Further, the present invention provides an isolated nucleic acid molecule coding for the C-terminal part of an acetolactate synthase, selected from the group consisting of:

a) a nucleic acid sequence according to SEQ ID No. 43 or 55;
b) nucleic acid sequences coding for a protein according to SEQ ID No. 44 or 56 or a functional fragment thereof;
c) nucleic acid sequences hybridising to a complementary strand of the nucleic acid sequence according to SEQ ID No. 43 or 55 under stringent conditions; and
d) nucleic acid sequences which are at least 50% homologous to the nucleic acid sequence shown in SEQ ID No. 43 or 55.
Finally, the present invention relates to an isolated nucleic acid molecule comprising a nucleic acid sequence according to SEQ ID No. 49 or a fragment of said nucleic acid sequence coding for a functional barnase fragment.

The inventors of the present invention have surprisingly found that a vector which led to a high frequency of male sterile dicotyledonous plants among primary transformants did not produce male sterile plants when introduced into monocotyledonous plants.

Hence, specific adaptations were necessary to use the intein-mediated trans-splicing system in a monocotyledonous plant to induce male sterility. These adaptations mainly involve the use of flexible linker sequences. In preferred embodiments, additionally the nucleic acid sequence coding for the protein which provides for male sterility was adapted to the codon usage of monocotyledonous plants. These specific adaptations led to a high frequency of monocotyledonous plants displaying complete or partial male sterility. The male sterile phenotype was stably inherited and displayed a pronounced robustness against extreme temperature. In general, the vegetative phenotype of the sterile progeny was, under greenhouse conditions, indistinguishable from that of the fertile control plants.

In the process of the invention, the nucleotide sequence encoding a protein which provides for male sterility is split into two fragments, thus obtaining a 5' and a 3' part of the nucleotide sequence. Said 5’ part encodes the N-terminal part of the protein which provides for male sterility and said 3’ part encodes the C-terminal part of the protein which provides for male sterility.
Said nucleotide sequence is typically a coding sequence (or an open reading frame) of a protein providing male sterility. However, said nucleotide sequence may also contain one or more introns.

To obtain said 5' and 3' part of the nucleotide sequence, said nucleotide sequence is preferably split such that each obtained fragment, upon expression, is incapable of generating a protein which provides for male sterility in the absence of the other fragment. Each fragment contains a sequence portion necessary for the function of the protein providing for male sterility. For example, if said protein is an enzyme, each fragment preferably contains amino acids necessary for catalysis or substrate binding of the enzyme. The protein providing for male sterility may be split into said fragments in many different ways provided that expression of the male sterility requires all said fragments and binding thereof to each other. Structural and functional information known about the protein providing for male sterility may be helpful for finding a suitable splitting site of said nucleotide sequence. In any case, one can easily test experimentally whether a fragment generated by splitting a nucleotide sequence at a randomly chosen site is capable of providing male sterility by expressing the fragment in the tapetum of plants and investigating if these plants are able to develop viable pollen. A further assay for testing the functionality of the fragments is described in example 6, i.e. syringe infiltration of *Nicotiana benthamiana* leaves using Agrobacterium containing either an N-terminal part or a C-terminal fragment of the protein and detection of lesions in the leaves. If lesions occur upon infiltration only with the N-terminal or the C-terminal fragment, these fragments are not suitable for use in the method of the present invention.

Expression of male sterility requires the presence of both fragments in the same plant, preferably in the same cells thereof. Expression of male sterility further requires transcription and translation of said first and said second fragment and
binding of the translation products of said fragments with peptide bond formation to restore a functional protein.

This peptide bond formation is accomplished by intein-mediated trans-splicing. For this purpose, said first and said second expression cassette further code for inteins capable of mediating protein trans-splicing. By said trans-splicing, the proteins and polypeptides encoded by said first and said second fragments may be linked by peptide bond formation. Trans-splicing inteins may be selected from the nucleolar and organellar genomes of different organisms including eukaryotes, archaeabacteria and eubacteria. Inteins that may be used for performing this invention are listed at http://www.neb.com/neb/intcins.html. The nucleotide sequence coding for an intein may be split into a 5' and a 3' part that code for the 5' and the 3' part of the intein, respectively. Sequence portions not necessary for intein splicing (e.g. homing endonuclease domain) may be deleted. The intein coding sequence is split such that the 5' and the 3' parts are capable of trans-splicing. For selecting a suitable splitting site of the intein coding sequence, the considerations published by Southworth et al. (1998) EMBO J. 17: 918 - 926 may be followed. In constructing the first and the second expression cassette, the 5' intein coding sequence is linked to the 3' end of the first fragment coding for the N-terminal part of the protein which provides for male sterility and the 3' intein coding sequence is linked to the 5' end of the second fragment coding for the C-terminal part of a protein which provides for male sterility.

Herein, peptide bond means the amide linkage between the carboxyl group of one polypeptide and the amino group of another polypeptide.

Within the scope of the present invention, the term "male sterile plants" is intended to mean plants which are unable to produce functional pollen and therefore are
unable to self-pollinate. The male sterility enables the controlled breeding to obtain hybrid plants showing the heterosis effect. One can distinguish between nuclear male sterility which is due to a mutation in the nuclear genome and cytoplasmic male sterility which is due to a mutation in the mitochondrial genome.

The term "monocotyledonous plant" is intended to comprise any monocotyledonous plant, preferably agricultural, food or feed plants. More preferably, the monocotyledonous plant is selected from the group consisting of Hordeum (barley), Avena (oat), Triticum (wheat), Secale (rye), Oryza (rice), Sorghum (millet), Zea (corn), Panicum, Pennisetum, Setaria and others. Other preferred monocotyledonous plants are plants of the genus Lolium such as *Lolium multiflorum*, *Lolium perenne* and *Lolium hybridum*. Most preferably, the monocotyledonous plant is a *Triticum aestivum* plant.

The first and the second expression cassette can be introduced into a monocotyledonous plant or plant cell by various means, e.g. by transformation of a plant or plant cell with two expression vectors, one of which carrying the first expression cassette and the other one carrying the second expression cassette. Alternatively, one expression vector carrying both the first and second expression cassette may be transformed. Further, a plant carrying the first expression cassette may be crossed with a monocotyledonous plant carrying the second expression cassette and the progeny of this cross which contains both the first and the second expression cassette will then be male sterile. The plants carrying the first or the second expression cassette, respectively, have been produced by transformation with the appropriate expression vectors or are progeny of plants produced by such transformation.

Furthermore, the first and the second expression cassette could also be introduced into a plant cell by cell fusion which may be the fusion of germ cells or of somatic
cells, wherein one cell carries the first expression cassette and another cell carries the second expression cassette.

For the introduction of DNA into a plant host cell there are a number of well-known techniques available and the person skilled in the art can determine the appropriate method in each case without any problem. These techniques include the transformation of plant cells with T-DNA by using *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* as a transformation agent, the fusion of protoplasts, the direct gene transfer of isolated DNA into protoplasts, the electroporation of DNA, the introduction of DNA by means of the biolistic method, as well as other possibilities. Thereby, both stable and transient transformants can be generated.

For the injection and electroporation of DNA into plant cells there are no special requirements per se for the plasmids used. The same applies for direct gene transfer. Simple plasmids, such as pUC derivates may be used. If, however, whole plants are to be regenerated from such transformed cells, the presence of a selectable marker gene is necessary. The person skilled in the art is acquainted with the current selection markers, and will have no problem in selecting an appropriate marker. Standard selection markers are those which mediate resistance to a biocide or an antibiotic such as kanamycin, G418, bleomycin, hygromycin, methotrexate, glyphosate, streptomycin, sulfonyl urea, gentamycin or phosphinotricin and suchlike, to the transformed plant cell.

Dependent upon the method of introduction of the desired gene into the plant cell, other DNA sequences may be required. For example, if the Ti or Ri plasmid is used for the transformation of the plant cell, at least the right flanking region, often however the right and the left flanking region of the T-DNA contained in the Ti or Ri plasmid must be linked as a flanking region with the gene to be introduced.
If agrobacteria are used for the transformations, the DNA to be introduced must be cloned in special plasmids, either in an intermediary or in a binary vector. Based on sequences which are homologous to sequences in the T-DNA, the intermediary vectors can be integrated into the Ti or Ri plasmid of the agrobacteria by homologous recombination. This plasmid also contains the vir region necessary for the transfer of the T-DNA. Intermediary vectors cannot replicate in agrobacteria. By means of a helper plasmid, the intermediary vector can be transferred to \textit{Agrobacterium tumefaciens} (conjugation). Binary vectors can replicate in \textit{E. coli} as well as in agrobacteria. They contain a selection marker gene and a linker or polylinker which are framed by the right and left T-DNA border regions. They can be transformed directly into the agrobacteria (Holsters et al. (1978) Molecular and General Genetics 163: 181-187). The agrobacterium serving as a host cell should contain a plasmid which carries a vir region. The vir region is necessary for the transfer of the T-DNA into the plant cell. T-DNA can also be present. This type of transformed agrobacterium is used for the transformation of plant cells. The use of T-DNA for the transformation of plant cells has been intensively investigated and is described sufficiently in EP 120 515. Both monocotyledonous and dicotyledonous plants or their cells are very accessible to transformation by means of vectors based on agrobacteria (Chan et al. (1993) Plant Mol. Biol. 22: 491-506).

For the transfer of DNA into the plant cell, plant explants can be cultivated specifically for this purpose with \textit{Agrobacterium tumefaciens} or \textit{Agrobacterium rhizogenes}. From the infected plant material (for example, pieces of leaf, stem segments, roots, but also protoplasts or suspension-cultivated plant cells) whole plants can be regenerated in an appropriate medium which can contain antibiotics or biocides for the selection of transformed cells. The regeneration of the plants takes place according to standard regeneration methods and using the common nutrient
solutions. The plants and plant cells obtained in this way can be examined for the presence of the DNA introduced.

The person skilled in the art is acquainted with other possibilities for the introduction of foreign DNA using the biolistic method or by protoplast transformation (see L. Willmitzer (1993) Transgenic Plants in: Biotechnology, A Multi-Volume Comprehensive Treatise (publisher: H.J. Rehm et al.), volume 2, 627-659, VCH Weinheim, Germany).


The transformed cells grow within the plant in the normal way (see also McCormick et al. (1986) Plant Cell Reports 5: 81-84). The resulting plants can be raised in the normal way and be crossed with plants which have the same transformed genetic disposition or other genetic dispositions. The resulting hybrid individuals have the respective phenotypical properties.

Two or more generations should be raised in order to ensure that the phenotypical feature remains stable and is inherited. Seeds should be harvested as well so as to ensure that the respective phenotype or other characteristics are maintained.

Similarly, by using the standard methods, transgenic lines can be determined which are homozygous for the expression cassettes of the present invention and their
phenotypical characteristics with regard to male sterility is investigated and compared with that from hemizygous lines.

Within the scope of the present invention, the term "expression cassette" means a nucleic acid molecule which contains all elements which are necessary for the expression of a gene, i.e. the gene to be expressed under the control of a suitable promoter and optionally further regulatory sequences such as termination sequences. An expression cassette of the present invention may be part of an expression vector which is transferred into a plant cell or may be integrated into the chromosome of a transgenic plant after transformation. The terms "first expression cassette", "second expression cassette", "third expression cassette" and "fourth expression cassette" are only used to distinguish the different expression cassettes comprising different elements, but are not intended to indicate any spatial relationship or order between the expression cassettes, i.e. the second expression cassette may be located 5' of the first expression cassette within an expression vector or a chromosome.

