Transgenic Research **10:** 513–521, 2001.

T-DNA activation tagging as a tool to isolate regulators of a metabolic pathway from a genetically non-tractable plant species

Leslie van der Fits, Frédérique Hilliou & Johan Memelink*
Institute of Molecular Plant Sciences, Clusius Laboratory, Leiden University, Wassenaarseweg 64, 2333 AL, Leiden, The Netherlands

Received 30 October 2000; accepted 12 March 2001

Key words: activation tagging, *Catharanthus roseus*, ORCA, secondary metabolism, terpenoid indole alkaloids, tryptophan decarboxylase

Abstract

T-DNA activation tagging is a method used to generate dominant mutations in plants or plant cells by the insertion of a T-DNA which carries constitutive enhancer elements that can cause transcriptional activation of flanking plant genes. We applied this approach to the species *Catharanthus roseus* (L.) G. Don (Madagascar periwinkle), in an attempt to isolate regulators of genes that are involved in the biosynthesis of secondary metabolites of the terpenoid indole alkaloid (TIA) class. Several TIAs have pharmaceutically interesting activities, including the anti-tumour agents vincristine and vinblastine. The use of suspension-cultured cells enabled us to screen in a relatively easy way hundreds of thousands of T-DNA-tagged cells for resistance to a toxic substrate of one of the TIA biosynthetic enzymes: tryptophan decarboxylase. This screening yielded several interesting tagged cell lines. Further characterisation of one of the tagged cell lines led to the isolation of *Orca3*, a gene encoding an AP2/ERF-domain transcription factor that acts as a master regulator of primary and secondary metabolism. The T-DNA activation tagging results described in detail in this paper illustrate the usefulness of this approach to isolate regulators of a complex metabolic pathway from a genetically non-tractable plant species.

Introduction

One of the most direct ways of dissecting complex biological processes in plants is generation and analysis of genetic mutants. Mutations resulting from T-DNA or transposon tagging or chemical mutagenesis usually cause loss of function, and are therefore recessive. Consequently, the mutant phenotype can only be visualised following selfing of the mutated plants. This demands a substantial amount of effort, and is not possible for all plant species. Another drawback of classical mutagenesis is that mutation of functionally redundant genes does not lead to phenotypically altered plants.

Many of these disadvantages are circumvented by an alternative approach to generate mutants, called T-DNA activation tagging. A T-DNA carrying a promoter reading towards its border is introduced into plant cells. Upon random T-DNA integration into the genome, flanking plant sequences can be transcribed, which can result in a dominant mutation. Therefore, in contrast to classical mutagenesis, mutants generated by T-DNA activation tagging allow direct selection for the desired phenotype in the primary transformants. Furthermore, a phenotype can result from T-DNA activation tagging of a functionally redundant gene, allowing its analysis and cloning. T-DNA activation tagging has been successfully used to isolate genes involved in ABA (Furini et al., 1997) and cytokinin (Kakimoto, 1996) signaling, and to isolate *Arabidopsis* mutants with altered phenotypes (Kardailsky et al., 1999; Neff et al., 1999; Weigel et al., 2000).

Knowledge about regulation of plant secondary metabolism is still limited. Biosynthesis of terpenoid indole alkaloids (TIAs), a class of secondary

* Author for correspondence:
E-mail: memelink@rulbim.leidenuniv.nl

metabolites, is highly regulated in the plant species *Catharanthus roseus*, and depends on plant cell type and environmental conditions. TIA biosynthesis is induced by fungal elicitors (Moreno et al., 1995) and by the plant stress hormone jasmonate (Aerts et al., 1994; Gantet et al., 1998; Vazquez-Flota & De Luca, 1998). This regulation occurs largely, if not completely, at the level of gene expression. Several biosynthetic genes were shown to be induced by fungal elicitors (Pasquali et al., 1992) and methyljasmonate (Menke et al., 1999; Geerlings et al., 2000; van der Fits & Memelink, 2000). Several TIAs have pharmaceutically interesting activities, such as the anti-tumour compounds vincristine and vinblastine. Knowledge about the regulation of TIA biosynthesis can potentially be applied to achieve cheaper and more stable production of these expensive phytochemicals.

Here we describe a T-DNA activation tagging approach to isolate regulators of TIA biosynthetic genes in *C. roseus*. The gene encoding the TIA biosynthetic enzyme tryptophan decarboxylase (TDC) was used as a marker for mutant selection (Goddijn et al., 1993). TDC converts L-tryptophan into tryptamine, one of the first steps in TIA biosynthesis. TDC can use certain L-tryptophan derivatives as a substrate, such as 4-methyl-tryptophan (4-mT) (Sasse et al., 1983a). This compound is toxic for plant cells, and is converted by TDC into the non-toxic 4-methyl-tryptamine. *C. roseus* suspension-cultured cells were transformed with a T-DNA construct carrying enhancer elements from the cauliflower mosaic virus (CaMV) 35S RNA promoter located near the right border, and subsequently selected on 4-mT. Resistant cell lines were further screened for high *Tdc* expression by northern blot analysis. This research strategy has resulted in the isolation of *Orca3*, a gene encoding an AP2/ERF-domain transcription factor that acts as a master regulator of primary and secondary metabolism in *C. roseus* (van der Fits & Memelink, 2000). Here, the T-DNA activation tagging strategy is described in detail, and the results indicate the usefulness of this approach to isolate regulators of a complex metabolic pathway from a genetically non-tractable plant species.

