G00001

A method for inhibiting the expression of target genes in plants

The present invention relates to a method for inhibiting the expression of target genes in plants and to DNA constructs used to implement this method.

Plant transcription factors mostly belong to gene families containing numerous members. An example how large such families can grow is the *myb*-family comprised of at least 97 genes in *Arabidopsis.* Also homeobox and MADS box genes are found in multiple copies comprising several subfamilies in all plant species analyzed so far. In genetic screens for loss of function alleles only few of these genes have been associated with a mutant phenotype. One explanation for numerous silent family members may be genetic redundancy; each gene specific contributions may be subtle or dependent on special conditions. This fraction of silent genes in plant as others genomes provides a major challenge in molecular biology because these are foreseen to oppose a functional analysis also in reverse genetic approaches.

In the transformation of plants, to alter the structure and/or the function of whole plants, it may be desirable to block the expression of genes. This may be effected by altering either the sequence of the gene itself or of the transcriptional factor that controls its transcription. For that purpose the antisense technology, as described for example in the US patent 4,943,674, may be used. Site-directed mutagenesis may alternatively be used. It has for example been proposed to use in vitro altered transcriptional activators encoding proteins that bind to one or more components of the transcription complex in such a way that they competitively inhibit expression of target gene or genes in plants (EP 0 475 584). Co-suppression system may also be used to block expression of specific genes (Flavell R.B., 1994).

However these technologies require an accurate knowledge of the structure of the gene or of the transcriptional factor (namely of the DNA binding and transactivator domains) to allow the antisense sequence to specifically hybridize with the gene sequence, or to allow the mutation to only

159159

affect some properties of the protein (DNA binding or activator domain) respectively.

Other limitations of these methods are well known by one skilled in the art : in the case of redundancy in genes and/or functions, other related genes, non altered by these targeted technologies, conceal the expression of the target gene. As a result, the observed phenotype is wild-type instead of mutant as expected. Furthermore these techniques make the isolation of new genes, or the determination of new functions difficult, sometimes impossible.

The authors of the present invention now propose an alternative approach for blocking the expression of genes in plants which meets the above requirements and presents other advantages. The method of the invention distinguishes from previous technologies in that this is, as further described, a dominant approach with high penetrance. This method is independent from any DNA sequence homology with the targeted gene, which is crucial for antisense and cosuppression approaches, and moreover circumvents the problem of redundancy in genes and/or functions.

The method of the invention involves the use of a chimeric DNA construct comprising at least one repressor sequence in transcriptional fusion with at least one plant-specific sequence that codes for a protein or fragment thereof that bind to DNA or that activates transcription either by binding to DNA itself or by interacting with a DNA-binding protein ;

said repressor sequence being operably linked to elements allowing the transcription of said fused sequences.

The term "plant-specific sequence" means a sequence that originates from plants or has been modified (or shown) to function in plants, i.e. to actually bind to DNA in plants or to activate transcription either by binding to DNA itself or by interacting with a DNA-binding protein.

It is understood that the repressor sequence of said chimeric DNA construct preferably does not contain a sequence that codes for a second other DNA-binding domain, so that no undesirable interaction could occur.

In a preferred embodiment, said repressor sequence is at least the repressor domain of the Drosophila *engrailed* gene (*eng*).

The product of the Drosophila *engrailed (eng)* locus was shown to be involved in the regulatory interactions that govern early embryogenesis (Poole et al, 1985). Eng is a homeodomain-containing transcription factor that is required for cell fate specification through out development of the fly. A pioneering experiment was performed in *Drosophila* where the DNA-binding homeodomain of the *engrailed (eng)* gene was replaced by its *Fushi tarazu (Ftz)* counterpart, and expression of the chimeric gene controlled by heat shock promoter. After induction, the sequence specificity of the *Ftz* homeodomain directed the strong transcriptional *eng* repressor function to *Ftz* target genes resulting in phenocopies of *Ftz* loss of function alleles in transgenic progeny (John et al., 1995). Thus, the Eng protein was shown to be an active repressor in Drosophilia (John et al, 1995) and cultured animal cells (Han & Manley, 1993). In 1996, Conlon et al use an approach in which the DNA-binding domain of a transcription activator is fused to the *engrailed* repressor domain to assist in the analysis of Xenopus and Zebrafish transcription factors (Conlon et al, 1996). But no stable transformation has been carried out until now in plants, in order to obtain transgenic plants with heritable mutant phenotype.

According to the present invention, the term "repression domain of the Drosophila *engrailed* gene *(eng)*" means a fragment of the Drosophila *engrailed* gene or of a derivative sequence thereof, said fragment comprising a nucleotide sequence that encodes a polypeptide interfering with the general transcription machinery and transcriptional activators. Said fragment advantageously comprises a minimal repression sequence coding for a polypeptide of 55 residues (Poole et al., 1985, Han K. et al., 1993). The whole *engrailed* gene may also be used but most preferably without its homeodomain.

"A derivative sequence" is understood as meaning a sequence which differs from the sequence of the Drosophila *engrailed* gene by mutation, insertion, deletion or substitution of one or more bases, or by the degeneracy of the genetic code so long as it codes for a polypeptide which is substantially the same as the Engrailed product.

Engrailed related domains, which share the same repression activity, may also be used, such as those described in Smith et al. (1996).

In another embodiment, the repressor sequence that is used may be a sequence that codes for at least the Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins (Witzgall et al., 1994) for at least the RE-1-silencing transcription factor (REST) (Thiel et al., 1998), or for at least the BTB (for Broad-complex Tramtrac and Bric) domain, also known as POZ-domain (Ahmad et al., 1998 ; Huynh & Bardwell, 1998).

Chimeric DNA constructs further comprising corepressor sequences associated with said repressor sequence are also encompassed. Among the corepressors of interest, one may cite KAP-1 (Friedman et al., 1996), groucho (Tolkunova et al., 1998), or KOX-1 (Moosmann et al., 1997), or N-Cor and SMRT (Huynd and Bardwell, 1998).