Preferably, the expression vector is selected from the group consisting of plasmids, cosmids, (recombinant) viruses and other vectors known in the field of gene technology, with which nucleic acid molecules can be transferred to plants or plant cells. The term "vector" also comprises so-called minichromosomes which are linear or circular DNA fragments which contain centromer sequences of the respective plant in addition to the transgene. Minichromosomes are stable in the nucleus and are passed on to the daughter cells during cell division. They are transferred by standard methods of transformation. Most preferably, the vector is selected from the group consisting of pBR322, pUC vectors, M13mp vectors or vectors being derived from the Ti plasmid or the Ri plasmid of agrobacteria.
In order to prepare the introduction of foreign genes into higher plants or the cells of the same, a large number of cloning vectors are available which contain a replication signal for *E. coli* and a marker gene for the selection of transformed bacterial cells. Examples of such vectors are pBR322, pUC series, M13mp series, pACYC184, etc. The required sequence can be introduced into the vector at an appropriate restriction site. The plasmid obtained is used for the transformation of *E. coli* cells. Transformed *E. coli* cells are cultivated in an appropriate medium, and finally harvested and lysed. The plasmid is recovered. As an analysis method for characterizing the plasmid DNA obtained, methods such as restriction analyses, gel electrophoreses and other biochemical/molecular biological methods are generally used. Following each manipulation the plasmid DNA can be cleaved and the DNA fragments obtained can be combined with other DNA sequences. Each plasmid DNA sequence can be cloned into the same or other plasmids. Standard cloning methods can be taken from Sambrook et al., 2001 (Molecular cloning: A laboratory manual, 3rd edition, Cold Spring Harbor Laboratory Press).

The term "tapetum-specific promoter" within the meaning of the present invention is understood to mean that a nucleic acid sequence under the control of a tapetum-specific promoter region is expressed in the tapetum of plants. Particularly, a promoter is also tapetum-specific within the meaning of the present invention if the promoter region preferentially leads to the expression of the nucleic acid sequence in the tapetum in comparison to other cell types and leads to a significantly increased expression such as at least two-fold, preferably at least five-fold and particularly preferably at least ten-fold and most preferably at least fifty-fold increased expression in tapetum in comparison to other cell types. The expression of a nucleic acid sequence in different tissues and organs can be determined with *in situ* detection techniques known to the person skilled in the art. For example, a reporter gene such as β-glucuronidase may be expressed under the control of the promoter to be
investigated and the activity of the reporter gene in different organs may be
determined.

The term "tapetum" is known to the expert and is intended to mean the highly
specialized, transient tissue surrounding the (micro-)spores and/or pollen grains
during their development. Supplementary information can be derived from any plant
anatomy or plant physiology book such as Strassburger, Lehrbuch der Botanik, 35.

Suitable tapetum-specific promoters are known to the person skilled in the art and
include the promoter of the rice osgβb gene (Tsuchiya et al. (1995) Plant Cell
Physiol. 36: 487 - 494), the pca55 promoter from corn (WO 92/13956) and the pEl
and pT72 promoters from rice (WO 92/13957). Preferably, the tapetum-specific
promoter is the promoter from the rice osg6B gene. However, the tapetum-specific
promoter of the present invention does not have to be derived from a mono-
cotyledonous plant, but can also be isolated from a dicotyledonous plant, as long as
the promoter is functional in cells of a monocotyledonous plant, i.e. as long as it is
capable of directing tapetum-specific expression of nucleic acid sequences
operatively linked thereto in monocotyledonous plants.

The "protein which provides for male sterility" may be any protein the expression of
which in the tapetum leads to male sterile plants by interfering with the function and
development of pollen. Examples of such genes include RNases, ribosomal inhibitor
36: 487 - 494). Preferably, the protein which provides for male sterility is an RNase.
Examples for RNases are barnase (Mariani et al. (1990) Nature 347: 737 - 741;
Mariani et al. (1992) Nature 357: 384 - 387), RNase T1 from Aspergillus oryzae
Even more preferably, the protein which provides for male sterility is barnase, i.e. ribonuclease from *Bacillus amyloliquefaciens*. The amino acid sequence of barnase from *Bacillus amyloliquefaciens* is shown in SEQ ID No. 2 and the native nucleic acid sequence coding for barnase is depicted in SEQ ID No. 1.

Preferably, the nucleic acid sequences coding for the N- and the C-terminal part of barnase are adapted to the codon usage of monocotyledonous plants, more preferably to the codon usage of *Triticum aestivum*. The nucleic acid sequence coding for the barnase which is adapted to the codon usage of *Triticum aestivum* is shown in SEQ ID No. 49.

To obtain said 5' and 3' part of the nucleotide sequence coding for barnase, said nucleotide sequence is preferably split such that each obtained fragment, upon expression, is incapable of generating a functional barnase protein in the absence of the other fragment. Each barnase fragment contains a sequence portion necessary for the barnase function. The barnase may be split into said fragments in many different ways provided that expression of the male sterility requires all said fragments and binding thereof to each other. One can easily test experimentally whether a fragment generated by splitting the barnase nucleotide sequence at a randomly chosen site is capable of providing male sterility by expressing the fragment in the tapetum of plants and investigating if these plants are able to develop viable pollen. A further assay for testing the functionality of the fragments is described in example 6 of this application. Preferably, the barnase is split into said fragments in the amino acid region between amino acid residues 30 and 40 of the mature protein, i.e. between amino acids 30 and 31, 31 and 32, 32 and 33, 33 and 34, 34 and 35, 35 and 36.
and 37, 38 and 39 or 39 and 40 of the amino acid sequence according to SEQ ID No. 50.

More preferably, the N-terminal part of the barnase comprises 36 amino acids, and the C-terminal part of the barnase comprises 75 amino acids. Alternatively, the N-terminal part of barnase may comprise 35 amino acids if no methionine is added to the N-terminus.

Most preferably, the N-terminal part of the barnase is encoded by a nucleic acid sequence according to SEQ ID No. 3 or 13 and the C-terminal part of the barnase is encoded by a nucleic acid sequence according to SEQ ID No. 5 or 27. The N-terminal part of the barnase preferably has the amino acid sequence according to SEQ ID No. 4 or 14 and the C-terminal part of the barnase preferably has the amino acid sequence according to SEQ ID No. 6 or 28.

However, the invention is also intended to comprise smaller fragments of the barnase protein or "functional barnase fragments", as long as the N- and the C-terminal parts of such fragment restore a functional, enzymatically active protein upon intein-mediated trans-splicing which protein leads to male sterility when expressed in the tapetum of plants. For example, the C-terminal part may lack one or more amino acid residues at its C-terminus and/or the N-terminal part may lack one or more amino acid residues at its N-terminus. Upon intein-mediated trans-splicing a protein smaller than the wild-type protein is formed from the two fragments which protein is functional and enzymatically active. However, also deletions within the N- and/or the C-terminal part of barnase are conceivable as long as they do not affect the ability of the N- and the C-terminal part to restore a functional protein.
The term "operatively linked" is understood to denote that the sequences linking the different nucleic acids used are selected in such a way that the function of the respectively linked nucleic acid segment is maintained. In case, for example, the nucleotide sequence coding for the N- or C-terminal part of the protein which provides for male sterility is to be expressed in a cell, it has to be observed that no sequences which would lead to a termination of the transcription are located between the promoter sequence and the nucleotide sequence coding for the N- or C-terminal part of the protein which provides for male sterility.

The "termination sequences" are sequences which ensure that the transcription or the translation is properly terminated. If the transferred nucleic acids are to be translated, the termination sequences are typically stop codons and corresponding regulatory sequences; if the transferred nucleic acids are only to be transcribed, they are generally poly-A sequences. Preferably, the termination sequence is selected from the octopine synthase terminator and the nopaline synthase terminator (Jones et al. (1992) Transgenic Res. 1: 285 - 297).

The "flexible linker sequence" within the scope of the present invention is a short flexible peptide which is used to bridge the N- and the C-terminal part of the protein which provides for male sterility without serious steric interference. The flexible linker sequence brings the two splice junctions in close proximity and helps to precisely align all reacting groups. Hence, efficient splicing is supported (Chong and Xu (1997) J. Biol. Chem. 272: 15587 - 15590). The most widely used linker designs have sequences consisting essentially of stretches of glycine (G) and serine (S) residues, because hydrophilic amino acids allow hydrogen bonding to the solvent and glycines provide the necessary flexibility. These properties prevent the penetration of the linker peptide into the hydrophobic interface formed in the association of the domains. Furthermore, the linkers are not able to form an ordered secondary
structure. The term "essentially consist of glycine and serine residues" is intended to mean that at least 60% or 65%, preferably 70% or 75%, more preferably 80%, 82%, 84%, 86% or 88%, even more preferably 90%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% of the amino acid residues within the flexible linker sequence are glycine and/or serine residues. Most preferably, all amino acid residues within the flexible linker sequence are glycine and/or serine residues.

The length of the linker sequences should be selected such that the linker peptides are too short to allow pairing of the domains of the same amino acid chain, but favour pairing between domains in two adjacent chains. Thus, the length of the linker sequence may be 1 to 20 amino acids, preferably 2 to 15 amino acids, more preferably 3 to 10 amino acids, even more preferably 4 to 8 amino acids and most preferably 5 amino acids.

Most preferably, the flexible linker sequence has the amino acid sequence GGGGS.

The flexible linker sequence is present in at least one copy, preferably 1 to 5 copies, more preferably 2 to 4 copies and most preferably three copies. If multiple copies of the flexible linker sequence are used every copy is preferably encoded by different codons to avoid interference.

The tapetum-specific promoter used in the second expression cassette may be the same as the one used in the first expression cassette or it may be different from the promoter used in the first expression cassette, preferably the same tapetum-specific promoter is used in the first and the second expression cassette to achieve equimolar expression of the N- and the C-terminal fragments.
"Inteins" are proteins that are able to catalyze trans-splicing events between separate protein fragments as they are able to excise themselves from precursor molecules and ligate the flanking protein sequences in a process termed protein-splicing (Southworth et al. (1998) EMBO J. 17: 918 - 926; Wu et al. (1998) Proc. Natl. Acad. Sci. USA 95: 9226 - 9231; Saleh and Perler (2006) Chem. Rec. 6: 183 - 193; Perler (1998) Cell 92: 1 - 4). At present, more than 400 inteins are known. They are distributed among the genomes of different organisms including eukaryotes, archaeabacteria and eubacteria. An overview of the currently available inteins is given by the homepage http://tools.neb.com/inbase/indcx.php (inbase, the intein database (Perler (2002) Nucleic Acids Res. 30: 383 - 384)).

Preferably, the first intein is selected from the group consisting of DnaB and DnaE, both from the green-blue algae of Synechocystis spec. More preferably, the first intein is DnaB and most preferably, the first intein is encoded by a nucleic acid sequence according to SEQ ID No. 7 or has the amino acid sequence according to SEQ ID No. 8. The nucleic acid sequence coding for the N-terminal part of DnaB is depicted in SEQ ID No. 9 and the nucleic acid sequence coding for the C-terminal part of DnaB is shown in SEQ ID No. 11.

In a preferred embodiment of the present invention, the nucleic acid sequences coding for the N- and the C-terminal part of the protein which provides for male sterility and the nucleic acid sequences coding for the N- and the C-terminal part of the first intein are adapted to the codon usage of monocotyledonous plants.

If, for example, the nucleic acid or the first and the second expression cassette are to be introduced into a wheat plant, the codon usage is adapted to the codon usage of the wheat plant, if the first and the second expression cassette are introduced into a rye plant, the nucleic acid sequences are adapted to the codon usage of the rye plant.
However, the use of a nucleic acid sequence the codon usage of which has been adapted to the codon usage in wheat may also have advantages in monocotyledonous plant species closely related to wheat, such as rye or barley.

Preferably, the nucleic acid sequences coding for the N- and the C-terminal part of the protein which provides for male sterility and the nucleic acid sequences coding for the N- and the C-terminal part of the first intein are adapted to the codon usage of Triticeae, more preferably they are adapted to the codon usage of plants of the genus Triticum and most preferably they are adapted to the codon usage in Triticum aestivum.