Materials and methods

Construction of plasmids

To provide the tagging vector Tag-2B4A1 with unique restriction sites for plasmid rescue, a pUC21 (Vieira & Messing, 1991) derivative containing the *hpt* gene from pTRA151 (Zheng et al., 1991) driven by the Δ34S promoter (van der Fits & Memelink, 1997), was supplemented with a *Xho*I/*Eco*RI fragment containing the multi-cloning site from pIC20R (Marsh et al., 1984). In the *Bgl*II site of this plasmid, a *Bam*HI/*Bgl*II fragment carrying a CaMV 35S promoter derivative consisting of multimerized enhancer domains (van der Fits & Memelink, 1997) was inserted. The resulting plasmid was linearized at the unique *Bgl*II site and inserted into the *Bam*HI site of pSDM14 (Offringa, 1992), a pBIN19-derived binary vector containing synthetic octopine Ti-plasmid-derived border repeats. A schematic representation of the T-DNA of the tagging vector is shown in Figure 1. For the determination of the 4-mT selection window, cells were transformed with a binary vector derivative of pMOG22 containing CaMV 35S-*Tdc* or the empty vector pMOG22 (Zeneca MOGEN, Leiden, the Netherlands).

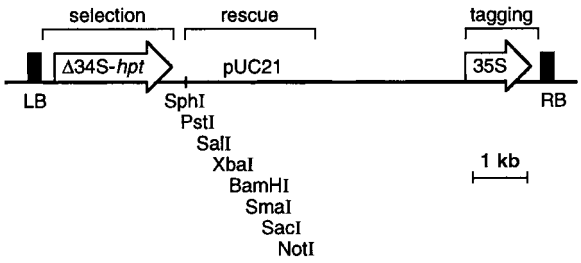

Figure 1. Schematic representation of the T-DNA of tagging vector Tag-2B4A1. Δ34S-hpt: hygromycin resistance marker driven by a deletion derivative of the figwort mosaic virus 34S promoter, pUC21: *E. coli* pUC21 vector, LB and RB: left and right T-DNA border sequences. 35S represents the CaMV 35S promoter derivative consisting of multimerized B and A1 enhancer domains. The unique restriction sites that can be used for plasmid rescue are indicated in the figure. The functional applicabilities of the different parts of the tagging construct are indicated at the top of the figure.

Cell culture and transformation

C. roseus cell line BIX was transformed using *Agrobacterium tumefaciens* strain LBA4404 containing the compatible plasmid pBBR1MCS carrying a constitutive *virG* gene (van der Fits et al., 2000) and the tagging vector Tag-2B4A1 (LBA4404.pBBR*virG*N54D.Tag2B4A1). After transformation, cells were grown on paper filters (Whatman No.5) on medium containing 50 μg/ml hygromycin, 100 μg/ml vancomycin, 400 μg/ml cefotaxim and, if indicated, 4-methyl-tryptophan (Sigma). In

contrast to the antibiotics, 4-mT was added to the medium before autoclaving. Individual calli were transferred onto new plates, and grown until they reached a diameter of approximately 1.5 cm. At that time, the calli were transferred to liquid selection medium (5 ml) and grown on a gyratory shaker. Depending on the growth of the cells, the culture volume was gradually increased to 20 ml. Further subculturing was performed weekly by transfer of 5 ml of cell culture into 20 ml of medium. The cell suspension cultures obtained showed some variation in characteristics (i.e., colour, growth rate, cell aggregate size).

RNA extraction and northern blot analysis

RNA extraction and northern blot analysis was performed as described (Menke et al., 1999), loading 10 μg of RNA onto the gels. Northern blots were probed with ^{32}P-labeled DNA probes for *Tdc*, *Str* and *Rps9* (Menke et al., 1999).

Plasmid rescue and Southern blot analysis

Genomic DNA was isolated from cell suspension cultures of 4-mT-resistant cell lines by CsCl-EtBr-ultracentrifugation. For plasmid rescue, 10 μg of genomic DNA was digested using the enzymes indicated. Upon self-ligation of the restriction mixture, DNA was electroporated into *E. coli* strain NM554 and cells were subsequently selected on medium containing carbenicillin. Southern blot analysis was performed as described by Memelink et al. (1994) using the *gusA* gene as a probe.

Results

Construction of the T-DNA activation tagging vector

An activation tagging vector was constructed (Figure 1) that was similar to pPCVICEn4HPT (Kakimoto, 1996). The T-DNA of vector Tag-2B4A1 contains a strong enhancer of transcription flanking the right T-DNA border, consisting of two copies of the B domain and four copies of the A1 domain of the CaMV 35S promoter (van der Fits & Memelink, 1997). Since this CaMV 35S promoter derivative lacks a TATA box and ATG start codon, integration of this T-DNA in either orientation in the vicinity of a gene can activate expression from the endogenous core promoter of this gene.