In one embodiment, said repressor sequence is fused to a plant-specific sequence that encodes a DNA-binding protein domain. The term "DNA binding protein domain" refers to the whole protein that contains such a DNA-binding domain but also to parts of it independently of their ability to bind to DNA.

In a preferred embodiment, a plant transcription factor, or a DNA-binding fragment thereof, is used as the DNA-binding domain containing protein. A complete transcription factor is preferably used, as it preserves interactions between proteins involved in regulation of transcription.

In an alternative embodiment, the repressor sequence is fused to a sequence that codes for a protein or parts thereof that activates transcription by binding to a DNA-binding protein such as a transcription factor. Such proteins may either increase the affinity of DNA-binding proteins to their target sites or mediate signalling to the transcriptional initiation complex and are generally called coactivators.

In the preferred embodiment wherein the DNA-binding domain containing protein is a plant transcription factor, said transcription factor may preferably be selected from the group consisting of :

- STM ("Shootmeristemless"), member of the Knotted class of homeodomain proteins, which is an essential gene for development and function of the shoot apical meristem ;

- AP3 ("APETALA3"), member of the MADS box proteins, which is a floral organ identity B function gene essential for the development of petals and stamen. AP3 protein is known to form heterodimers with the PISTILLATA (PI) gene product.

STM and AP3 encode plant transcription factors with well-characterized loss of function alleles. Both genes are not only members of different classes of plant transcription factors but deficiencies in both genes cause phenotypic effects at different developmental stages, early in the vegetative phase or rather late in the floral development.

- ZmHox - member of the homeobox proteins, more particularly ZmHox 1a/1b and 2a/2b, which are expressed in maize meristems and proliferating cells from the early embryo to late reproductive organs. This expression pattern suggests a contribution to plant growth and morphogenesis;

- Ms-41-A and Zm-41-A. These two factors are associated with male fertility. Recent results assertain that Ms41-A protein from *Arabidopsis* is a transcription factor belonging the family recently described as ARF1 family for Auxin Response Factor 1 (Ulmasov et al., 1997). Ms41-A analogues may also be used, such as those described in WO 97/23618 or the Arabidopsis gene Monopteros which encodes a transcription factor mediating embryo axis formation and vascular development (Hardtke et al., 1998), and the Arabidopsis gene ETTIN involved in floral development (Sessions, 1997).

The plants transformed with a chimeric DNA construct of the invention comprising such Ms-41-A factor are expected to be male sterile.

In one embodiment, the chimeric DNA construct of the invention may further comprise a sequence that encodes at least the hormone-binding domain of a steroid hormone receptor, said sequence being in frame with the fusion construct consisting of at least one repressor sequence in transcriptional fusion with at least one plant-specific sequence that codes for a protein or

fragment thereof that activates transcription either by binding to DNA itself or by interacting with a DNA-binding protein ;

whereby the nuclear localization of the chimeric protein, that is the product of the translation of said fusion construct is dependent on the application of said steroid hormone or analogues thereof.

The steroid hormone receptor that is used may be for example the glycocorticoid receptor (GR). In that case, dexamethasone may be used to induce the nuclear translocation of the chimeric protein.

Optionally, we can use the amino acid sequences which, when included in a protein, function to promote transport of the protein to the nucleus are known in the art and are termed nuclear localization signals (NLS).

The repression domain of the Drosophila *engrailed* gene (*eng*) is operably linked to elements allowing its expression. Such elements may more particularly comprise a promoter and polyadenylation signals.

In one embodiment, the promoter that is used is a constitutive promoter.

In another embodiment, the promoter that is used is a tissue-specific promoter or a developmentally regulated promoter. This allows a conditional loss of functions that may be desired, for example to design new plant varieties.

In still other embodiments, the promoter that is used is an inducible promoter.

The polyadenylation that can be used may be for example the 35S polyA terminator of cauliflower mosaic virus (CaMV), as disclosed in Franck et al. (1980) and NOS polyA terminator, that corresponds to the non coding 3' region of nopalin synthase of Ti plasmid of *Agrobacterium tumefaciens* (Depicker et al., 1982).

Among the preferred promoters that can be used, one can cite for example :

a) <u>constitutive promoters :</u>

- 35S promoter, or advantageously the double 35S constitutive promoter of CaMV as described in Kay et al. (1987) ;

- rice actin promoter followed by the rice actin intron (PAR-IAR) included in plasmid pAct1-F4 as described in Mc Elroy et al. (1991);
- the constitutive promoter EF-1α of the gene encoding for plant elongation factor described in WO 90/02172 or in Axelos et al. (1989);
- chimeric superpromoter PSP (Ni et al., 1995) constituted by the fusion of a triple repeat of the transcriptional activity element from the promoter of the gene of *Agrobacterium tumefaciens* octopin syntase, the transcriptional activating element of the promoter of the gene *Agrobacterium tumefaciens* mannopin synthase ; and
- ubiquitin promoter from sunflower (Binet et al., 1991).

b) <u>specific promoters :</u>

• *seed-specific promoters :*

- PCRU promoter of radish cruciferin gene allowing the expression specifically in seeds, as described in Depigny-This et al. (1992) ;
- HMWG promoter (High Molecular Weight Glutenin) from barley (Anderson et al., 1989) ;
- the promoter of maize γzein (Pγzein) included in pγ63 plasmid in Reina et al. (1990) allowing the expression in albumen of maize seeds ;
- PGEA1 and PGEA6 promoters corresponding to the non coding 5' region of the genes GEA1 and GEA6, expressed in the grains in *Arabidopsis thaliana* (Gaubier et al., 1993) ; and
- β-phaseolin promoter (Riggs et al., 1989).