The genetic code is redundant, as 20 amino acids are specified by 61 triplet codons. Thus, most of the 20 proteinogenic amino acids are coded by several base triplets (codons). However, the synonymous codons which specify an individual amino acid are not used with the same frequency in a specific organism, but there are preferred codons, which are used frequently, and codons which are used less frequently. Said differences in codon usage may be due to selective evolutionary pressures, and, in particular, to the efficiency of translation. One reason for the lower translation efficiency of rarely occurring codons could be that the corresponding aminoacyl-tRNA pools are depleted and are therefore no longer available for protein synthesis.

Furthermore, different organisms prefer different codons. Thus, for example, the expression of a recombinant DNA originating from a mammalian cell often proceeds only suboptimally in E. coli cells. Therefore, the replacement of infrequently used codons by frequently used codons can enhance expression in some cases.

The DNA sequence of a larger number of genes of many organisms is known and there are tables, from which the frequency of the usage of specific codons in the
respective organism can be taken. With the aid of said tables, protein sequences can be relatively exactly back-translated to form a DNA sequence, which contains the codons preferred in the respective organism for the different amino acids of the protein. Tables for codon usage can, inter alia, be found at the following internet address: http://www.kazusa.or.ip/codon/index.html. There are programs available also for reverse translation of a protein sequence, for example the amino acid sequence of barnase, to form a degenerate DNA sequence, like for instance at http.V/www.ente^ ion.com/bioinformatics/backtranslation.php; or http://www.hgmp.mrc.ac.ub /Software.EMBOSS/Apps/backtranscq.html.

The optimized nucleic acid sequence coding for the N-terminal part of the barnase is shown in SEQ ID No. 13 and the nucleic acid sequence coding for the C-terminal part of the barnase is shown in SEQ ID No. 27. Further, the optimized nucleic acid sequence coding for the N-terminal part of the DnaB intein is shown in SEQ ID No. 17 and the optimized nucleic acid sequence coding for the C-terminal part of the DnaB intein is shown in SEQ ID No. 19.

Preferably, stretches containing exon sequences are inserted between the nucleic acid sequences coding for the C-terminal part of the intein and the C-terminal part of the protein which provides for male sterility and between the N-terminal part of the protein which provides for male sterility and the N-terminal part of the intein. As protein-splicing is dependent on the chemical nature of the splice-site junction amino acids, the insertion of stretches containing exon sequences is supposed to increase the efficiency of protein trans-splicing (Sun et al. (2001) Appl. Environ. Microbiol. 67: 1025 - 1029). Thus, the inserted exon sequences provide the parts of the exteins which are advantageous for the function of the corresponding intein.
Hence, preferably the first expression cassette comprises a nucleic acid sequence coding for an amino acid sequence comprising the amino acid sequence RESG or fragments of said sequence from the DnaB extein and the second expression cassette comprises a nucleic acid sequence coding for an amino acid sequence comprising an amino acid sequence selected from the group consisting of SEEQDHG and SIEQD or fragments of said sequences. Preferably, the second expression cassette comprises a nucleic acid sequence coding for an amino acid sequence comprising the amino acid sequence SIEQD or fragments of said sequence. "Fragments of the RESG extein sequence" may be for example ESG or SG. "Fragments of the SIEQD extein sequence" may be for example IEQD, EQD or QD. "An amino acid sequence comprising the amino acid sequence RESG" or " an amino acid sequence comprising an amino acid sequence selected from the group consisting of SEEQDHG and SIEQD" may be an amino acid sequence with one or more amino acid residues in addition to the RESG, SEEQDHG and SIEQD sequence, respectively. Preferably, the one or more additional amino acid residues are located in the N-terminus of the RESG, SEEQDHG and SIEQD sequence, respectively.

In a further preferred embodiment, the first and the second expression cassette do not comprise any nucleic acid sequences coding for amino acids other than the above mentioned, i.e. the N- and the C-terminal part of the protein which provides for male sterility, the N- and the C-terminal part of the first intein and the nucleic acid sequence coding for at least one copy of the flexible linker sequence as well as optionally sequences from the extein sequence as described above. Particularly, any amino acids that have been inserted into the original vector as a result of the cloning strategies and that are not present in the native amino acid sequence of the protein which provides for male sterility are removed. These are the amino acid residues D and V between the N-terminal part of the protein which provides for male sterility and the N-terminal part of the first intein.
In one embodiment of the present invention, a nucleic acid sequence coding for a protein which is a phenotypical marker is introduced into said plants together with the first and the second expression cassette. "A protein which is a phenotypical marker" is intended to mean a protein the expression of which leads to a property of the plant which allows the identification of plants that have been transformed with the nucleic acid sequence coding for the phenotypical marker. Suitable phenotypical markers are for example proteins conferring herbicide resistance, proteins involved in anthocyanin synthesis or reporter proteins such as glucuronidase, luciferase and green fluorescent protein. Preferably, the phenotypical marker is a protein which confers herbicide resistance.

The protein which confers herbicide resistance can be used for the selection of plants which contain both the N- and the C-terminal part of the protein which provides for male sterility. The protein which confers herbicide resistance may be selected from the group consisting of acetolactate synthase, 5-enolpyruvylshikimate-3-phosphate synthase, phosphinotricin acetyl transferase (BAR), betainaldehyde dehydrogenase (BADH), dihydrofolate reductase (DFRl) and glyphosate oxidoreductase.

Preferably, the protein which confers herbicide resistance is acetolactate synthase which is the first enzyme in the synthesis of branched chain amino acids. This protein can be engineered to confer resistance to sulphonylureas and imidazolinones (Tan et al. (2005) Pest Manag Sci. 61: 246 - 257). Preferably, the acetolactate synthase gene is derived from rice (GeneBank Accession No. AP008208, Oryza sativa; japonica cultivar group; genomic DNA, chromosome 2, position 18335903 - 18337834) and has been engineered according to the teachings in Tan et al. (2005) Pest Manag Sci. 61: 246 - 257 to confer resistance to sulphonylureas and imidazolinones. Most preferably, a mutation of tryptophane to leucin was introduced
on position 548 so that the protein is able to confer herbicide resistance. However, also other mutations on other positions are conceivable, for example a tryptophane to serine mutation on position 548.

Preferably, the nucleic acid sequence coding for the acetolactate synthase is selected from the group consisting of:

a) a nucleic acid sequence according to SEQ ID No. 31 or 51;

b) nucleic acid sequences coding for a protein according to SEQ ID No. 32 or 52 or a functional fragment thereof;

c) nucleic acid sequences hybridizing to a complementary strand of the nucleic acid sequence according to SEQ ID No. 31 or 51 under stringent conditions; and

d) nucleic acid sequences which are at least 50% homologous to the nucleic acid sequence shown in SEQ ID No. 31 or 51.

More preferably, the protein encoded by the nucleic acid sequence comprises a mutation that confers herbicide resistance. Most preferably, the mutation is a tryptophan to leucine mutation on position 548 of the amino acid sequence according to SEQ ID NO. 32 or 52 or a tryptophan to serine mutation on position 548 of the amino acid sequence according to SEQ ID NO. 32 or 52.

According to the present invention, the term "homologous" is generally understood to denote that the nucleic acid or amino acid sequence of a DNA molecule or of a protein is identical to the nucleic acid or amino acid sequences of acetolactate synthase or functionally equivalent parts thereof by at least 50%, preferably by at least 55%, further preferably by at least 60%, also preferably by at least 70%, 80% or 85%, especially preferably by at least 90%, 91%, 92% or 93%, particularly
preferably by at least 94%, 95%, 96% or 97% and most preferably by at least 98% or 99%. Preferably, homology is determined over the entire sequence length of acetolactate synthase.

"Identity of two proteins" is understood to denote the identity of the amino acids over a particular protein region, preferably over the entire protein length, in particular the identity calculated by comparison with the aid of the Lasergene software by DNA Star Inc., Madison, Wisconsin (USA) using the CLUSTAL method (Higgins et al. (1989) Comput. Appl. Biosci. 5 (2): 151).

Nucleic acid molecules are identical if they have identical nucleotides in the same 5' to 3' order.

Thus, homology is preferably calculated over the entire amino acid or nucleic acid sequence region. Besides the programs mentioned above, the person skilled in the art knows further programs based on different algorithms for comparing different sequences. Herein, the algorithms by Needleman and Wunsch, or Smith and Waterman yield particularly reliable results. For said sequence comparisons, for example, the program PileUp (Feng and Doolittle, J. Mol. Evolution. (1987) 25: 351-360; Higgins et al. (1989) CABIOS 5: 151-153) or the programs Gap and Best Fit (Needleman and Wunsch (1970) J. Mol. Biol. 48: 443-453 and Smith and Waterman (1981) Adv. Appl. Math. 2: 482-489), which are contained in the GCG Software Package by the Genetics Computer Group (575 Science Drive, Madison, Wisconsin, USA 5371 1), can also be used.

The Clustal W program, as can be called up at http://www.ebi.ac.uk/clustalw, was used for the sequence alignments conducted within the scope of the present
invention. The parameters of said default homepage remained unaltered for the alignments.

A further object of the present invention are nucleic acid molecules, which hybridize under stringent conditions with, or are substantially complementary to, those nucleic acid molecules coding for acetolactate synthase or functionally equivalent parts thereof. The term "complementarity" describes the capability of a nucleic acid molecule of hybridizing with another nucleic acid molecule due to hydrogen bonds formed between complementary bases. The person skilled in the art is aware of the fact that two nucleic acid molecules do not have to have a 100% complementarity in order to be able to hybridize with each other. Preferably, a nucleic acid sequence, which is supposed to hybridize with another nucleic acid sequence, is complementary to the latter by at least 40%, by at least 50%, by at least 60%, preferably by at least 70%, especially preferably by at least 80%, also especially preferably by at least 90%, particularly preferably by at least 95%, and most preferably by at least 98% or 100%.

Stringent in vitro hybridization conditions are known to the person skilled in the art and can be taken from the literature (see, for example, Sambrook et al., vide supra).

The term "specific hybridization" relates to the fact that a molecule preferably binds to a specific nucleic acid sequence under stringent conditions, provided that said nucleic acid sequence is part of a complex mixture of, for example, DNA or RNA molecules.

Thus, the term "stringent conditions" relates to conditions, under which a nucleic acid sequence preferably binds to a target sequence, but not, or at least in a significantly reduced manner, to other sequences.
Stringent conditions are dependent on the circumstances. Longer sequences hybridize specifically at higher temperatures. In general, stringent conditions are selected in such a way that the hybridization temperature is about 5°C below the melting point ($T_m$) for the specific sequence at a defined ionic strength and a defined pH value. $T_m$ is the temperature (at a defined pH value, a defined ionic strength, and a defined nucleic acid concentration), at which 50% of the molecules, which are complementary to a target sequence, hybridize with said target sequence. Typically, stringent conditions comprise salt concentrations between 0.01 and 1.0 M sodium ions (or ions of another salt) and a pH value between 7.0 and 8.3. The temperature is at least 30°C for short molecules (for example, for those comprising between 10 and 50 nucleotides). In addition, stringent conditions may comprise the addition of destabilizing agents, like for example formamide. Typical hybridization and washing buffers are of the following composition.

15  **Pre-hybridization solution:**

0.5 % SDS

5 x SSC

50 mM NaPO$_4$, pH 6.8

0.1% Na pyrophosphate

5 x Denhardt's Reagent

100 µg/ml salmon sperm

**Hybridization solution:**

Pre-hybridization solution

1 x 10$^6$ cpm/ml probe (5-10 min, 95°C)

**20 x SSC:**

3 M NaCl

0.3 M sodium citrate

ad pH 7 with HCl
50x Denhardt's Reagent: 5 g Ficoll
5 g polyvinyl pyrrolidone
5 g Bovine Serum Albumin
A d 500 ml A. dest.