A battery of unique restriction sites is present (Figure 1), which, together with complete *E. coli* pUC21 plasmid sequences, enables the isolation of putatively tagged genes from the mutant cell lines by plasmid rescue.

Determination of the 4-mT selection window

In order to establish the effective 4-mT concentration for selection of *C. roseus* cell lines with increased TDC enzyme activity, the 4-mT selection window was determined. *C. roseus* cell suspensions were transformed with the *Tdc* gene under control of the CaMV 35S promoter, and, as a negative control, with an empty vector carrying only the hygromycin resistance gene. Figure 2A shows the results after several weeks of selection at various 4-mT concentrations. Cells overexpressing the *Tdc* gene showed normal growth on medium containing 0.2 mM 4-mT, whereas growth of control cells was strongly inhibited at this 4-mT concentration.

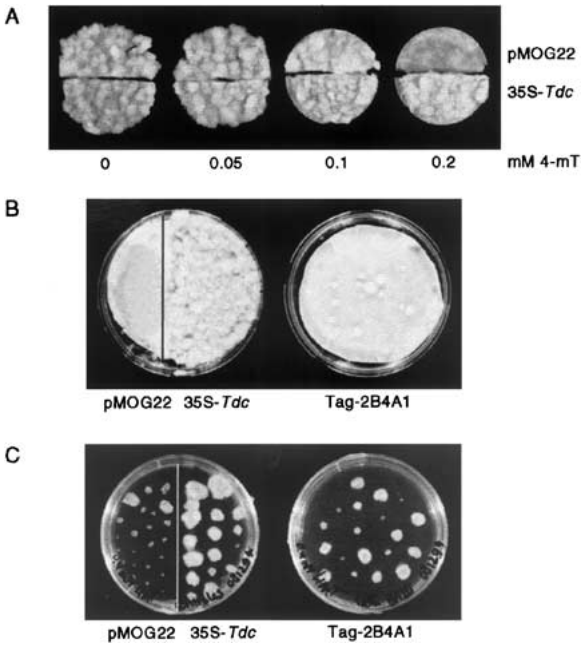

Figure 2. Selection of *C. roseus* cells with increased *Tdc* expression on 4-mT. (A) Cells were transformed with CaMV 35S-*Tdc* or with the empty vector pMOG22. Cells on filters were grown for several weeks on medium containing 0, 0.05, 0.1 and 0.2 mM 4-mT. (B) Cells, transformed with CaMV 35S-*Tdc*, the empty vector pMOG22 or with Tag-2B4A1 were selected on 0.4 mM 4-mT. (C) Calli obtained from filters as shown in Figure 2B were further selected on medium containing 0.4 mM 4-mT.

Table 1. *Tdc* and *Str* expression levels in tagged 4-mT-resistant *C. roseus* cell lines

Cell line	*Tdc* expression	*Str* expression	Cell line	*Tdc* expression	*Str* expression
46	++++	++++	II-39	+++	−
A6	++	−	II-55	++	−
A16	+++	+	IV-18	++++	−
A42	++++	+++	V-3	+++	−
A46	+++	−	V-8	++++	+++
B17	++++	−	V-9	+++	−
C18	+++	−	VI-1	++++	−
II-25	+++	−	VI-6	+++	++
II-35	+++	−	VI-8	++++	−
II-37	+++	+	VI-13	+++	−

Tdc or *Str* expression levels were analysed several times by northern blotting and the average level is depicted. The relative expression levels were quantified visually and indicated with + and −, ranging from highly increased expression (+ + ++) to normal expression (−).

Isolation of 4-mT-resistant cell lines by T-DNA activation tagging

As judged from the experiment shown in Figure 2A, a concentration of 0.2 mM 4-mT was initially used in the tagging experiments with vector Tag-2B4A1. However, the thickness of the cell layers on the filters appeared to be critical for cell survival. On 0.2 mM 4-mT, negative control cells often showed growth. Since most of this growth was inhibited by 0.4 mM 4-mT, and the *Tdc*-overexpressing cells still showed normal growth at this concentration, the cells transformed with the tagging vector were further selected on 0.4 mM 4-mT.

After 6–8 weeks of selection, actively growing calli that were putatively resistant to 4-mT appeared against a background of dead cells. An example of a typical tagging experiment is shown in Figure 2B. Occasionally calli arose on negative control filters of cells transformed with the empty vector, demonstrating that, at this stage, the 4-mT selection was not stringent. Growing calli were transferred onto new plates containing 0.4 mM 4-mT. At this stage 4-mT selection was more stringent, resulting in growth arrest of all negative control calli, whereas almost all 35S-*Tdc* calli showed continued growth (Figure 2C). Some of the calli transformed with Tag-2B4A1 stopped growing, suggesting that these escaped the 4-mT selection in the first round, whereas other calli showed vigorous growth, demonstrating that these were 4-mT-resistant (Figure 2C). In total, six independent tagging experiments were performed, that generated an estimated total number of stable Tag-2B4A1 transformants of 400,000–500,000. These numbers were estimated based on stable transformation frequencies as calculated in van der Fits et al. (2000). After 4-mT selection, in total 281 growing calli were obtained. Roman numerals in the codes for the different resistant lines (see Table 1) refer to the independent tagging experiments from which these were isolated. The first tagging experiment was not labelled with a Roman numeral, and was subdivided into A, B and C groups. Control transformations with a T-DNA without 35S enhancer elements next to the right border gave a negligible amount of resistant calli, indicating that the large majority of actively growing calli obtained after transformation with vector Tag-2B4A1 arose from actual T-DNA activation tagging events.