• *specific promoters which drive expression in particular plant tissues which are involved in the control of fertility* :

Among the promoters of interest one may cite the Brassicaceae A3 or A9 promoter described in WO 92 11379, the A6 promoter described in WO 93 02197, or TA29, TA26, TA13 promoters described in WO 89 10396 ;

- the Ms41-A anther-specific promoter described in WO 97/23618, which may also be used in male sterility systems ;

- a dehiscence-zone specific promoter such as the one described in EP 692 030.

c) **inducible promoters :**

- a promoter inducible in stress conditions, for example heat shock, wound or interaction with pathogens (Kuhlemeier et al., 1987, WO 94/21793) ; and
- an ethanol-inducible promoter (Salter et al., 1998) ; and
- the PR1a promoter inducible by salicylic acid for example (US 5,689,044).

The DNA construct of the invention is advantageously inserted in a vector, e.g. a plasmid, for use in plant cell transformation.

The transformation of plant cells may be effected by transferring the above vectors in protoplasts, in particular after incubating those protoplasts in a solution of polyethyleneglycol (PEG) in the presence of divalent cations (Ca^{2+}) as described in Krens et al. (1982).

The transformation of plant cells may also be effected by electroporation as described in Fromm et al. (1986).

A gene gun may also be used allowing the projection of metal particles coated with a DNA construct of the invention, whereby genes are delivered into cell nucleus (Sanford et al., 1988).

Another method for transforming plant cells is cytoplasmic or nuclear micro-injection.

In a preferred embodiment, plant cells are transformed with a DNA construct of the invention, by means of a host cell infecting said plant cells. A further subject of the present invention is thus a host cell transformed with a chimeric DNA construct as previously described. Advantageously, the above host cell is *Agrobacterium tumefaciens*, as used in particular in the methods of Bevan et al. (1984) and An (1986), or *Agrobacterium rhizogenes*, in particular as used in the method of Jouanin et al. (1987).

Plant cell transformation is preferably effected by transferring the *Agrobacterium tumefaciens* T region of an extra-chromosomal circular plasmid that induces tumors (Ti) e.g by using a binary system.

For that purpose, two vectors are constructed. In one of these vectors, the T-DNA region is removed by deletion except for the right and left borders, a marker gene being inserted between the two borders to allow the selection in plant cells. The other partner of the binary system is a helper Ti plasmid which is a modified plasmid that has no T-DNA longer but still contains the virulence genes *vir*, necessary for transforming a plant cell. This plasmid is maintained in *Agrobacterium*.

The present invention also provides a transgenic plant or parts thereof, said plant being transformed with a DNA construct of the invention, or deriving from a plant initially transformed with a DNA construct of the invention. Such transgenic plants exhibit heritable phenotypes.

The term "deriving" refers to plants of the following generations, as long as the parent plant is fertile.

The term "parts" of transgenic plants refer in particular to leaves, fruits, seeds, roots or cells that have been genetically transformed.

The preferred plants that are used for transformation may be for example selected from the group consisting of Arabidopsis thaliana, rice, tobacco, maize, Brassica, wheat, tomato and flowers (Petunia, rose, carnation).

The present invention thus provides a method for obtaining a transgenic plant, wherein a DNA construct of the invention is transferred and expressed in a plant cell and said cell is cultured under conditions for regenerating a whole transgenic plant.

The conditions for regenerating a whole plant from a plant cell are well-known by one skilled in the art.

A chimeric DNA construct of the invention may be used for inhibiting the expression of a target gene in the genome of a plant, the transcription of which involves a DNA-binding protein domain as previously defined.

Inhibiting the expression of a target gene in the genome of a plant may be desired in many purposes. Many transcription factors known by the man skilled in the art play a role in the control of metabolic pathways (starch, lipids, amino acids...) or are involved in the plant development, or in the plant sensibility to pathogen. For example, blocking a gene whose expression is necessary for pollen or another formation (e.g. the Ms-41-A transcription factor, as above described) produces male sterility. Blocking the gene controlled by the AP-3 transcription factor also leads to male sterility. As another example, blocking the gene which codes for the enzyme which catalyses the conversion of sugars to starch can be used to produced sweet corn (see EP 475 584).

It is also possible to obtain transgenic plants with enriched content in lysine by using, according to the invention, the opaque 2 transcription factor (Schmidt et al., 1990), involved in the control of the expression of certain zeins. One could further use Myb-related transcription factors involved in the control of anthocyanin biosynthesis in flowers (Martin et al., 1991 ; Matin, 1997), to modify their colour. Use of other members of Myb-related transcription factors playing a role in the regulation of phenylpropanoid and lignin biosynthesis (Tamagnone et al., 1998) could also be interesting. Some others could be involved in cellular development and senescence.

The chimeric construct, fusion of the repressor domain to the plant transcription factor or part of it, can be advantageously used to identify essential protein-protein interaction domains and interacting protein partners in planta in transgenic plants.

In a preferred embodiment, expression of a chimeric protein comprising the repressor domain in fusion to a plant transcription factor is known to cause a dominant phenocopy. By the deletion of increasing parts of plant sequences encoding the transcription factor and their repression in fusion to the repressor domain in transgenic plants, essential protein domains can be identified. These are DNA-binding domains but also others, like protein-protein interaction domains, if the transcription factor is part of multi-component complex. The method according to the invention is therefore also suitable to study protein-protein interactions in planta.

The chimeric DNA construct of the invention can be used to produce phenocopies of loss of function mutants in genes involved in transcriptional control, by reversing the biological function from activation to repression and by providing chimeric fusion proteins in excess over the endogenous function.

Phenocopy is an artificial (transgenic) situation mimicking a mutant phenotype. The term phenocopy is used by Smith et al. (1996), referring to the original description of John et al. (1995), describing a Ftz loss of function phenotype mimicked by the Eng Ftz chimeric protein.

The phenocopy caused by the expression of the chimeric fusion protein allows to associate a biological function to a given transcription factor or gene involved in transcriptional control.

A chimeric DNA construct of the invention may more particularly be used in a method for determining the function of a transcription factor in plants, comprising the steps of :

i) fusing a sequence encoding said transcription factor to a repressor sequence to form a DNA construct as previously defined;

ii) transforming plant cells with said DNA construct;

iii) culturing the plants obtained from the transformed cells and observing a phenocopy of a mutation correlated with the loss of expression of genes controlled by said transcription factor.