A typical hybridization procedure is conducted as follows:

Optional: washing the blot 30 min in 1 x SSC / 0.1% SDS at 65°C

Pre-hybridization: at least 2 h at 50-55°C

Hybridization: overnight at 55-60°C

Washing:
5 min 2 x SSC / 0.1% SDS Hybridization temp,
30 min 2 x SSC / 0.1% SDS Hybridization temp,
30 min 1 x SSC / 0.1% SDS Hybridization temp,
45 min 0.2 x SSC / 0.1% SDS 65°C
5 min 0.1 x SSC Room temp.

The term "functional fragment of a protein which confers herbicide resistance" or "functionally equivalent parts of a protein which confers herbicide resistance" is intended to mean that the fragment of the protein is still able to confer herbicide resistance when expressed in plants.

The protein which confers herbicide resistance is expressed under the control of a promoter which is functional in cells of a monocotyledonous plant. As described above with respect to the tapetum-specific promoter, this promoter does not have to
be derived from a monocotyledonous plant as long it can govern the expression of a nucleic acid sequence operatively linked thereto in cells of a monocotyledonous plant. Preferably, constitutive promoters such as the 35S promoter, the actin promoter or the ubiquitin promoter are used, however, other promoters can of course be used which are obtainable from different sources such as plants or plant viruses or fungi and which are suitable for the expression of genes in monocotyledonous plants. The choice of promoter and other regulatory sequences determines the local and temporal expression pattern of the gene of the protein which confers a herbicide resistance. Besides constitutive promoters, also tissue-specific promoters such as the phosphoenolpyruvate promoter or the fructose-1,6-bisphosphatase promoter or inducible promoters are conceivable. Preferably, the protein conferring herbicide resistance is expressed under the control of a rice actinl promoter (McElroy et al. (1990) Plant Cell 2: 163 - 171).

In a preferred embodiment of the present invention, the nucleic acid sequence coding for the protein which is a phenotypical marker is also adapted to the codon usage of monocotyledonous plants. More preferably, the codon usage is adapted to the codon usage of the plant into which the first and the second expression cassette are introduced, preferably to the codon usage of Triticeae, even more preferably to the codon usage of the genus Triticum and most preferably to the codon usage of *Triticum aestivum*.

In one embodiment of the present invention the protein which confers herbicide resistance is also split into two fragments which are assembled to a functional protein by intein-mediated trans-splicing.

It has already been shown that functional acetylactate synthase and functional 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) can be reconstituted from two

Hence, a third and a fourth expression cassette are introduced into a monocotyledonous plant in addition to the first and the second expression cassette. The third expression cassette comprises the following elements in 5’ to 3’ orientation:

- a promoter functional in cells of a monocotyledonous plant;
- operatively linked thereto a nucleic acid sequence coding for an N-terminal part of a protein which confers herbicide resistance;
- a nucleic acid sequence coding for the N-terminal part of the second intein; and
- optionally, operatively linked thereto a terminator sequence functional in said plant cells.

The fourth expression cassette comprises the following elements in 5’ to 3’ orientation:

- a promoter functional in cells of a monocotyledonous plant;
- operatively linked thereto a nucleic acid sequence coding for the C-terminal part of said second intein;
- a nucleic acid sequence coding for a C-terminal part of said protein which confers herbicide resistance; and
- optionally, operatively linked thereto a terminator sequence functional in said plant cells.
Thus, the expression from the third and the fourth expression cassette will restore a herbicide resistance gene upon trans-splicing catalysed by the second intein.

Preferably, the protein which confers herbicide resistance is acetolactate synthase (ALS). More preferably, the acetolactate synthase is split into said fragments in the amino acid region between amino acid residues 400 and 410 of the mature protein according to SEQ ID NO. 32 or 52, i.e. between amino acids 400 and 401, 401 and 402, 402 and 403, 403 and 404, 404 and 405, 405 and 406, 406 and 407, 407 and 408, 408 and 409 or 409 and 410 of the amino acid sequence according to SEQ ID No. 52 or 52. Most preferably, the N-terminal part of the acetolactate synthase has 403 amino acids and the C-terminal part of the acetolactate synthase has 421 amino acids.

Preferably, stretches containing exon sequences are inserted between the nucleic acid sequences coding for the C-terminal part of the second intein and the C-terminal part of the protein which is a phenotypical marker and between the N-terminal part of the protein which is a phenotypical marker and the N-terminal part of the second intein. These exon sequences provide the parts of the extein which are advantageous for the function of the intein. Hence, preferably the third expression cassette comprises a nucleic acid sequence coding for an amino acid sequence comprising the amino acid sequence DVKFAEY or fragments of said sequence from the DnaE extein and the second expression cassette comprises a nucleic acid sequence coding for an amino acid sequence comprising the amino acid sequence CFNHG or fragments of said sequence. "Fragments of the DVKFAEY extein sequences" may be for example VKFAEY, KFAEY, FAEY, AEY or EY. "Fragments of the CFNHG extein sequences" may be for example FNHG, NHG and HG. "An amino acid sequence comprising the amino acid sequence DVKFAEY " or " an amino acid sequence comprising the amino acid sequence CFNHG" may be an amino acid with one or
more amino acid residues in addition to the DVKFAEY or CFNHG sequence, respectively. Preferably, the one or more amino acid residues in addition are located on the N-terminus of the DVKFAEY or CFNHG sequence, respectively.

Preferably, the fourth expression cassette further comprises a nucleic acid sequence coding for at least one copy of a flexible linker sequence which nucleic acid sequence is located between the nucleic acid sequence coding for the C-terminal part of said second intein and the nucleic acid sequence coding for a C-terminal part of said protein which is a phenotypical marker. The flexible linker is defined as above with respect to the second expression cassette.

Some proteins which are phenotypical markers, such as acetolactate synthase, are expressed in the cytosol and transported to the chloroplast via an N-terminal signal sequence. Hence, for targeting the C-terminal part of said proteins fused to the C-terminal part of the second intein to the chloroplast, an artificial chloroplast targeting sequence has to be fused to this fusion protein. Preferably, the artificial chloroplast targeting sequence has the amino acid sequence MASSMLSSAAAVTASAAQASMVAPFTGLKSAASFPVTRKQNNLDITSIAS NGGRVQCA or is a functional fragment thereof which is still capable of directing the transport of proteins to the chloroplast.

The second intein the N- and the C-terminal part of which are present in the third and the fourth expression cassette, respectively, is other than the first intein the N- and the C-terminal part of which is present in the first and second expression cassette so that the first and the second intein do not cross-react with each other and do not lead to mis-spliced products due to the universal nature of interaction between the intein parts. This means that for example if a DnaB is used as the first intein, DnaE is used as the second intein and vice versa.
In one embodiment of the present invention the first, the second, the third and the fourth expression cassette are all located on the same expression vector. Preferably, the expression vector comprises a first part comprising the first expression cassette and the third or fourth expression cassette, and a second part comprising the second expression cassette and the third (if the first part comprises the fourth expression cassette) or fourth expression cassette (if the first part comprises the third expression cassette). More preferably, both the first part and the second part are flanked by recombinase recognition sites.

A "recombinase" is an enzyme which catalyzes a recombination process. In particular, the recombinase recognizes certain nucleic acid sequences, the so called "recombinase recognition sites", and leads to a recombination of these sites. One recombinase which could be used for catalyzing a recombination is the Streptomyces phage PhiC31 integrase which catalyzes the irreversible recombination between attB and attP recognition sites. The site-specific recombination between these sites leads to the deletion of sequences located between the participating recombinase recognition sites.

In an embodiment of the present invention, an expression vector comprising a first part comprising the first expression cassette and the third or fourth expression cassette, and a second part comprising the second expression cassette and the third (if the first part comprises the fourth expression cassette) or fourth expression cassette (if the first part comprises the third expression cassette), wherein both the first part and the second part are flanked by recombinase recognition sites is first introduced into a plant to form a pro-locus. Upon crossing said plants carrying the pro-locus with a plant expressing a site-specific recombinase such as Streptomyces phage PhiC31, the recombinase will catalyze the irreversible site-specific recombination
between the recombinase recognition sites and derivative loci will be obtained which contain either the first or the second part. For example, a first derivative locus contains a nucleic acid sequence coding for the N-terminal part of the protein which provides for male sterility and a nucleic acid sequence coding for the N-terminal part of a protein which confers herbicide resistance and a second derivative locus contains a nucleic acid sequence coding for the C-terminal part of the protein which provides for male sterility and the C-terminal part of the protein which confers herbicide resistance. Furthermore, the first and the second derivative locus are located on identical loci on homologous chromosomes, i.e. they are "linked in repulsion". For example in diploid plants, this means that one derivative locus is located on locus X on chromosome Y and the other derivative locus is located on locus X' on the other chromosome Y'. However, the processes of the present invention may not only be applied to diploid plants, but also to polyploid plants. "Identical loci" are loci between which no recombination is possible. A scheme for the production of the pro-locus and the derivative loci is shown in Figure 8.

In another embodiment of the present invention, the protein which confers herbicide resistance is not split. The first part of the expression vector then comprises either the nucleic acid sequence coding for the N-terminal part of the protein which provides for male sterility alone or together with the nucleic acid sequence coding for the full-length protein which confers herbicide resistance. The second part of the expression vector comprises either the nucleic acid sequence coding for the C-terminal part of the protein which provides for male sterility alone or together with the nucleic acid sequence coding for the full-length protein which confers herbicide resistance (if the latter nucleic acid sequence is not present in the first part). Upon recombination, two derivative loci will be formed. One of these derivative loci will then contain the nucleic acid sequence coding for the full-length protein which confers herbicide resistance together with either the nucleic acid sequence coding for the N-terminal
part of the protein which provides for male sterility or the nucleic acid sequence coding for the C-terminal part of the protein which provides for male sterility, while the other derivative locus only contains either the nucleic acid sequence coding for the N-terminal part of the protein which provides for male sterility or the nucleic acid sequence coding for the C-terminal part of the protein which provides for male sterility.

In one embodiment of the present invention the male sterile plants are selected by applying a suitable herbicide, i.e. a herbicide against which the plants may be resistant by expression of the protein which confers herbicide resistance, to said plants. Only the plants which contain a functional protein which confers herbicide resistance to said plants will survive the selection. If the protein which confers herbicide resistance is acetolactate synthase, the herbicide used for selection of the plants may be selected from the groups of sulphonylureas and imidazolinones, such as primisulphuronmethyl or imazethapyr.

The present invention further relates to a method of producing monocotyledonous hybrid plants, comprising the steps of:

a) producing a male sterile monocotyledonous plant by a method described above; and

b) crossing the male sterile monocotyledonous plant of step a) with a male fertile monocotyledonous plant.

"Hybrid plants" are plants which are the progeny of two genetically unidentical parents. Hybrid plants are produced by cross-pollination of genetically different parental lines. As the plants of the present invention are male sterile, the self-pollination of these plants is avoided and it is possible to perform directed crosses
with a selected male fertile parent. The term "male fertile parent" means that this parent is able to produce functional pollen which can be used to pollinate the female sexual organs of the male sterile plant.

In the process of the present invention, the hybrid seeds produced may be 100% fertile, as either only the N-terminal fragment or only the C-terminal fragment of the protein which provides for male sterility is expressed in the progeny. The male fertile plant used for producing the hybrid plants should not contain a fragment of a nucleotide sequence encoding the protein which provides for male sterility and, optionally, not contain a fragment of a nucleotide sequence encoding the protein which confers herbicide resistance.

The hybrid seed growing on the hybrid plants may then be harvested.