Identification of cell lines with increased Tdc and Str expression levels

Resistance for 4-mT can be caused by different mechanisms. For example, increased tryptophan production, impaired 4-mT uptake or increased TDC protein stability can result in tolerance for 4-mT. However, for the isolation of regulators of *Tdc* expression, we were only interested in cell lines showing high *Tdc* expression. To obtain sufficient biomass for subsequent analysis, suspension cultures resistant to 0.2 mM 4-mT were generated. This was successful for 180 of the 281 4-mT-resistant calli. Northern blot analysis of RNA extracted from these suspensions showed some variation in *Tdc* expression. Most 4-mT-resistant cultures showed a similar *Tdc* expression as found in

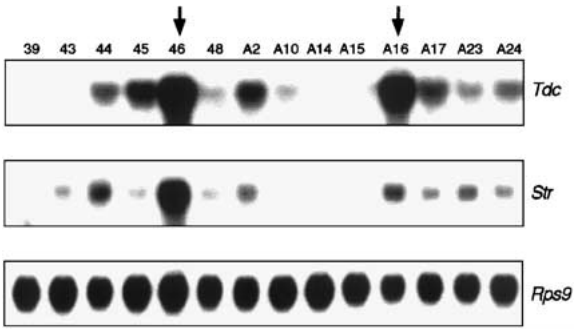

Figure 3. Identification of 4-mT-resistant lines with increased TIA biosynthetic gene expression via northern blot analysis. An example of a northern blot containing RNA extracted from 4-mT-resistant suspension cultures was probed with *Tdc*, *Str* and *Rps9*. Numbers above the lanes indicate the different 4-mT-resistant cell lines. Cell lines with significantly increased *Tdc* expression are marked with arrows.

untransformed control cell cultures (data not shown). However, certain cultures showed a clearly increased *Tdc* mRNA level. In the example shown in Figure 3, *Tdc* expression in lines 46 and A16 was significantly increased compared to the expression in other cell lines. Analysis of the level of *Rps9* mRNA, encoding the 40S ribosomal protein S9, showed equal loading of RNA. Analysis of cell cultures with putatively increased *Tdc* expression after several more subculture periods revealed that the *Tdc* expression levels were not always reproducible. In total, 20 lines reproducibly showed high *Tdc* expression (Table 1), suggesting that the observed 4-mT resistance in these lines was caused by *Tdc* overexpression.

To test whether a central regulator of TIA biosynthetic genes was activated by the T-DNA tag, additional northern blot analysis was performed using as a probe the *Str* gene, encoding the TIA biosynthetic enzyme strictosidine synthase (STR). This gene was not used as a selectable marker in the initial screen. In the example shown in Figure 3, only line 46 showed increased *Str* mRNA accumulation, compared to the other cell lines. In total, in three out of 20 cell lines, *Str* expression was significantly increased, whereas in another three lines a slight elevation of *Str* expression was observed (Table 1).

Further characterisation by plasmid rescue and sequence analysis

Some cell lines with increased *Tdc* expression (Table 1) were selected for isolation and analysis of the tagged genes. An overview of the molecular characterisation of these lines is shown in Table 2. Southern blot analysis indicated that T-DNA copy number ranged from 1 to 5. The majority of the lines contained a single copy.

Line 46 reproducibly showed high expression levels of *Tdc* and *Str*. Southern blot analysis suggested that one T-DNA copy was inserted in the genome of this cell line (data not shown). Plasmid rescue with the restriction enzyme *Xba*I, which is unique in the T-DNA tag, resulted in isolation of 1.6 kb of T-DNA flanking plant DNA. Sequence analysis revealed the presence of an open reading frame (ORF) encoding ORCA3, a transcription factor of the AP2/ERF-domain family (Riechmann & Meyerowitz, 1998). A detailed analysis of the phenotype of line 46 and characterisation of the tagged gene has been described previously (van der Fits & Memelink, 2000, 2001).

Repeated RNA analysis of line A42 showed reproducibly high *Tdc* expression (Table 1). However, variation was observed for *Str* expression, ranging from very high to normal expression levels. Southern blotting indicated that one T-DNA copy was integrated (data not shown). Plasmid rescue with *Xba*I resulted in the isolation of 4.8 kb of flanking plant DNA. Sequence comparisons with *Arabidopsis thaliana* genomic and EST databases showed that on this rescued DNA a gene is located that encodes a protein with similarity to an *Arabidopsis* protein of unknown function (data not shown). Further similarity with a protein encoded by an expressed sequence tag from rice confirms that this gene is conserved in different plant species and is actively transcribed. Further analysis should reveal whether the phenotype of cell line A42 is caused by overexpression of this gene.