The chimeric DNA construct of the invention allows the production of a cDNA library, said library being useful to isolate new genes controlled by the transcriptional activator as above described.

The present invention thus encompasses a method for identifying new genes in plants comprising the steps of :

i) obtaining transgenic plants transformed with a chimeric DNA construct of the invention ;

ii) comparing the RNA population from said transgenic plants with the RNA population of a plant that has not been transformed with a chimeric DNA construct of the invention, by amplifying (for example by Polymerase

Chain Reaction in a differential display approach (Liang et al., 1993) the RNAs repressed by expression of the chimeric DNA construct, identified as genes inactive in said transgenic plants but active in the plant that has not been transformed with a chimeric DNA construct of the invention.

The plants to be compared should genotypically be as identical as possible. Most preferably, the plants that have not been transformed with a chimeric DNA construct of the invention are transgenic plants transformed with a blank vector.

In a preferred embodiment, the method for identifying new genes in plants comprises the steps of :

i) obtaining transgenic plants transformed with a chimeric DNA construct comprising a sequence that encodes at least the hormone-binding domain of a steroid hormone receptor and/or comprising an inducible promoter;

ii) submitting said transgenic plants to an induction by means of a steroid hormone or analogues thereof and/or promoter inducer, whereby a phenocopy is created due to the loss of expression of target genes;

iii) comparing the RNA populations from said transgenic plants before and shortly after induction, by amplifying (for example by Polymerase Chain Reaction in a differential display approach (Liang et al., 1993) the RNAs repressed by expression of the chimeric DNA construct, identified as genes active before but inactive after induction.

In a preferred embodiment, a differential display approach associated with the AFLP technique described by Vos et al. (1995) or the PCR-Select System (Clontech) is used.

The preferred embodiment of the method of the invention wherein an inducible system is used advantageously allows to identify direct target genes controlled by the transcriptional factor, at any time during the plant development.

The creation of phenocopies by means of a chimeric DNA construct of the invention presents several other advantages :

Phenocopies are hardly sensitive to genetic redundancy. In contrast to mutagenesis approaches like gene machines where knock outs of single genes only give phenotypes in non-redundant situations, the phenocopy approach is expected to be informative in redundant situations. This may be important for species which are not or not truly diploid (e.g. wheat or maize respectively). Also in many Brassica crops the genome is triplicated relative to Arabidopsis.

The advantage in redundant situations is related to the fact that the phenocopy is caused by the chimeric protein making the method independent from DNA sequence homology, crucial for antisense and cosuppression approaches, which focuses effects to single or few genes (depending on the degree of sequence similarity). The newly created chimeric repressor protein may perform and compete with all kinds of interactions and functions related to the transcription factor fused with the repressor. It thus uses functional protein domains and redundant protein functions that partly or fully compensate for the loss of a single gene product (reduced by antisense or cosuppression) therefore should be also effected by the chimeric repressor protein.

Phenocopies support results obtained in gene machines, where many additional elements in the genome frequently interfere with the conclusive association of a phenotype to a specific insertion.

The phenocopy approach is also helpful in transferring information between species, e.g. from Arabidopsis to Brassica crops or between grass crops (rice, maize, wheat, barley) and in comparing regulatory networks between species. By this way, it is possible to detect similarities and differences between groups of genes regulated by the same transcription factor within several species.

The high penetrance (80 %) and dominance of phenocopies in primary transformants distinguishes this method from lower frequency antisense and cosupression approaches, where a few individual progeny are selected.

The below examples and figures illustrate the invention without limiting its scope in any way.

LEGENDS OF FIGURES

Figure 1 represents a top view into the shoot apical meristem of a young wild-type Arabidopsis seedling. In addition to both cotyledons, it has developed four elaborated leaves plus two additional leaf primordia in the center covering the functional shoot apical meristem.

Figure 2 represents a top view into the shoot apical meristem of Arabidopsis phenocopy eng-STM of the same age as in figure 1. This phenocopy has two horizontal cotyledones and a single leaf in vertical position. A functional shoot apical meristem is missing and has presumably been consumed by initiation of the single leaf primordium.

Figure 3 represents a flower of wild-type *Arabidopsis thaliana* with outermost sepals (green), white petals, yellow stamen and central carpel.

Figure 4 represents a flower of *Arabidopsis thaliana* phenocopy eng-AP3. A single sepal has been removed frontally, two neighbouring normal sepals are slightly displaced left and right to allow insight into the second and third floral whorls. Petals and stamen are obviously replaced by sepaloid organs close to the central carpel and small filamentous structures, one visible at the bottom between the two sepaloid petals.

Figure 5 represents a eng-STM construct.

Figure 6 represents a eng-AP3 construct.

EXAMPLES

Example 1 : Construct eng-STM

Creation of pRT Ωeng

The cDNA clone that is used is D_2B clone as described in Poole et al, 1985. This clone was obtained by inserting a 2 kb EcoRI fragment of the engrailed cDNA in a pEMBL vector. The engrailed cDNA sequence is identical to the data base entry (Genbank Access. M 10017) except for the lack of the 3'

end downstream the EcoRI site gaattc, i.e the lack of the 3' end from nucleotide 2014 included.

From the cDNA clone D_2B clone a 929 Bp AflII-BamHI fragment was isolated and inserted into NotI/BamHI cleaved pRTΩNot-Asc vector. Compatibility of the NotI/AflII sites was achieved by fill in reaction with Klenow enzyme. The cDNA fragment covers the natural translation start and 298 amino acid residues of the engrailed protein. The BamHI site or a following XbaI site can be used to create translational fusions to proteins of interest.

The construct obtained was pRTΩeng.

The pRTΩNot-Asc is described in Überlacker et al. (1996). It is a derivative of pRT100 (Töpfer et al., 1987) and contains the CaMV 35S promoter and a polyA signal of Cabb B.