"Transgenic plants" and "transgenic plant cells" can be any monocotyledonous plant or plant cell, preferably agricultural plants or cells from agricultural plants, into which a nucleic acid molecule or at least one expression cassette has been introduced. These plants can be produced by any of the methods which have been described herein. The invention is further directed to transgenic parts of this plant such as leaves and blossoms, transgenic propagation material such as protoplasts, calli, fruit, seeds, tubers, root stocks, germs, pollen, cuttings and transgenic progeny of the plant.

The plant cells according to the invention include differentiated and undifferentiated plant cells including protoplasts which were produced by the method according to the invention and which have integrated the expression cassettes of the present invention into the plant genome, or have received these as autonomously replicating molecules.
Of course, plant cells which contain the nucleic acid molecules according to the invention and plant cells (including protoplasts, calli, suspension cultures and suchlike) can further be cultivated.
Brief Description of the Drawings

Figure 1: Structure of the T-DNA constructs

All vectors were cloned into pBIN19-based binary vectors between the T-DNA left and right borders (LB and RB).

Bar-N, Bar-C, gene fragments from the *Bacillus amyloliquifaciens* barnase gene coding for the N- and C-terminal fragment of barnase, respectively (synthetic sequence with codon usage adapted for wheat; except in case of pICH13688 that carries a native sequence from *Bacillus amyloliquifaciens*); ALS-N, ALS-C, gene fragments coding for the N- and C-terminal fragment of a mutated version of *Oryza sativa* ALS or of *Arabidopsis thaliana* ALS (in case of pICH13688 only); IntN, IntC, gene fragments from the from the DnaE and DnaB genes of *Synechocystis* sp. coding for N- and C-terminal intein sequences (DnaB sequences fused to barnase are synthetic with codon usage adapted for wheat except in case of pICH13688); Pact, rice actin 1 promoter; Ptap, tapetum-specific promoter osgβB from rice; Tnos, nopaline synthase terminator; Toes, octopine synthase terminator; CT, artificial chloroplast-targeting signal; Pubi, maize ubiquitin promoter, HptII, hygromycin phosphotransferase gene; Pspm, maize spm promoter; linker, nucleic acid sequence coding for a flexible (GGGGS)_n linker cloned in frame to the barnase C-fragment.

Figure 2: Analysis of the \(\pi\alpha\tau\)-splicing reaction efficiency using a transient assay in *Nicotiana benthamiana*

Two complementary barnase constructs (N- and C-terminal barnase fragments fused to the N- and C-terminal fragments of the DnaB intein, respectively) were infiltrated into leafs of *N. benthamiana*. The N-terminal gene fragments are expressed from a
35S promoter (pICH12795) whereas the C-terminal gene fragments are expressed from the weaker *Arabidopsis thaliana* Act2 promoter. The cytotoxicity increases with the size of the flexible GGGGS linkers introduced between the nucleic acid sequences coding for the C-terminal fragment of the DnaB intein and the C-terminal fragment of barnase (a-c). The structure of the infiltrated constructs is depicted in d. Construct pICH12795 was supplied by Icon Genetics.

Figure 3: Analysis of the prorø-splicing reaction efficiency using a transient assay in *Nicotiana benthamiana*

Mixtures of agrobacteria carrying different plasmids were infiltrated into leaves of 7-8 weeks old *N. benthamiana* plants. At the infiltration point, the epidermis was carefully scratched in order to facilitate efficient delivery of the agrobacterium solution into the plant tissue. After infiltration, the plants were grown under greenhouse conditions.

a-d, effects of linker-insertions between the domains of the mature protein, e, f, effects of mutating key trans-splicing residues. The leaf morphology is depicted 4 (b), 10 (c) or 12 (a, d, e, f) days after infiltration. The structure of the infiltrated constructs is shown in g. ACT2, *Arabidopsis* actin 2 promoter; 35S, Cauliflower Mosaik Virus 35S promoter.

Figure 4: Phenotype of male-sterile wheat plants

Morphology of wheat spikes from a male fertile control plant (that was generated by *in vitro* culture methods, e.g. that grew up under identical conditions like transgenic plants) and a transgenic male sterile plant containing T-DNA pICH27371 are shown in a and b, respectively. The spike of the transgenic plant displays the typical "open floret" phenotype and contains no seed. The deformed pollen produced by plants
carrying a functional barnase protein can be clearly distinguished from the pollen of wildtype plants in the Alexander's vitality tests (c, d) and in SEM analysis (e-h). The left-hand column shows the phenotype of the control plant and the right-hand column shows the phenotype of the male sterile plants.

Figure 5: Efficiency of differently modified split-barnase-systems in primary transformants (To)

The proportion of T₀ plants displaying complete sterility (black sections of the bars) is compared to the proportion of plants carrying both sterile and fertile flowers (gray) and such plants that are completely fertile (white) for all plasmids investigated. Total number of T₀ individuals investigated is given in parentheses.

Figure 6: Molecular characterization of wheat lines transgenic for pICH27371

Progeny from T₀ plants pICH27371-1219 (b, c) and pICH27371-1058 (d) segregates into male-sterile (S) and fertile (F) individuals. Total DNA from plants pICH27371-1219 and pICH27371-1058 was digested with EcoRI and BamHI, respectively, and a Southern Blot analysis was performed. The EcoRI digested DNA was hybridized with a probe homologous to the barnase-N (b, d) and the BamHI digested DNA was hybridized with the barnase-C sequence (c), respectively. Regions of homology are shown by rectangles in the schematic illustration (a). For pICH27371-1058, the presence of the N- or C-terminal barnase gene fragment as proven by PCR is indicated. Wt, wildtype control plant; To, primary transformed plant.

Figure 7: Maps of the plasmids used for transformation

a) pICH13688
Figure 8: Selection of plants carrying pICH25881 (split ALS).

The picture depicts a T₂ plant that is selected after two preceding selections of the Ti and the To ancestors (proving that the phenotype is stably inherited over three generations). Plants carrying the plasmid pICH27371 with the non-split ALS gene were used as a positive control for the selection. Selection conditions are described in the examples.

Figure 9: Split ALS from rice confers herbicide resistance in wheat

F₂ plants of a line carrying pICH25881 (split ALS) derived from selection on medium containing 0.5 µM PSM. Plants were grown in phytochambers under a regime of 16 h light, 24°C and 8 h dark, 16°C. Herbicide resistant plants displayed normal root development (6-10). F₂ plants that lost the T-DNA due to segregation have stunted roots and did not survive on herbicide containing medium (1-5). pICH25301 (F₂), pICH21721 (Fi), control plants carrying a continuous ALS transgene. Plants are shown 21 days after transferring plantlets on the selective medium. Wt, wildtype control plant.

Figure 10: Strategy for the maintenance of the male sterility and for the production of hybrid seed
a) An expression construct containing the first, second, third and fourth expression cassettes of the present invention (corresponding to plasmid pICH25881) is transformed into plant cells to form a prolocus. Upon crossing with a plant expressing recombinase, a derivatization of the prolocus can occur in the progeny plant, thus forming two isoloci (Al, A2) on identical loci on homologous chromosomes in different progeny plants.

b) Production of F3 plants containing either the N-terminal part (Al) or the C-terminal part (A2) of barnase using a site-specific recombinase (integrase) and producing male sterile F4 plants using the F3 plants

c) Use of the male sterile plants to produce fertile hybrid seed by crossing with a male fertile parent

d) Maintenance of the male sterile line by crossing it with a fertile plant containing only the N-terminal part of barnase (Al)

Figure 11: Molecular proof of the PhiC31-induced recombination process

a) Schematic illustration of the positions of primer binding sites (symbolized by arrows) before (pICH27371) and after recombination (pICH-27371-N, pICH27371-C).

b) PCRs were carried out on pICH27371 plasmid-DNA (P) and on total DNA from plants carrying pICH27371 that have been crossed with pICH13130. Recombination of the N-terminus (i.e. between attPi and attB) was detected with Primer Rec 1, Rec 2 in plants 5 and 6; recombination of the C-terminus (i.e. between attP2 and attB) was detected with Primer Rec 3, Rec 4 in plants 1-4.
Figure 12: Sequence analysis of PCR products

In total, 25 PCR-Products resulting from the PCR Rec 1 - Rec 2 and 25 PCR products resulting from the PCR Rec 3 - Rec 4 were analyzed. Without exception, all sequences represented the expected recombination product containing the sequences attL or attR, respectively. The diagram includes two original abi sequencing files (sequencing reaction performed by the IPK-Gatersleben PGRC Sequencing Service).

EXAMPLES

1) Plant material and growth conditions

Spring wheat \((Triticum aestivum\ L.,\ cultivar\ "Bobwhite")\) was used throughout this study. For standard breeding, plants were grown under greenhouse conditions with 16 h of light at 20\(^\circ\) C and 8 h of darkness at 16\(^\circ\) C. For assaying the temperature sensitivity of male-sterile phenotypes, plants were grown in phytochambers at 35\(^\circ\) C for 16 h in light and at 20\(^\circ\) C for 8h in the dark. For DNA isolation, plant tissues were harvested, frozen in liquid nitrogen, and stored at -80\(^\circ\)C.

2) Standard molecular biology techniques

Standard molecular biology procedures were performed as described in Sambrook et al, 2001 (Molecular cloning: A laboratory manual, 3\(^{rd}\) edition, Cold Spring Harbor Laboratory Press).
3) Construction of the vector plasmids

All vectors used in this study (Fig. 1) are pBIN19-based binary vectors and were constructed using standard recombinant DNA methods. The vectors are derivatives of plasmid pICH13688 (Fig. 1). The structure and construction of pICH13688 is described in Gils et al. (2008) Plant Biotechnol J. 6(3): 226-35. The T-DNA sequence of pICH13688 is deposited in the EMBL/GENBANK database under the accession number AM888351.

Barnase system For inducing male sterility, the vectors contain split barnase genes which are fused to N- and C-terminal fragments of the Ssp DnaB intein. Protein splicing is depending on the chemical nature of the splice site junction amino acids. Hence, the insertion of stretches containing exon sequences is supposed to increase the efficiency of protein παρα-splicing (Sun et al. (2001) Appl. Environ. Microbiol. 67: 1025-1029). The amino acid sequence SIEQD was inserted between DnaB IntC and Bar-C. Six of the seven amino acids belong to the adjacent extein sequence of the DnaB gene. Similar, the amino acid sequence RESG from the DnaB extein was introduced between DnaB IntN and Bar-N. Additionally, a methionine starting codon was added at the 5’ terminus of the DnaB IntC sequence.

Both barnase-intein fusions were transcribed from a tapetum-specific rice promoter (Tsuchiya et al. (1995) Plant Cell Physiol. 36: 487-494).

ALS-System Plasmid pICH21721 contains a rice ALS gene (Gene Bank accession number AP008208. Oryza sativa; japonica cultivar-group; genomic DNA, chromosome 2, position 18335903-18337834). At position 548, the mutation Trp>Leu was introduced according to Tan et al. (2005) Pest Manage. Sci. 61: 246-257. The ALS gene is expressed under control of the constitutive rice Actin 1

Vector pICH25881 contains two fragments of the rice ALS gene that are expressed individually from a rice actin promoter. ALS-N (see Fig. 1) contains 403 amino acids of the N-terminal end of the mature ALS protein (GenBank accession no AP008208, position 18335903-18337111) whereas ALS-C contains the residues 404-644 (GenBank accession no AP008208, position 18337112-18337834). For linking both parts of the ALS by rαø-splicing, Ssp DnaE intein fragments were cloned in frame to the ALS segments, including extein sequence stretches, as described in (Gils et al. (2008) Plant Biotechnol J. 6(3):226-35). For targeting the C-terminal ALS-intein fusion protein to the chloroplast, an artificial chloroplast targeting sequence (MASSMLSSAAVVATRASAAQASMVAPFTGLKSAAASFPVTRKQNNLDITSIA SNGGRVQCA) was fused to the segment. Like the complete ALS gene in pICH27371, the expression of both ALS protein segments are controlled by a rice Actin 1 promoter and terminated by nos terminators.