In line B17 *Tdc* expression was increased, whereas *Str* expression was normal (Table 1). For this line, Southern blot analysis was not performed. Plasmid rescue with *Xba*I yielded 11 independent clones all containing a 2 kb fragment of flanking plant DNA, suggesting that this mutant cell line contained a single T-DNA copy. The isolated flanking DNA did not contain ORFs and did not show homology with sequences in databases.

Line II-37 contained a single T-DNA copy (data not shown). Plasmid rescue with *Xba*I recovered 0.6 kb of flanking DNA. This DNA consisted mainly of heavily re-arranged fragments derived from the activation vector, including parts of the B domain of the CaMV 35S promoter, the *nos* terminator, and an *is1* insertion element occurring in the *nptIII* gene of the pBIN19-derived binary vector. This indicated that DNA duplication and recombination events took place

Table 2. Summary of the molecular characterisation of interesting mutant cell lines

Cell line	# copies Southern[a]	# copies rescued[b]	Size rescued (kb)[c]	Genes[d]
46	1	1	1.6	*Orca3*
A16	3	ND	ND	ND
A42	1	1	4.8	Unknown gene[e]
A46	Inconclusive	ND	ND	ND
B17	ND	1	2.0	No homology
II-37	1	1	0.2	No homology
V-8	1	ND	ND	ND
VI-1	5	3	0.2/0.9/8.5	ND
VI-6	2	ND	ND	ND

[a]The number of T-DNA copies present in the different cell lines as determined by Southern blot analysis.
[b]The number of different sequences of flanking plant DNA isolated by plasmid rescue.
[c]The size of the flanking plant DNA isolated by plasmid rescue.
[d] Genes that are present on the rescued plant DNA.
[e]Sequence encoding a protein with similarity to a protein from *A. thaliana* with unknown function.
ND: not determined.

during T-DNA insertion in this cell line. Flanking the *Xba*I site was a DNA sequence of 190 bp without homology to sequences in the database, which could be a segment of *C.roseus* genomic DNA flanking the T-DNA insertion site.

Plasmid rescue and sequence analysis of lines VI-1 and VI-8 revealed that they originated from the same primary transformant, since identical sequences of flanking plant DNA were rescued from both lines. Probably these cell lines were generated by splitting the original callus during the initial selection stage. The T-DNA copy number of these lines was estimated at five, based on Southern blot analysis (data not shown). By plasmid rescue with *Xba*I, three different T-DNA flanking plant sequences were cloned, with sizes of 200 bp, 900 bp and 8500 bp. These were not further analysed, since the presence of multiple T-DNAs complicates functional analysis of the rescued sequences.

Discussion

Genetic studies in plants are predominantly performed using the model plant species *Arabidopsis thaliana*. Among the many advantages of using this particular species is the ability to generate large numbers of independently transformed plants. This allows extensive searches for mutants that are generated using classical mutagenesis (e.g. by chemical agents, T-DNA or transposon tagging) or by T-DNA activation tagging (Weigel et al., 2000). Here we show that T-DNA activation tagging can also be applied to genetically non-tractable plant species such as *C.roseus*. One of the main prerequisites is the ability to generate sufficiently high numbers of transformants to screen for the desired phenotype. We achieved this by using an improved *Agrobacterium tumefaciens* strain that is capable of very efficient transformation of *C.roseus* cell suspension cultures (van der Fits et al., 2000). Furthermore, the use of suspension cultures enabled us to screen hundreds of thousands of transformed cells with relative ease. This is in contrast to the around 10-fold lower number of *A. thaliana* T-DNA activation-tagged plants that were screened for altered phenotypes (Weigel et al., 2000), which is a laborious and time-consuming activity.

A summary of the selection strategy followed, together with the numbers of mutants isolated, is depicted in Figure 4. Transformed cells were initially selected for resistance against 4-mT. The concentration of 4-mT which was effective for selection of cell lines with increased *Tdc* expression in this study (0.4 mM) is in the same range as concentrations used

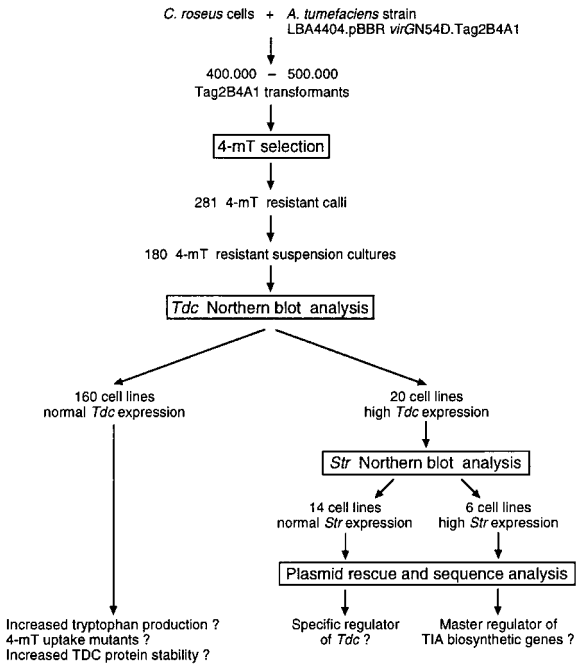

Figure 4. Schematic representation of the selection procedure and the numbers of transformants and mutants that were generated. Possible mechanisms for passing the specific selection steps are indicated at the bottom of the figure.

previously (0.05–1.0 mM) (Sasse et al., 1983b; Berlin et al., 1987; Goddijn et al., 1993).