ENG-STM cloning strategy

The sequence of the *STM* coding region that is used is disclosed in Long et al, 1996 and is identical to the database entry (Genbank – Access U32344).

To fuse the *STM* coding region to the engrailed repressor domain the frame of the unique XbaI cloning site in pRT Ωeng had to be shifted which was achieved by inserting a GATCTCGA adaptor into the upstream BamHI site, which was destroyed. The STM coding region previously amplified by reverse transcriptase PCR from Arabidopsis RNA with an 5' terminal XbaI (upstream the translation start ATG) was inserted into the XbaI site of the adpted pRT Ωeng. The STM coding sequences used contain a BamHI site immediately preceeding the natural translation stop codon, which was used subsequently to create the c-terminal GR fusion.

ENG-STM without homeodomain (HD)

The homeodomain of STM was removed in an independent experiment. For that purpose, the plasmid pRTΩ eng-STM was first partially digested with HindII, secondly to completion with BamHI and ends were ligated after fill in reaction with Klenow enzyme. Thus, the construction eng-STM (ΔHD) was obtained.

Example 2 : Construct eng-AP3

ENG-AP3 cloning strategy

The sequence of the *AP3* coding region that is used is disclosed in Jack et al., (1992) and is identical to the data base entry (Genbank – Access – D21125).

The AP3 cDNA clone coding region is fused to a myc epitope at its carboxy-terminus. The protein coding region including the myc epitope was amplified by PCR with primers adding terminal BamHI sites. The resulting BamHI fragment was inserted into the BamHI site of pRTΩeng. This construct was designated as pRTΩeng-AP3.

Example 3 : Construct eng-STM-GR

ENG-STM-GR

For fusion of the chimeric eng-STM polypeptide with the GR domain (Lloyd et al, 1994) a XhoI site in front of the TMV Ω leader was converted into a BamHI site The eng-STM coding region was then inserted into pBI-ΔGR (Simon et al., 1996 ; Schena et al., 1991) as a BamHI fragment. The artificial BamHI site (see above) in front of the natural STM STOP codon fuses the eng-STM poypeptide in frame with the GR domain.

Example 4 : Obtention of transgenic plants

A – *Arabidopsis thaliana*

Both constructs, ENG-STM and ENG-AP3, were inserted into pGPTV-Bar Asc (Überlacker et al., 1996) by use of the AscI sites flanking the expression cassette and subsequently transferred into *Agrobacterium tumefaciens*, GV3101. Infiltration of *Arabidopsis* immature inflorescences followed the protocol of Bechtold et al., (1993) with minor modifications, as follows :

A single *A. tumefaciens* colony was grown at small scale (5 ml) in liquid YEB medium (5 g beef extract, 5 g saccharose, 1 g yeast extract, 1 g bacto-tryptone, 2 mM $MgSO_4$ per liter) at 28° C for 48 hours. From this preculture 0.1 ml each are used to inoculate 4 x 500 ml LB medium (5 g yeast extract, 10 g bacto-tryptone, 10 g NaCl per liter) which are grown with vigorous shaking for 16 hours at 28°C. Bacteria are pelleted at 500 x g and resuspended in infiltration medium (50 g saccharose, 0.2 ml silvet per liter) to OD_{600} = 0.8. Twenty plants were infiltrated for each constructs. Immature Arabidopsis inflorescences (ecotype Columbia) are therefore inserted upside down into the bacterial suspension in infiltration medium. Vacuum is produced with an oil pump for 5 minutes and released rapidly, this step is repeated 3 times. Subsequently, plants are transferred to the green house until maturation and seed harvest. Seeds are vernalized and transgenic progeny selected for BASTA (eng/stm-GR; kanamycin) resistance.

As shown on Figures 1 to 4, the phenotypes observed upon expression of eng-STM and eng-AP3 are those expected. Transgenic *eng-STM* plants exhibited abnormalities from the earliest seedling stage on. Compared to wild type (Figure 1), the petiole of this single leaf originates centrally not laterally the petioles of both cotyledons. Exceptionally shoot apical meristem (SAM) activity in *eng-STM* plants was completely terminated after initiation of a single leaf. Generally, however, subsequent primordia all develop in the same fashion and with accidental timing. These trans-dominant alterations phenocopy weak *stm* alleles and have been reproduced in various independent large-scale experiments (< 1000 transgenic T1 progeny). Generally *stm* phenocopies are found in 75 % to 90 % of the primary transformants. Alterations may be subtle early in seedling development but are manifested with increasing numbers of leave primordia until the primary SAM ultimately looses its activity in the vast majority of transgenic *eng-STM* plants.

Concomitantly or after a short break axillary meristems are activated. The resulting secondary axes exhibit weaker phenotypic alterations and often develop quite normal rosettes. A peculiar feature of these axillary meristems is switching of meristem identity back and force between vegetative and inflorescence state resulting in air-borne rosettes. The numbers of

inflorescence/vegetative cycles may vary between individual transgenic plants and exceptionally cycling may perpetuate to plant death. However, the majority of plants ultimately develop fertile flowers sometimes at terminal position. Other peculiarities are a reduced leaf size, alterations in leaf shape, fused leaves or fasciation of the inflorescence. To confirm heredity of the dominant-negative phenotype progeny of 20 T_1 plants exhibiting mild to strong *stm* phenocopies were analysed in the next generation. BASTA resistant T_2 progeny were obtained from 18 T_1 plants and phenocopies recovered in 16 families. The strength of phenotypic alterations and absolute numbers of affected progeny varied between different families, a variability presumably related to the expression level of the transgene, which may also account for the failure to recover phenocopies in the two residual families.

In contrast to *eng-STM* transgenic plants the phenotypic alterations caused by ectopic expression of the *eng-AP3* construct were exclusively restricted to the flower. In 70% of the primary transformants, petals were converted to sepals in the second floral whorl and filamentous structures replaced stamens in the third whorl (compare wild type in Figure 3 with transgenic plant in 4). Both homeotic transformations are reminiscent of strong *ap3* alleles and indicate that ectopic expression of the chimeric *eng-AP3* protein trans-dominantly copies the *ap3* loss of function phenotype. Due to the lack of functional stamen those *ap3*-phenocopy plants required backcrossing, but strictly segregating with the BASTA resistance marker this trans-dominant phenotype was transmitted to the next generation. The flowers from the plants transformed with the eng-AP3 construct are male sterile.