In case of vectors pICH24581, pICH25301 and pICH27371, the barnase-intein fusions are encoded by synthetic sequences in which the codon usage was adapted to the codon bias of Triticum aestivum genes (chemically synthesized by GENART, Regensburg, Germany). Additionally, amino acids that have been inserted into the original vector as a result of cloning strategies and are not present in the native barnase sequence (D and V between Bar-N and DnaB IntN in case of plasmid pICH13688) were removed. Furthermore, vectors pICH25301 and pICH27371 contain flexible linker sequences that are introduced in frame between the dnaB InteinC and the Barnase-C fragment. The linkers should not exhibit a propensity for ordered secondary structure or any tendency to interfere with domain folding. Thus,
the sequence Gly-Gly-Gly-Gly-Ser was selected to bridge the 5'- and 3' barnase fragments either in a double [(Gly-Gly-Gly-Gly-SeT₂); pICH25301] or triple [(Gly-Gly-Gly-Gly-Ser)₃; pICH27371] configuration. In order to avoid repetitiveness of the sequences, a different codon usage was used for each of the GGGGS units.

Vectors for transient barnase expression assays Vector pHW21 (N-terminal vector for transient assays) contains a fusion of Bar-N and DnaB IntN, as described for the vectors pICH24581, pICH25301, pICH27371 and pICH25881 (Figure 1). The expression of the barnase- intein fusion was controlled by the cauliflower mosaic virus 35S promoter and an octopine synthase terminator. Using PCR site-directed mutagenesis, the N-terminal amino acid of DnaB Int (Cl) was changed from cysteine in pHW21 to alanine, resulting in vector pHW211. The C-terminal vectors for transient assays contain a fusion of DnaB IntC and Bar-C. The intein-barnase fusion sequence is cloned between an Arabidopsis actin2 promoter and a nopaline synthase terminator. In pICH24591, pICH24601, pICH24612 and pHW231, flexible linkers composed of one to three GGGGS stretches were cloned in frame between the extein stretch SEEQD and Bar-C. By PCR site-directed mutagenesis, the C-terminal amino acid of DnaB Int (Asn154) and the following residue of the C-terminal extein stretch (Ser+1) were both changed to alanine in the case of pICH24431 (resulting in pHW221, Figure 3g) and pICH24612 (resulting in pHW231, Figure 3g).

4) Genetic transformation of wheat plants via biolistic particle bombardment

Callus culture maintenance

Immature seeds of wheat were surface-sterilized by immersing them in 70 % ethanol for 3 min. The procedure was followed by incubation in 2.5 % sodium hypochlorite solution, including 0.01 % SDS, with shaking at 125 rpm for 7 min and subsequently by three washing steps in sterile distilled water. Immature embryos (1.0-2.5 mm in
length, semitransparent) were excised aseptically and placed, with scutellum-side up, on MS culture medium (Duchefa, M0222; (Murashige and Skoog (1962) Physiol Plant 15(3): 473-497), containing 30 g/l sucrose, 2 mg/l 2,4-D (2,4-dichlorophenoxy-acetic acid) and 0.25 % phytagel for solidification. Embryos that develop compact nodular calli were selected using a stereomicroscope and used for bombardment 14-21 days after isolation. The cultures were kept in the dark at 25° C.

Microprojectile bombardment of immature embryos

The gold coating procedure was done according to Sanford et al. (1993) Methods Enzymol. 217: 483-503 and following the original protocol of Bio-Rad (Munich, Germany).

Standard procedure

For particle coating, 50 µl of gold suspension (0.6 Micron gold in 50% glycerol, 60 mg/ml) was mixed with 10 µl (1 µg/µl) plasmid-DNA, 50 µl 2.5 M CaCl₂ and 20 µl 0.1 M spermidine. The mixture was shaken for 2 min, followed by incubation at room temperature for 30 min, brief centrifugation and washing with 70% and 99.5% ethanol. Finally, the gold particle pellet was suspended in 60 µl of 99.5% ethanol. For one bombardment procedure, 6 µl of the suspension was used. All manipulations were done at room temperature.

Microprojectile bombardment was performed utilizing the Biolistic PDS-1000/He Particle Delivery System (Bio-Rad, Munich, Germany). Prior to the bombardment, the immature embryos were pre-treated for four hours on MS medium supplemented with 100 g/l sucrose.

Approximately 50 Embryos were placed in the centre of a plate to form a circle with a diameter of about 10 mm. The shooting was carried out using a helium pressure of 900 psi, with 15 mm distance from a macrocarrier launch point to the stopping screen and 60 mm distance from the stopping screen to the target tissue. The distance between rupture disk and launch point of the macrocarrier was 12 mm. Finally, 16
hours after treatment, the calli were transferred to MS medium containing 60 g/l sucrose and grown in dark conditions for one week at 25°C.

**Selection and regeneration**

For obtaining plants transgenic for all vectors except pICH13688, callus selection was carried out in vitro on medium containing primisulphuronmethyl (PSM) and imazethapyr (IMA), both belonging to the class of pyrimidinylsulphonylurea herbicide. Concentrations: 0.5 µM PSM + 1.0 µM IMA (first selection); 1.0 µM PSM + 1.0 µM IMA (second selection), 0.5 µM PSM + 1.5 µM IMA (subsequent selections) Selection of plants carrying pICH13688 was carried out by callus selection on medium containing 150 mg/l hygromycin B.

The cultures were kept in the dark at 22°C. After 5-6 successive callus selection steps (total time: 4-6 months) callus tissue was subcultured in MS regeneration medium supplemented with 1 mg/l kinetin, 7 mg/l zeatin. Regenerating plantlets were transferred to jars with half strength hormone-free MS medium containing 0.5 µM PSM + 1.0 µM IMA or 50 mg/l hygromycin B respectively. Fully developed plantlets were acclimated for 7-10 days at room temperature in liquid medium containing four-fold diluted MS salts. Plants with developed roots were transferred into soil and grown under greenhouse conditions to maturity.

For assaying the herbicide resistance of progeny plants, embryos were isolated from seeds and selected in vitro.

**Selection of transgenic wheat plants by a split ALS gene system**

By callus selection on PSM and IMA, lines carrying the vector pICH25881 displayed an herbicide tolerant phenotype that was indistinguishable from the phenotype of control plants (Figure 9). The herbicide resistance was stably inherited over three generations (To-F3) and the plants displayed normal vegetative development. From these results it can be deduced that, despite the low overall efficiency, a functional
ALS protein can be produced by the assembly of two precursor proteins and that the chosen barnase-intein junction site is principally suitable for the split ALS system.

5) Analysis of wheat transformants carrying T-DNA pICH13688

In case of *A. thaliana*, transformation of pICH13688 led to a high frequency of male-sterile plants among primary transformants. Therefore, initial wheat transformation experiments were carried out using this vector. To be able to test the system in wheat, and since it was uncertain whether ALS selection would work in this species (given that the ALS of pICH13688 is derived from an *Arabidopsis* gene), the plants were selected on hygromycin. After growing the primary transformants to maturity, the function of the split-barnase system was tested by pollen activity assays. All 83 plants carrying the vector pICH13688 displayed full fertile pollen that was non distinguishable from wildtype control plants (see Fig. 5) and produced seed. Therefore, it can be concluded that the split-barnase system of pICH13688 that is inducing pollen abortion in dicotyledonous species is not efficient in wheat.

6) Delivery of *Agrobacterium* into plants for transient assays

Prior to transforming new vector constructs in wheat plants, the effect of (GGGGS)$_n$ linker sequences was tested by a rapid transient assay based on the agroinfiltration of *Nicotiana benthamiana* leaves.

T-DNA-vectors were transformed into *Agrobacterium* strain GV3101:pMP90. The infiltration of *Agrobacterium* into *N. benthamiana* plants was performed according to a protocol described in Marillonnet et al. (2004) Proc. Natl. Acad. Sci. USA 101: 6852-6857.
Flexible glycine/serine linker sequences of varying length \([\text{GGGGS}_1]; \text{(GGGGS)}^\wedge; \text{(GGGGS)}_3\] were introduced into control vectors in such a way that they connect the TN- and C-terminal domains of the mature barnase-protein after ligation via trans-splicing (see Figs. 2 and 3). Combinations of T-DNAs that carry the N- and C-terminal parts of the barnase gene (under the control of constitutively expressed promoters) were introduced into leaves of *N. benthamiana* by syringe infiltration using *Agrobacterium* (Figs. 2 and 3). Delivered alone, N- or C-terminal vectors do not exhibit cytotoxicity (Gils et al. (2008) Plant Biotechnol J. 6(3):226-35).

However, when vectors containing the N- and C-terminal fragments were delivered by mixing the *Agrobacterium* suspensions, lesions are obtained in the affected tissue. The extent of cytotoxicity correlates with the length of the flexible linkers that are introduced in the C-terminal vectors. Infiltration of pICH 12795 in combination with pICH 14601 (no linker), results in minor effects. When constructs containing linkers were delivered into the plant tissue, first lesions occurred 4-5 days after infiltration. Usually, the tissue was completely destroyed 14 days after infiltration, depending on the linker size.

Furthermore, by site-directed mutagenesis, it could be demonstrated that that a trans-splicing mechanism is responsible for an efficient protein assembly in the split DnaB intein-system. In order to determine whether protein splicing or IPC (intein mediated protein splicing via affinity domains) results in the reconstitution of barnase activity, key residues of the split DnaB Int system were exchanged without manipulating the affinity domain. For the majority of inteins, three conserved residues are essential for trans-splicing (serine, threonine or cysteine at the intein N-terminus; asparagine or glutamine at the intein C-terminus; serine, threonine or cysteine as the first extein residue following the C-terminal splice site; Perler, *Cell*, 92, 1-4, 2002). By using the *Synechocystis* sp. DnaE intein to complement inactive EPSPS precursor proteins
in *E. coli*, Chen et al. (*Gene*, 263, 39-48., 2001) demonstrated that the exchange of the three key splicing residues to alanine blocks *rrn0*-splicing, but, nevertheless, association of N- and C-terminal splicing domains led to a robust phenotype. However, in the present transient assays, the delivery of vectors carrying *DnaB* Int mutants did not result in apparent symptoms (see Figures 2 e,f). Hence, it can be concluded that the system of the invention requires intein-mediated splicing.

7) Molecular analysis of transformants

Isolation of total plant DNA

For DNA isolation, 300 mg leaf material of young plants was shock-frozen. Homogenization was carried out using a TissueLyser® from Qiagen (Hilden, Germany). Total plant DNA was isolated following a modified protocol from Dellaporta et al. (1983) Plant. Mol. Biol, Pep. 1: 19-29.

Primers used a) for detection of the N-terminal barnase-Intein fusion were *Barnase-N FW* (GCATCGATATGGCCCAAGTG); *dnaB Intein-N REV* (GAGCTGGAGGGAGGAGGATTCG) b) for amplification of the C-terminal barnase-intein fusion gene sequence were *Barnase-C REV* (GATCTTG GGTGAAGTCTGTAG); *dnaB Intein-C FW* (GGGACTCCATCGTGTCCATCC) c) for detection of the N-terminal ALS gene sequence were *ALS-N FW* (GTCAGCGACGTGTCGCTAC) and *ALS-N-REV* (GTCTCTCA ATCAAAGGAACG) and d) of the C-terminal ALS gene sequence were *RiceALS-C FW* (GCAATATGCCATTACAGGTG) *RiceALS-C REV* (CACG GACTGCAGGAATATTG). Finally, for amplification of an ALS-fragment that covers both N- and C-terminal parts of the ALS sequence, *ALS-N-3 FW*
(GATTCTCTATGTCGGTGGTG) and ALS-N-5 'REV (GCGACAGAATTGCTTGAGCAG) were used.