Out of 180 4-mT-resistant cell suspensions tested, 20 showed increased *Tdc* expression. In some of the remaining 160 cell lines, 4-mT resistance might be caused by overproduction of tryptophan. In *C. roseus*, a correlation has been observed between toxicity of a tryptophan analogue and its inhibiting effect on anthranilate synthase (AS), an enzyme involved in tryptophan biosynthesis (Sasse et al., 1983a). It was suggested that false feedback inhibition, preventing synthesis of L-tryptophan for protein biosynthesis, is the important factor for the toxic action of the tryptophan analogues. Therefore, selection of cell lines with a toxic tryptophan analogue will probably result in isolation of lines with increased levels of endogenous tryptophan and an altered AS. Cells resistant to 5- or α-methyl-tryptophan often show altered feedback control of AS and, as a result, the cells contain high levels of tryptophan (Wakasa & Widholm, 1987; Kreps & Town, 1992; Kang & Kameya, 1993; Tam et al., 1995). In contrast to 5- and α-methyl-tryptophan, 4-methyl-tryptophan can act as substrate for TDC (Sasse et al., 1983a). Therefore, this particular tryptophan analogue can be used to select for cell lines with increased TDC activity (Sasse et al., 1983b; Berlin et al., 1987; Goddijn et al., 1993). All *C. roseus* cell cultures that were selected on 4-mT by Sasse et al. (1983b) contained, besides enhanced TDC activity, increased amounts of tryptophan and tryptamine. The cell cultures were cross-resistant to other toxic tryptophan analogues that cannot be metabolised by TDC, suggesting that the 4-mT resistance in these lines was not (only) due to enhanced detoxification by TDC.

We counterselected against such undesired mutants by performing additional screens for gene expression levels. Starting with 180 4-mT-resistant cell lines, the number of interesting cell lines was reduced to 20 by screening for high *Tdc* expression levels. Fourteen of these cell lines with increased *Tdc* expression showed normal *Str* expression. This suggests that *Tdc* and *Str* are not always coordinately regulated, but that specific regulators of individual genes exist. Only six lines showed increased expression of both genes. Apparently, the frequency of tagging components of a shared signal transduction pathway leading to regulation of several, or possibly all, TIA biosynthetic genes is low. Possibly, overexpression of multiple secondary metabolite biosynthetic genes, and the resulting production of secondary metabolites, is detrimental for the cells. Since the T-DNA activation tagging approach selects for actively growing cells, it can be speculated that there is negative selection pressure for TIA-producing cells. Alternatively, the number of target genes that act as master regulators for induction of multiple genes operating in the TIA biosynthetic pathway may be limited. The screening results also show the importance of successive selection steps, especially of counter-selection for 4-mT- resistant cell lines without increased *Tdc* expression. This reduced the amount of initial lines to a manageable number, and more importantly, selected those lines in which a master regulator of TIA biosynthetic gene expression was likely to be tagged.

A straightforward method to prove linkage between the activation tag and *Tdc* overexpression is analysis of the effects of retransformation of the rescued plasmid into wild type *C. roseus* cells (van der Fits & Memelink, 2000). A similar approach was followed by Furini et al. (1997) for analysis of a desiccation-tolerant *Craterostigma plantagineum* mutant that was obtained by T-DNA activation tagging. Introduction of a genomic clone containing the T-DNA/plant DNA junction of the tagged mutant into wild type *C. plantagineum* yielded desiccation-tolerant cells. However, sequence analysis of a cDNA

corresponding to the tagged genomic region in the desiccation-tolerant mutant did not reveal large ORFs nor homology to sequences in the database (Furini et al., 1997). The authors suggested that the phenotype may be caused by an unusual RNA. In analogy, the analysed rescued sequence from line B17 might encode unusual RNA. Alternatively, the tagged gene in this line might be located more distantly from the T-DNA border. In one of the *A. thaliana* cytokinin-independent mutants isolated by T-DNA activation tagging, the T-DNA activation tag was located as far as 4 kb downstream of the activated gene (Kakimoto, 1996). The T-DNA activation vector used here only contained enhancer sequences near the right border, lacking a TATA box and an ATG start codon. Therefore, using Tag-2B4A1, activation tagging of genes near the left border sequence may be possible. Plasmid rescue resulted only in isolation of plant DNA immediately flanking the right border sequence. Thus, for the identification of the tagged genes in some of the 4-mT resistant cell lines described here, it may be necessary to isolate DNA sequences flanking the left border. This can be achieved relatively easily by screening a genomic *C. roseus* library with the plant DNA isolated by plasmid rescue.