In order to identify the chimeric eng-AP3 protein in the cell, one could advantageously use the Myc epitope, which would be helpful to study for example whether the chimeric protein is expressed, whether it is cytoplasmatic or nuclear, whether it forms a heterodimer with the pistilata gene product or whether nuclar import is dependent on pistilata.

In summary the ectopic expression of both chimeric constructs, *eng-STM* and *eng-AP3*, results in stable, inheritable trans-dominant functions phenocopying *stm* or *ap3* loss of function alleles, respectively. Excess of

chimeric repressor transcription factor fusions can efficiently displace the native gene products from target genes.

Although expressed from a constitutive promoter both chimeric repressors eng-STM and eng-AP3 are obviously recruited by either DNA sequence specificity, protein-protein interactions or both to their natural target genes which are repressed resulting in the phenocopy of the respective loss of function mutant phenotype.

B – Maize

Genetic transformation of maize, whatever the method (electroporation, Agrobacterium, microfibres, particle gun) generally involves the use of an undifferentiated cells in rapid divisions that can still regenerate into a complete plant. This type of cells constitutes the embryogenic callus (of type II) of maize. Such callus is obtained from immature embryos having the genotype HI II or A188 x B73 according to the method and the media described in Dennehey et al, 1994, and may be multiplied and maintained by successive prickings every two-weeks on the initiation medium.

Plants are regenerated from these callus by modifying the hormonal and osmotic equilibrium of the cells according to the method described by Vain et al, 1989.

These plants are then acclimated in greenhouse where they can be crossed or self-pollinated.

A method of genetic transformation leading to a stable integration of the modified genes in the genome of the plant is used. This method involves the use of a gene gun. The target cells are fragments of callus having a surface area of 10 to 20 mm^2. Four hours before bombarding, the fragments are laid in the center of a Petri dish containing a culture medium identical to the initiation medium, further containing 0.2 M of mannitol + 0.2 M of sorbitol (16 fragments per dish). Tissues are then bombarded as previously described. The dishes are then sealed using Scellofrais® and cultivated in dark at 27°C. The first planting is effected 24 hours afterwards, then every two-weeks during three months in a medium identical to the initiation medium but containing a selective agent. The

selective agents may be for example active ingredients of herbicidal agents (Basta®, Round up®) or antibiotics (hygromycin, kanamycin).

After three months or sometimes sooner, callus that have not been inhibited by the selected agents, develop. These are usually composed of cells resulting from the division of a cell having integrated in its genome one or more copies of gene of interest. The frequency of obtaining such callus is about 0.8 callus per bombarded dish. Callus are identified, individualized, amplified and cultivated so as to regenerate plants. In order to avoid any interference with non transformed cells, all these steps are effected in culture media that contain the selective agent. The regenerated plants are acclimated and cultivated in the greenhouse were they can be crossed or self-pollinated.

In an other embodiment, the transformation of plants may be carried out with *Agrobacterium tumefaciens* and immature embryos, as described by Ishida et al., (1996).

Example 5 : Expression of the endogenous STM and AP3 genes

One explanation for these phenocopies might be homology dependent cosuppression mediated by high RNA levels of the transgene. To exclude this, transgenic eng-STM or eng-AP3 phenocopy plants were subjected to RNA gel blot and RT-PCR analysis. The chimeric transcripts derived from the CaMV 35S promoter were easily detectable in total or poly(A)$^+$ RNA from pooled seedlings or flowers (12 individual progeny) probed with the STM or AP3 coding regions. Generally the high abundance of chimeric transcripts interfered with the detection of the shorter native mRNA in these gel blot experiments. Discrimination was achieved by probing with the natural 3'UTR sequences that are lacking in the transgenes. The result obtained for the STM 3'UTR confirms transcription of the native STM gene although at a significantly reduced level.

This reduction is not due to absence of the STM or AP3 transcripts in individual T_1 progeny, because the native transcripts could be detected by RT-PCR experiments in each individual phenocopy plant. To

distinguish the native mRNA from the chimeric transcript a forward primers located in the protein coding region of either the STM or the AP3 transcripts was combined with reverse primers residing in both native 3' untranslated regions respectively or in the Cabb-B polyadenylation signal of pRT Not (Überlacker et al., 1996). Single tube RT-PCR experiments with three primer combinations failed to detect the endogenous transcripts because the unique forward primer achieved saturation by excess of the chimeric amplicons. In contrast the predicted PCR products, 456 bp STM and 466 bp AP3, were easily detected with the gene specific primer pairs. We have not tried to quantify these RT-PCR experiments because we wanted only to confirm presence of the native transcript in each transgenic plant, but the amount of amplicon differs between individual phenocopy plants. The level of endogenous gene transcripts is always significantly lower than in wild type, a reduction, however, which coincides with significant morphological changes in the expressing organs. In eng-AP3 transgenic flowers the appropriate organs, petals and stamen, are lacking and converted to sepals or small carpeloid filaments.

Example 6 : Phenocopies rely on the incorporation of the chimeric eng-STM protein into the nuclear compartment

Although the changes in size or identity of the expressing organs might explain the lower transcript levels observed in *stm* and *ap3* phenocopy plants we aimed to elaborate further evidence against homology based gene silencing. The chimeric *eng-STM* polypeptide therefore was expressed in C-terminal fusion with the hormone-binding domain of the glucocorticoid receptor in transgenic *Arabidopsis* plants. Due to this addition the resulting *eng-STM-GR* polypeptide should be cytoplasmatic and incorporated into the nuclear compartment only after hormone application. A linkage between the dexamethasone treatment and alterations in SAM activity thus should depend on the nuclear import of the chimeric *eng-STM-GR* protein and be incompatible with cosuppression.