PCR analyses were performed in a thermocycler (DNA Engine™ PTC-0200, Bio-Rad, Munich, Germany). Amplification was carried out for 35 cycles (94°C for 1 min; 55°C for 1 min; 72°C for 1-2 min).

8) Scanning electron microscopy (SEM)

Probes of T. aestivum were dehydrated in an ethanol series, followed by critical point drying in a Bal-Tec critical point dryer (Bal-Tec AG, Balzers, Switzerland). Dried specimens were attached onto carbon coated aluminium sample blocks and coated in an Edwards S150B sputter coater (Edwards High Vacuum Inc., Crowlesy, West Sussex, UK). Probes were examined in a Hitachi S4100 SEM (Hisco Europe, Ratingen, Germany) at 5 kV acceleration voltage. Digital recordings were made and saved as tif-files.

9) Fertility assays

To check for the viability of pollen, anthers of control wild-type and sterile plants were stained with Alexander stain (Alexander (1969) Stain Technol. 44: 117-122). The assays were transformed prior to 2-3 days before anthesis.
10) Cross pollinations

Pollination of male-sterile plants was performed by tearing anthers of untransformed bobwhite plants with tweezers just before anthesis and placing one anther into the closed flower of the male-sterile crossing partner.

11) Expression of N- and C-terminal barnase fragments from T-DNA pICH24581, 25301 and 27371 confers pollen ablation in wheat plants

From the results obtained with the transient tests it could be concluded that the cytotoxicity of the split-barnase system can be significantly increased by introducing flexible amino acid stretches. Therefore, T-DNA vectors containing codon optimized barnase-intcin fusions and flexible linkers \((\text{GGGS})_2\) and \((\text{GGGS})_4\) (Fig. 1) were constructed and delivered into wheat plants by biolistic bombardment. 1385 primary transformants (To) were assayed for pollen fertility. Pollen of at least three spikes of each To-plant was stained in double test series according to (Alexander (1969) Stain Technol. 44: 117-122) (Fig. 4). In contrast to pICH13688, it was possible to generate male sterility by transforming pICH24581, pICH25301 or pICH27371. Thus, it could be demonstrated that the constructs are functional with regard to the inlein-mediated complementation of the barnase peptide fragments.

Sterile pollen can be identified unambiguously by a transparent phenotype in vitality-staining assays (Fig 4 d) and a degenerated phenotype as been detected in scanning electron microscopy analyses (Fig. 4 f, h). The amount of pollen produced by male sterile plants was not reduced compared to that of wildtype plants that were generated by in vitro culture. Expression of barnase does not influence the vegetative phenotype in regard to germination frequency, plant height, leaf size, flowering time and tillering. Typically, to allow cross-pollination, florets of male-sterile plants open as a consequence of inhibited self-fertilization (Fig. 4 b).
From these observations it is concluded that ταφ-splicing of the barnase fragments fosters an efficient ligation of the protein fragments with the concomitant correct folding of the mature protein, that the barnase protein is catalytically functional, and that the activity is limited exclusively to the tapetum.

Strikingly, the frequency of male sterile phenotypes varied among the classes of primary transformed plants, depending on the T-DNA that was introduced. Here, the results obtained from the transient assays (Figures 2 and 3) were confirmed (Fig. 5). The highest frequency (45%) of plants displaying complete or partial male sterility is achieved by transforming pICH27371 which contains codon-optimized barnase-intein fusions and a triple GGGGS linker, followed by pICH25301 (double linker GGGGS); optimized barnase-intein fusion sequence: 32% and pICH24581 (no linker, optimized barnase -intein fusion sequence: 10%).

For exploiting the stability of the phenotype under extreme environmental conditions, plants carrying pJCH24581, pSch25301 and pICH27371 were grown in phytochambers from germination to maturity at 35 °C (8 h) and 20 °C (8 h. see material and methods). Although some of the plants displayed stress symptoms like reduced tillering and early senescence, all individuals carrying Bar-N and Bar-C were male-sterile whereas, in contrary, the control plants formed vital pollen and seed.

12) Inheritance of the male-sterile phenotype: Analysis of Ti generation

The male-sterile plants were backcrossed with wildtype plants. Except a negligible number of cases (< 2%), all backcrosses led to vital seed, showing that the sterility was restricted to the male gametes and did not affect reproducibility of the plants in general.
In 69 lines (76% of the cases), the pollen sterile phenotype was inherited to the Ti generation. This demonstrates that pollen ablation caused by tissue culture effects (somaclonal variation) is a rare event and does not affect the overall statistic to a significant extent. None of the Ti plants displaying male sterility lacked the barnase transgene fragments. In the majorities of Ti progeny populations, the male-sterile phenotype can be strictly correlated with the presence of both transgene fragments. Segregation of the transgenes leads to reversion of fertility in all cases. Interestingly, some of the plants carry only an N-terminal or a C-terminal fragment of the barnase gene. Such T-DNA-truncation results from a fragmentation of the plasmids caused by physical forces during biolistic delivery. As expected, such individuals form vital pollen.

An example for the proof of a linkage between the male-sterile phenotype and the barnase gene fragments is given for two Ti populations in Fig. 5. However, a number of lines produced also Fi plants that were fertile although they inherited both barnase-fragments (as been shown by PCR). Southern blot analysis of the Fi revealed that the barnase locus of fertile plants differed from that of the male-sterile plants (Fig. 6 d; line pICH27371-1058, plants 4, 8, 14). This indicates the presence of an inactive barnase locus. This result is important since it indicates that the phenotypical differences were not due to epigenetic effects (that may cause uncontrollable variations in transgene expression and therefore would limit the applicability of the system).

13) Proof of site-specific recombination at the stably integrated prolocus

Wheat lines carrying Streptomyces phage PhiC31 integrase were used to induce irreversible site-specific recombination reactions at the stably integrated T-DNA-loci pICH27371 according to the scheme in Figure 10). The identification of suitable double haploid integrase wheat lines was performed with a transient test assay and is
described in Rubtsova et al. (2008) Plant Cell Rep. 27: 1821-1831. The arrangement of the att sequences allows two alternative recombination reactions that may lead to the deletion of either the 3'- or the 5'-part of the T-DNA-locus.

Exposure of the prolocus to a second T-DNA encoding a *Streptomyces* PhiC31 integrase as a result of sexual hybridization led to a derivatization of the target-T-DNA at a high frequency. In total, 30 wheat transformants carrying independent target-T-DNA integrations displayed intrachromosomal recombination, as been proven by PCR-analyses and product sequencing (Figures 11,12). The recombination products could be recovered in subsequent generations; thus an inheritable "genetic switch" was induced by the PhiC31 Integrase in trans. The results demonstrate the feasibility of the α#-integrase as a site specific recombination system for the establishment of the hybrid breeding system in wheat.
1. Method of producing male sterile monocotyledonous plants, comprising the steps of:

a) introducing into a monocotyledonous plant or plant cell a first expression cassette comprising the following elements in 5' to 3' orientation:
   - a tapetum-specific promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for an N-terminal part of a protein which provides for male sterility;
   - a nucleic acid sequence coding for the N-terminal part of a first intein; and
   - optionally, operatively linked thereto a terminator sequence functional in plant cells;

and

b) introducing into said plant or plant cell a second expression cassette comprising the following elements in 5’ to 3’ orientation:
   - a tapetum-specific promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for a C-terminal part of said first intein
   - a nucleic acid sequence coding for at least one copy of a flexible linker sequence;
   - a nucleic acid sequence coding for a C-terminal part of said protein which provides for male sterility; and
- optionally, operatively linked thereto a terminator sequence functional in plant cells.

2. Method according to claim 1, wherein the nucleic acid sequences coding for the N- and the C-terminal part of the protein which provides for male sterility and the nucleic acid sequences coding for the N- and the C-terminal part of the first intein are adapted to the codon usage of monocotyledonous plants.

3. Method according to claim 2, wherein the monocotyledonous plants are Triticeae and wherein the nucleic acid sequences coding for the N- and the C-terminal part of the protein which provides for male sterility and the nucleic acid sequences coding for the N- and the C-terminal part of the first intein are adapted to the codon usage of Triticeae.

4. Method according to any of the preceding claims, wherein the protein which provides for male sterility is an RNase.

5. Method according to claim 4, wherein the RNase is barnase.

6. Method according to claim 4, wherein the nucleic acid sequence coding for the N-terminal part of the barnase is SEQ ID No. 3 or 13 or a functional fragment thereof and wherein the nucleic acid sequence coding for the C-terminal part of the barnase is SEQ ID No. 5 or 27 or a functional fragment thereof.

7. Method according to any of the preceding claims, wherein said first intein is selected from the group consisting of DnaB and DnaE.
8. Method according to any of the preceding claims, wherein the flexible linker sequence essentially consists of glycine and serine residues.

9. Method according to claim 8, wherein the flexible linker sequence is GGGGS.

10. Method according to any of the preceding claims, wherein in said first expression cassette a nucleic acid sequence coding for an amino acid sequence comprising the amino acid sequence RESG or fragments of said sequence is inserted between the nucleic acid sequence coding for the N-terminal part of said protein which provides for male sterility and the nucleic acid sequence coding for a N-terminal part of said first intein.

11. Method according to any of the preceding claims, wherein in said second expression cassette a nucleic acid sequence coding for an amino acid sequence comprising the amino acid sequence SIEQD or fragments of said sequence is inserted between the nucleic acid sequence coding for a C-terminal part of said first intein and the nucleic acid sequence coding for the C-terminal part of said protein which provides for male sterility.

12. Method according to any of the preceding claims, further comprising introducing a nucleic acid sequence coding for a protein which is a phenotypical marker.

13. Method according to claim 12, wherein the phenotypical marker is a protein conferring herbicide resistance.
14. Method according to claim 13, wherein the nucleic acid sequence coding for a protein which confers herbicide resistance is introduced into the monocotyledonous plant or plant cell by the steps of:
c) introducing into said plant or plant cell a third expression cassette comprising the following elements in 5' to 3' orientation:
   - a promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for an N-terminal part of a protein which confers herbicide resistance;
   - a nucleic acid sequence coding for the N-terminal part of a second intein; and
   - optionally, operatively linked thereto a terminator sequence functional in said plant cells; and
d) introducing into said plant or plant cell a fourth expression cassette comprising the following elements in 5' to 3' orientation:
   - a promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for the C-terminal part of said second intein;
   - a nucleic acid sequence coding for a C-terminal part of said protein which confers herbicide resistance; and
   - optionally, operatively linked thereto a terminator sequence functional in said plant cells.

15. Method according to claim 14, wherein the fourth expression cassette further comprises a nucleic acid sequence coding for at least one copy of a flexible linker sequence which nucleic acid sequence is located between the nucleic acid sequence coding for the C-terminal part of said second intein and the nucleic acid sequence coding for a C-terminal part of said protein which confers herbicide resistance.
16. Method according to any of claims 13 to 15, wherein the protein which confers herbicide resistance is acetolactate synthase.

17. Method according to claim 16, wherein the nucleic acid sequence coding for the acetolactate synthase is selected from the group consisting of:
   a) a nucleic acid sequence according to SEQ ID No. 31 or 51;
   b) nucleic acid sequences coding for a protein according to SEQ ID No. 32 or 52 or a functional fragment thereof;
   c) nucleic acid sequences hybridising to a complementary strand of the nucleic acid sequence according to SEQ ID No. 31 or 51 under stringent conditions; and
   d) nucleic acid sequences which are at least 50% homologous to the nucleic acid sequence shown in SEQ ID No. 31 or 51.

18. Method according to any of claims 14 to 17, wherein said second intein is selected from the group consisting of DnaB and DnaE.