The haploid genome size of *C. roseus* is 8×10^8 bp according to a recent estimation (Frank van Iren, personal communication). This is 2.5-fold smaller than a previous estimation by Galbraith et al. (1983). Thus, by generation of 400,000–500,000 independent transformants with on average two T-DNA copies per genome (van der Fits et al., 2000), the probability of finding a T-DNA insert in a given 2 kb DNA region is around 90% using the formula $P = 1 - (1 - [x/800,000])^n$ (where P = probability; x = length of DNA region in kb; n = number of T-DNA insertions in population; and T-DNA integration is assumed to be random). Since T-DNA preferentially inserts into transcribed regions of the plant genome (Koncz et al., 1989), T-DNA density near or into genes may even be higher. Based on these calculations, it can be concluded that T-DNA activation tagging of the *C. roseus* genome by construct Tag-2B4A1 was close to saturation. Thus, there is a probability of around 90% that a mutant in which the endogenous *Tdc* gene was activated by the T-DNA tag, is present among the cell lines that have only high *Tdc* expression.

In this paper, it is shown that T-DNA activation tagging resulted in isolation of mutant cell lines with the desired phenotype from the genetically non-tractable plant species *C. roseus*. Further characterisation of the tagged gene from line 46 (van der Fits & Memelink, 2000, 2001) demonstrated that the activation tagging approach described in detail here, indeed resulted in the isolation of a regulator of TIA biosynthetic genes. In our strategy, the target gene (i.e. *Tdc*) was used immediately as a selection gene by the use of a toxic substrate. For several metabolic pathways, toxic substrates and inhibitors of biosynthetic enzymes have been described, and these enzymes can thus be used as targets in a screen for regulators of metabolism. For example, several toxic analogues of anthranilate, an intermediate in the tryptophan biosynthetic pathway, have been described and used in genetic studies (e.g., Li & Last, 1996). Therefore, use of these substrates in a T-DNA activation tagging approach might result in the isolation of regulators of tryptophan metabolism. For many enzymes operating in metabolic pathways however, toxic substrates are not available. To apply the T-DNA activation tagging approach for the isolation of transcriptional regulators of these enzymes, the promoter of the corresponding gene can be cloned in front of a known selectable marker (such as the hygromycin or kanamycin resistance gene). After determination of the appropriate concentration for selection of cells with increased expression of the selection gene, T-DNA activation tagging experiments can be performed. T-DNA activation tagging using a tagging construct with constitutively active enhancers lacking a TATA box sequence, will predominately yield positive regulators of the reporter gene. In addition, by replacing the selectable reporter gene (e.g., hygromycin or kanamycin resistance gene) by a negative selectable marker (e.g., cytosine deaminase (Schlaman & Hooykaas, 1997)), it should also be possible to identify repressors of gene expression using T-DNA activation tagging.

Acknowledgements

The authors thank Peter Hock for preparing the figures and Ward de Winter for assistance with tissue culturing. F. Hilliou was financially supported by Marie Curie research training grant MCFI-1999-01097 from the European Community, whereas L. van der Fits was financially supported by the ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry of Agriculture, Nature Management and Fishery in the framework of an industrial relevant research program of the Association of Biotechnology Centres in the Netherlands (ABON).

References

Aerts RJ, Gisi D, De Carolis E, De Luca V and Baumann TW (1994) Methyl jasmonate vapor increases the developmentally controlled synthesis of alkaloids in *Catharanthus* and *Cinchona* seedlings. *Plant J* **5**: 635–643.

Berlin J, Mollenschott C, Sasse F, Witte L, Piehl G-W and Büntemeyer H (1987) Restoration of serotonin biosynthesis in cell suspension cultures of *Peganum harmala* by selection for 4-methyltryptophan-tolerant cell lines. *J Plant Physiol* **131**: 225–236.

Furini A, Koncz C, Salamini F and Bartels D (1997) High level transcription of a member of a repeated gene family confers dehydration tolerance to callus tissue of *Craterostigma plantagineum*. *EMBO J* **16**: 3599–3608.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP and Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissue. *Science* **220**: 1049–1051.

Gantet P, Imbault N, Thiersault M and Doireau P (1998) Necessity of a functional octadecanoic pathway for indole alkaloid synthesis by *Catharanthus roseus* cell suspensions cultured in an auxin-starved medium. *Plant Cell Physiol* **39**: 220–225.

Geerlings A, Martinez-Lozano Ibanez M, Memelink J, van der Heijden R and Verpoorte R (2000) Molecular cloning and analysis of strictosidine β-D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *J Biol Chem* **275**: 3051–3056.

Goddijn OJM, van der Duyn-Schouten PM, Schilperoort RA and Hoge JHC (1993) A chimaeric tryptophan decarboxylase gene as a novel selectable marker in plants. *Plant Mol Biol* **22**: 907–912.

Kakimoto T (1996) CKI1, a histidine kinase homolog implicated in cytokinin signal transduction. *Science* **274**: 982–985.

Kang KK and Kameya T (1993) Selection and characterization of a 5-methyltryptophan-resistant mutant in *Zea mays* L. *Euphytica* **69**: 95–101.