Some of the primary transformants (T_1) obtained with the *eng-STM-GR* transgene exhibited weak *stm* phenocopies. To test for

dexamethasone inducibility we focused on one T_2 family with phenotypically normal kanamycine resistant seedlings. After confirming their normal development to the 4-5 leaf stage 10 transgenic seedlings were sprayed with dexamethasone and photographed daily to document the developmental progress. As controls either wild-type plants treated with dexamethasone or untreated transgenic progeny were analysed over the same time interval. In the control plants, 4-6 leaf primordia newly appear within the depicted 6-day interval in these. In contrast after dexamethasone application only a single leaf primordia gets evident in the transgenic *eng-STM-GR* plants. This result was representative for 7 out of 10 transgenic progeny. The residual three transgenic progeny showed a normal development progress, however, RNA gel blot analysis revealed absence of the chimeric *eng-STM-GR* transcript, which was detected at high levels in the responsive progeny. These 3 escapes therefore are related to transgene silencing and the temporary arrest in leaf development appeared in strict correlation to the transcriptional activity of the transgene. Dexamethasone inducibility in this family, carrying a single T-DNA insertion is inherited to subsequent generations (up to T_4 now). But normal development is confined to heterozygous plants, while homozygous progeny often manifest a weak *stm* phenocopy late in vegetative development. We take this as an indication that the amount of the chimeric *eng-STM-GR* protein is not allowed to exceed a threshold level, an assumption also compatible with the few *stm* phenocopies observed among primary *eng-STM-GR* transformants.

Presence of the hormone in addition to the temporary arrest in SAM activity effects growth of leaf primordia. Petioles of immature leaves hardly elongated in comparison to control plant, the leaflets in *eng-STM-GR* plants consequently remain closely attached to the shoot axis. Also retarded is growth of the leaf lamina, which is often associated with leaf curling. Most of these defects in leaf development resemble observations in *stm* phenocopy plants, where initiated leaves generally grow slowly compared to wild-type plants or unaffected progeny. The only pronounced difference compared to *eng-STM* transgenic plants is the lack in petiole elongation, which, however, may be a consequence of the C-terminal fusion with the GR hormone-binding

domain. The dexamethasone inducibility strongly favours the assumption that the chimeric *eng-STM* protein exerts a dominant-negative function internally the plant cell nucleus.

Example 7 : Expression of the C-terminal deletion polypeptide eng-STM (ΔHD)

To further investigate the mechanism transgenic plants were raised expressing the C-terminal deletion polypeptide *eng-STM (ΔHD)*. This deletion protein still spans 296 aa of *STM* protein sequence including the conserved KNOX (Bürglin, 1997) and ELK domains (Kerstetter et al., 1994) but lacks the DNA-binding homeodomain. Transcription of the *eng-STM (ΔHD)* transgene was confirmed in BASTA resistant progeny by RT-PCR resulting in the predicted 192 bp amplicon (compared to 374 bp for *eng-STM*). The obtained primary transformants exhibited *stm* phenocopies in frequency and strength indistinguishable from those observed with the full-length protein. With this construction, the trans-dominant negative function provided by the chimeric *eng-STM* protein is therefore independent of the DNA-binding activity contributed by the *STM* homeodomain.

Further results obtained in transgenic *Arabidopsis* indicate that only 50 amino acids residues of the *STM* protein are sufficient to mediate *stm* phenocopies in fusion with the *eng* N-terminus.

REFERENCES

- Ahmad et al, (1998), Proc. Natl. Acad. Sci. USA, 95, 1213-1218
- An G., (1986), Plant Physiol., 81, 86-91
- Anderson et al, (1989), T.A.G, 77, 689-700
- Axelos et al, (1989), Mol. Gen. Genet., 219, 106-112
- Bechtold et al, (1993), C.R. Acad Science, 316, 1194-1199
- Bevan et al, (1984), Nucleic Acid Research, 12, 8711-8721.
- Binet et al, (1991), Plant Science, 79, 87-94
- Bürglin, (1997), Nucleic Acids Research, 25, 4173-4180
- Conlon et al, (1996), Development, 122, 2427-2435
- Dennehey et al, (1994), Plant Cell Tissue and Organ Culture, 36, 1-7
- Depicker et al, (1982), J. Mol. Appl. Genet., 1, 561-573
- Depigny-This et al, (1992), Plant. Mol. Biol., 20, 467-479
- Flavell (1994), Proc. Natl. Acad. Sci. USA, 91, 3490-3496
- Franck et al, (1980), Cell, 21, 285-294
- Friedmann et al, (1996), Genes Dev., 10, 2067-2078
- Fromm et al, (1986), Nature, 319, 791-793
- Gaubier et al, (1993), Mol. Gen., 238, 409-418
- Han & Manley, (1993), EMBO J., 12, 2723-2733
- Hardtke et al, (1998), EMBO J., 17, 1405-1411
- Huynh & Bardwell, (1998), Oncogene, 17, 2473-2484
- Ishida et al, (1996), Nature Biotechnology, 14, 745-750
- Jack et al, (1992), Cell, 68, 683-697
- John et al, (1995), Development, 121, 1801-1813
- Jouanin, (1987), Plant. Sci., 53, 53-63
- Kay, (1987), Science, 236, 1299-1302
- Kerstetter et al, (1994), The Plant Cell, 6, 1877-1887
- Krens et al, (1982), Nature, 296, 72-74
- Kuhlemeier et al, (1987), Ann. Rev. Plant. Physiol., 38, 221-257
- Liang et al, (1993), Nucleic Acids Research, 21, 3269-3275
- Lloyd et al, (1994), Science, 266, 436-439