19. Method according to any of claims 14 to 18, wherein said second and third or fourth expression cassettes are deleted from a prolocus by site-specific recombination and wherein said first and third or fourth expression cassettes are deleted from said prolocus by site-specific recombination.

20. Method according to any of claims 14 to 19, wherein said first and said third or fourth expression cassettes are located on a first locus on a first homologous chromosome and wherein said second and said third or fourth expression cassettes are located on second locus on a second homologous chromosome, wherein said first and second loci are identical loci on said homologous chromosomes.
21. Method according to any of claims 14 to 20, further comprising the step of selecting male sterile plants by applying a herbicide to said plants.

22. Method of producing monocotyledonous hybrid plants, comprising the steps of:
   a) producing a male sterile monocotyledonous plant by a method according to any of claims 1 to 21; and
   b) crossing the male sterile monocotyledonous plant of step a) with a male fertile monocotyledonous plant.

23. Transgenic plant produced by a method according to any of claims 1 to 22.

24. Transgenic monocotyledonous plant comprising a nucleic acid sequence coding for an N-terminal part of barnase.

25. Transgenic monocotyledonous plant comprising a nucleic acid sequence coding for a C-terminal part of barnase.

26. Transgenic monocotyledonous plant comprising a nucleic acid sequence coding for an N-terminal part of acetolactate synthase.

27. Transgenic monocotyledonous plant comprising a nucleic acid sequence coding for a C-terminal part of acetolactate synthase.

28. Transgenic monocotyledonous plant according to any of claims 24 to 27, wherein the nucleic acid sequences coding for the N-terminal part and the C-terminal part of barnase and/or the nucleic acid sequences coding for the N-terminal
part and the C-terminal part of acetolactate synthase are adapted to the codon usage of monocotyledonous plants.

29. Transgenic monocotyledonous plant according to claim 28, wherein the monocotyledonous plant is from the family of Triticeae and wherein the nucleic acid sequences coding for the N-terminal part and the C-terminal part of barnase and/or the nucleic acid sequences coding for the N-terminal part and the C-terminal part of acetolactate synthase are adapted to the codon usage of Triticeae.

30. Transgenic monocotyledonous plant according to any of claims 24, 28 or 29, wherein the nucleic acid sequence coding for an N-terminal part of the barnase is SEQ ID No. 3 or 13.

31. Transgenic monocotyledonous plant according to any of claims 25, 28 or 29, wherein the nucleic acid sequence coding for a C-terminal part of the barnase is SEQ ID No. 5 or 27.

32. Transgenic monocotyledonous plant according to any of claims 26, 28 or 29, wherein the nucleic acid sequence coding for an N-terminal part of acetolactate synthase is SEQ ID No. 33 or 53.

33. Transgenic monocotyledonous plant according to any of claims 27 to 29, wherein the nucleic acid sequence coding for a C-terminal part of acetolactate synthase is SEQ ID No. 43 or 55.

34. Recombinant nucleic acid molecule comprising the following elements:
   a) a first expression cassette comprising the following elements in 5' to 3' orientation:
- a tapetum-specific promoter functional in cells of a monocotyledonous plant;
- operatively linked thereto a nucleic acid sequence coding for an N-terminal part of a protein which provides for male sterility;
- a nucleic acid sequence coding for the N-terminal part of a first intein; and
- optionally, operatively linked thereto a terminator sequence functional in said plant cells;

and

b) a second expression cassette comprising the following elements in 5' to 3' orientation:
- a tapetum-specific promoter functional in cells of a monocotyledonous plant;
- operatively linked thereto a nucleic acid sequence coding for a C-terminal part of said first intein;
- a nucleic acid sequence coding for at least one copy of a flexible linker sequence;
- a nucleic acid sequence coding for a C-terminal part of said protein which provides for male sterility; and
- optionally, operatively linked thereto a terminator sequence functional in said plant cells.

35. Recombinant nucleic acid molecule according to claim 34, wherein the nucleic acid sequences coding for the N- and the C-terminal part of the protein which provides for male sterility and the nucleic acid sequences coding for the N- and the C-terminal part of the first intein have been adapted to the codon usage of monocotyledonous plants.

36. Recombinant nucleic acid molecule according to claim 35, wherein the monocotyledonous plants are Triticeae and wherein the nucleic acid sequences coding for the N- and the C-terminal part of the protein which provides for male
sterility and the nucleic acid sequences coding for the N- and the C-terminal part of
the first intein are adapted to the codon usage of Triticeae.

37. Recombinant nucleic acid molecule according to any of claims 34 to
5 36, wherein the protein which provides for male sterility is an RNase.

38. Recombinant nucleic acid molecule according to claim 37, wherein
the RNase is barnase.

39. Recombinant nucleic acid molecule according to any of claims 34 to
38, wherein said first intein is selected from the group consisting of DnaB and DnaE.

40. Recombinant nucleic acid molecule according to any of claims 34 to
39, wherein the flexible linker sequence essentially consists of glycine and serine
15 residues.

41. Recombinant nucleic acid molecule according to claim 40, wherein
the flexible linker sequence is GGGGS.

42. Recombinant nucleic acid molecule according to any of claims 34 to
41, wherein a nucleic acid sequence coding for an amino acid sequence comprising
the amino acid sequence RESG or fragments of said sequence is inserted between the
nucleic acid sequence coding for the N-terminal part of said protein which provides
for male sterility and the nucleic acid sequence coding for a N-terminal part from
said first intein in said first expression cassette.

43. Recombinant nucleic acid molecule according to any claims 34 to 42,
wherein a nucleic acid sequence coding for an amino acid sequence comprising the
amino acid sequence SIEQD or fragments of said sequence is inserted between the nucleic acid sequence coding for the C-terminal part of said protein which provides for male sterility and the nucleic acid sequence coding for a C-terminal part from said first intein in said second expression cassette.

44. Recombinant nucleic acid molecule according to any of claims 34 to 43, further comprising an expression cassette comprising the following elements in 5’ to 3’ orientation:
   - a promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for a protein which is a phenotypical marker; and
   - optionally, operatively linked thereto a terminator sequence functional in said plant cells.

45. Recombinant nucleic acid molecule according to claim 44, wherein the protein which is a phenotypical marker is a protein conferring herbicide resistance.

46. Recombinant nucleic acid molecule according to any of claims 34 to 43, further comprising:
   c) a third expression cassette comprising the following elements in 5’ to 3’ orientation:
      - a promoter functional in cells of a monocotyledonous plant;
      - operatively linked thereto a nucleic acid sequence coding for an N-terminal part of a protein which is a phenotypical marker;
      - a nucleic acid sequence coding for the N-terminal part of a second intein; and
      - optionally, operatively linked thereto a terminator sequence functional in said plant cells; and
d) a fourth expression cassette comprising the following elements in 5' to 3' orientation:
   - a promoter functional in cells of a monocotyledonous plant;
   - operatively linked thereto a nucleic acid sequence coding for the C-terminal part of said second intein;
   - a nucleic acid sequence coding for a C-terminal part of said protein which is a phenotypical marker; and
   - optionally, operatively linked thereto a terminator sequence functional in said plant cells.

47. Recombinant nucleic acid molecule according to claim 46, wherein the protein which is a phenotypical marker is a protein which confers herbicide resistance.

48. Recombinant nucleic acid molecule according to claim 45 or 47, wherein the protein which confers herbicide resistance is acetolactate synthase.

49. Recombinant nucleic acid molecule according to claim 48, wherein the nucleic acid sequence coding for the acetolactate synthase is selected from the group consisting of:
   a) nucleic acid molecules comprising a nucleic acid sequence according to SEQ ID No. 31 or 51;
   b) nucleic acid molecules coding for an acetolactate synthase from rice according to SEQ ID No. 32 or 52 or a functional fragment thereof;
   c) nucleic acid molecules hybridising to a complementary strand of the nucleic acid sequence according to SEQ ID No. 31 or 51 under stringent conditions; and
   d) nucleic acid molecules which are at least 50 % homologous to the nucleic acid sequence shown in SEQ ID No. 31 or 51.
50. Recombinant nucleic acid molecule according to any of claims 46 to 49, wherein said second intein is selected from the group consisting of DnaB and DnaE.

51. Recombinant nucleic acid molecule according to any of claims 46 to 50, wherein the recombinant nucleic acid molecule comprises a first part comprising the first expression cassette and the third or fourth expression cassette and a second part comprising the second expression cassette and the third or fourth expression cassette.

52. Recombinant nucleic acid molecule according to claim 51, wherein the first part and the second part are each flanked by recombinase recognition sites.

53. Transgenic plant comprising a recombinant nucleic acid molecule according to any of claims 34 to 52.

54. Isolated nucleic acid molecule comprising a nucleic acid sequence according to SEQ ID No. 49 or a fragment of said nucleic acid sequence coding for a functional barnase fragment.

55. Isolated nucleic acid molecule comprising a nucleic acid sequence according to SEQ ID No. 51 or a fragment of said nucleic acid sequence coding for a functional acetolactate synthase fragment.

56. Isolated nucleic acid molecule coding for the N-terminal part of an acetolactate synthase, selected from the group consisting of:
   a) a nucleic acid sequence according to SEQ ID No. 33 or 53;
b) nucleic acid sequences coding for a protein according to SEQ ID No. 34 or 54 or a functional fragment thereof;

c) nucleic acid sequences hybridising to a complementary strand of the nucleic acid sequence according to SEQ ID No. 33 or 53 under stringent conditions; and

d) nucleic acid sequences which are at least 50% homologous to the nucleic acid sequence shown in SEQ ID No. 33 or 53.

57. Isolated nucleic acid molecule coding for the C-terminal part of an acetolactate synthase, selected from the group consisting of:

a) a nucleic acid sequence according to SEQ ID No. 43 or 55;

b) nucleic acid sequences coding for a protein according to SEQ ID No. 44 or 56 or a functional fragment thereof;

c) nucleic acid sequences hybridising to a complementary strand of the nucleic acid sequence according to SEQ ID No. 43 or 55 under stringent conditions; and

d) nucleic acid sequences which are at least 50% homologous to the nucleic acid sequence shown in SEQ ID No. 43 or 55.
- Figure 2 -

1. pICH12795
2. pICH24431
3. pICH24591
4. pICH24601
5. pICH24612
-Figure 6-

**pICH27371-1219**

[Diagram showing genetic elements and probes]

**pICH27371-1058**

[Diagram showing genetic elements and probes]
- Figure 7 (continued) -

pICH25301

13863 bp
- Figure 7 (continued) -

d) pICH27371

pICH25881
16277 bp

SUBSTITUTE SHEET (RULE 26)
Figure 7 (continued)

plICH25881

16277 bp
Figure 10 -

a)
b)

**F1: Primary transformant**
Male sterile
Resistant

\[ A_1A_2 \] (cis)

\[ \times \] Integrase

**F2: Mosaic plants**

\[ ? \]

\[ X \]

**F3: Fertile plants**
Sensitive

\[ A_1 \]

\[ A_2 \]

\[ \times \]

**F4: Male sterile**
Resistant

\[ A_1 \]

\[ A_2 \] (trans)
c)

Male sterile \( A_1 \) \( A_2 \) \( X \) \( A_1 \) \( A_2 \) fertile

\[ X \]

\( A_1 \) \( A_1 \) \( A_2 \) \( A_2 \)

100% fertile
d)

Male sterile \( A_1 \) - \( A_2 \) - \( A_1 \)

\( X \)

Fertile Sensitive

\[ \downarrow \]

Male sterile \( A_1 \) - \( A_2 \) - \( A_1 \)

50% 50%

resistant
- Figure 11 -

(a) Diagram of pICH27371, pICH27371-N, and pICH27371-C constructs.

(b) Gel electrophoresis of DNA samples, with markers indicating 1 kb and 500 bp.