Kardailsky I, Shukla V, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ and Weigel D (1999) Activation tagging of the floral inducer *FT*. *Science* **286**: 1962–1965.

Koncz C, Martini N, Mayerhofer R, Koncz-Kalman Z, Korber H, Redei GP and Schell J (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc Natl Acad Sci USA* **86**: 8467–8471.

Kreps JA and Town CD (1992) Isolation and characterization of a mutant of *Arabidopsis thaliana* resistant to alpha-methyltryptophan. *Plant Physiol* **99**: 269–275.

Li J and Last RL (1996) The *Arabidopsis thaliana* trp5 mutant has a feedback-resistant anthranilate synthase and elevated soluble tryptophan. *Plant Physiol* **110**: 51–59.

Marsh JC, Erfle M and Wykes EJ (1984) The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. *Gene* **32**: 481–485.

Memelink J, Swords KMM, Staehelin LA and Hoge JHC (1994) Southern, northern and western blot analysis. In: Gelvin SB, Schilperoort RA and Verma DPS (eds.), *Plant Mol Biol Manual*. (pp. F1–23) Kluwer Academic Publishers, Dordrecht.

Menke FLH, Parchmann S, Mueller MJ, Kijne JW and Memelink J (1999) Involvement of the octadecanoid pathway and protein phosphorylation in fungal elicitor-induced expression of terpenoid indole alkaloid biosynthetic genes in *Catharanthus roseus*. *Plant Physiol* **119**: 1289–1296.

Moreno PRH, van der Heijden R and Verpoorte R (1995) Cell and tissue cultures of *Catharanthus roseus*: a literature survey. *Plant Cell Tissue Org Cult* **42**: 1–25.

Neff MM, Nguyen SM, Malancharuvil EJ, Fujioka S, Noguchi T, Seto H, Tsubuki M, Honda T, Takatsuto S, Yoshida S and Chory J (1999) *BAS1*: a gene regulating brassinosteroid levels and light responsiveness in *Arabidopsis*. *Proc Natl Acad Sci USA* **96**: 15316–15323.

Offringa R (1992) Gene targeting in plants using the *Agrobacterium* vector system. PhD thesis, Leiden University, The Netherlands.

Pasquali G, Goddijn OJM, de Waal A, Verpoorte R, Schilperoort RA, Hoge JHC and Memelink J (1992) Coordinated regulation of two indole alkaloid biosynthetic genes from *Catharanthus roseus* by auxin and elicitors. *Plant Mol Biol* **18**: 1121–1131.

Riechmann JL and Meyerowitz EM (1998) The AP2/EREBP family of plant transcription factors. *Biol Chem* **379**: 633–646.

Sasse F, Buchholz M and Berlin J (1983a) Site of action of growth inhibitory tryptophan analogues in *Catharanthus roseus* cell suspension cultures. *Z Naturforsch* **38**c: 910–915.

Sasse F, Buchholz M and Berlin J (1983b) Selection of cell lines of *Catharanthus roseus* with increased tryptophan decarboxylase activity. *Z Naturforsch* **38**c: 916–922.

Schlaman HRM and Hooykaas PJJ (1997) Effectiveness of the bacterial *codA* encoding cytosine deaminase as a negative selectable marker in *Agrobacterium*-mediated plant transformation. *Plant J* **11**: 1377–1385.

Tam YY, Slovin JP and Cohen JD(1995) Selection and characterization of α-methyltryptophan-resistant lines of *Lemna gibba* showing a rapid rate of indole-3-acetic acid turnover. *Plant Physiol* **107**: 77–85.

van der Fits L and Memelink J (1997) Comparison of the activities of CaMV 35S and FMV 34S promoter derivatives in *Catharanthus roseus* cells transiently and stably transformed by particle bombardment. *Plant Mol Biol* **33**: 943–946.

van der Fits L, Deakin EA, Hoge JHC and Memelink J (2000) The ternary transformation system: Constitutive *virG* on a compatible plasmid dramatically increases *Agrobacterium*-mediated plant transformation. *Plant Mol Biol* **43**: 495–502.

van der Fits L and Memelink J (2000) ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* **289**: 295–297.

van der Fits L and Memelink J (2001) The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element. *Plant J* **25**: 43–53.

Vazquez-Flota FA and De Luca V (1998) Jasmonate modulates development- and light-regulated alkaloid biosynthesis in *Catharanthus roseus*. *Phytochem* **49**: 395–402.

Vieira J and Messing J (1991) New pUC-derived cloning vectors with different selectable markers and DNA replication origins. *Gene* **100**: 189–194.

Wakasa K and Widholm JM (1987) A 5-methyltryptophan resistant rice mutant, MTR1, selected in tissue culture. *Theor Appl Genet* **74**: 49–54.

Weigel D, Hoon Ahn J, Blázquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrándiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Thuy Nguyen J, Sato S, Wang Z-Y, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF and Chory J (2000) Activation tagging in *Arabidopsis*. *Plant Physiol* **122**: 1003–1013.

Zheng Z, Hayashimoto A, Li Z and Murai N (1991) Hygromycin resistance gene cassettes for vector construction and selection of transformed rice protoplasts. *Plant Physiol* **97**: 832–835.