- Long et al, (1996), Nature, 379, 66-69
- Martin C. et al, (1991), Plant J., 1, 37-49
- Matin C., (1997), Trends Genet., 13, 67-73
- Mc Elroy, (1991), Mol. Gen. Genet., 231, 150-160
- Moosmann et al, (1997), Biol. Chem., 378, 669-677
- Ni et al, (1995), Plant J7, 661-676
- Poole et al, (1985), Cell, 40, 37-43
- Reina et al, (1990), Nucleic Acids Research, 18, 6426
- Riggs et al, (1989), Plant Science, 63, 47-57
- Salter et al, (1998), The Plant Journal, 16, 127-132
- Sanford, (1988), Trends in Biotechnology, 6, 299-302
- Schena et al, (1991), Proc. Natl. Acad. Sci. USA, 88, 10421-10425
- Schmidt et al, (1990), Proc. Natl. Acad. Sci. USA, 87, 46-50
- Sessions (1997), Development, 124, 4481-4491
- Simon et al, (1996), Nature, 384, 59-62
- Smith et al, (1996), Development, 122, 3141-3150
- Tamagnone et al, (1998), The Plant Cell, 10, 135-154
- Thiel et al, (1998), J. Biol. Chem., 273, 26891-26899
- Tolkunova et al, (1998), Mol. Cell. Biol., 18, 2804-2814
- Töpfer et al, (1987), Nucleic Acids Research, 15, 5890
- Überlacker et al, (1996), Molecular Breeding, 2, 293-295
- Ulmasov et al, (1997), Science, 276, 1865-1868
- Vain et al, (1989), Plant Cell Tissue and Organ Culture, 18, 143-151
- Vain et al, (1989), Journal of Plant Physiology, 135, 537-540
- Vos et al, (1995), Nucleic Acids Research, 23, 4407-4414
- Witzgall et al, (1994), Proc. Natl. Acad. Sci. USA, 91, 4514-4518

CLAIMS

1. A chimeric DNA construct comprising at least one repressor sequence in transcriptional fusion with at least one plant-specific sequence that codes for a protein or fragment thereof that binds to DNA or that activates transcription either by binding to DNA itself or by interacting with a DNA-binding protein ;
said repressor sequence being operably linked to elements allowing the transcription of said fused sequences.

2. A chimeric DNA construct according to claim 1, wherein said repressor sequence is at least the repressor domain of the Drosophila *engrailed* gene (*eng*).

3. A chimeric DNA construct according to claim 1, wherein said repressor sequence encodes for at least the Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins, or for at least the RE-1-silencing transcription factor (REST), or for at least the BTB (for Broad-complex Tramtrac and Bric) domain, also known as POZ-domain.

4. A chimeric DNA construct according to any of claims 1 to 3, wherein said plant-specific sequence codes for at least a DNA-binding domain of a plant transcription factor.

5. A chimeric DNA construct according to claim 4, wherein said transcription factor is selected from the group consisting of STM, AP3, ZmHox, Ms-41-A and Zm41-A.

6. A chimeric DNA construct according to any of claims 1 to 3, wherein said plant-specific sequence activates transcription by interacting with a DNA-binding protein.

7. A chimeric DNA construct according to any of claims 1 to 6, further comprising a sequence that encodes at least the hormone-binding domain of a steroid hormone receptor, said sequence being in frame with the fusion construct consisting of at least one repressor sequence in transcriptional fusion with at least one plant-specific sequence that codes for a protein or fragment thereof that activates transcription either by binding to DNA itself or by interacting with a DNA-binding protein ;
whereby the nuclear localization of the chimeric protein, that is the product of the translation of said fusion construct is dependent on the application of said steroid hormone or analogues thereof.

8. A chimeric DNA construct according to claim 7, wherein said hormone-binding domain of steroid hormone receptor is the hormone-binding domain of a glycocorticoid receptor.

9. A chimeric DNA construct according to any of claims 1 to 8, wherein the elements allowing the transcription of said repressor sequence comprise a constitutive promoter.

10. A host cell transformed with a DNA construct according to any of claim 1 to 9.

11. A transgenic plant or parts thereof, said plant being transformed with a DNA construct according to any of claims 1 to 9, or deriving from a plant initially transformed with a DNA construct according to any of claims 1 to 9.

12. A method for obtaining a transgenic plant according to any of claims 10 or 11, wherein a DNA construct according to any of claims 1 to 9 is transferred and expressed in a plant cell and said cell is cultured under conditions for regenerating a whole transgenic plant.

13. Use of a chimeric DNA construct according to any of claims 1 to 9 for inhibiting the expression of a target gene in the genome of a plant, the transcription of which is activated by a protein encoded by a plant-specific sequence as defined in any of claims 1 to 9.

14. A method for determining the function of a transcription factor in plants, comprising the steps of:

i) fusing a sequence encoding said transcription factor to a repressor sequence to form a DNA construct as defined in any of claims 4 to 9 ;

ii) transforming plant cells with said DNA construct;

iii) culturing the plants obtained from the transformed cells and observing a phenocopy of a mutation correlated with the loss of expression of genes controlled by said transcription factor.

15. A method for identifying new genes in plants, comprising the steps of :

i) obtaining transgenic plants transformed with a chimeric DNA construct according to any of claims 1 to 6 ;

ii) comparing the RNA population from said transgenic plants with the RNA population of a plant that has not been transformed with a chimeric DNA construct according to any of claims 1 to 6, by amplifying the RNAs repressed by expression of the chimeric DNA construct, identified as genes inactive in said transgenic plants obtained in step i) but active in the plant that has not been transformed with a chimeric DNA construct.

16. A method for identifying new genes in plants, comprising the steps of :

i) obtaining transgenic plants transformed with a chimeric DNA construct according to claims 7 or 8, comprising a sequence that encodes at least the hormone-binding domain of a steroid hormone receptor, and/or according to claim 9, comprising an inducible promoter ;

ii) submitting said transgenic plants to an induction by means of a steroid hormone or analogues thereof and/or promoter inducer, whereby a phenocopy is created due the loss of expression of target genes ;

iii) comparing the RNA populations from said transgenic plants before and shortly after induction, by amplifying the RNAs repressed by expression of the chimeric DNA construct, identified as genes active before but inactive after induction.