FEBS Letters 579 (2005) 4622–4628 FEBS 29826

A transposon-based activation-tagging population in *Arabidopsis thaliana* (TAMARA) and its application in the identification of dominant developmental and metabolic mutations

Anja Schneider[a,1], Thomas Kirch[b,1], Tamara Gigolashvili[a], Hans-Peter Mock[c], Uwe Sonnewald[d], Rüdiger Simon[e,*], Ulf-Ingo Flügge[a,*], Wolfgang Werr[b,*]

[a] *Botanisches Institut, Universität zu Köln, Gyrhofstr. 15, D-50931 Köln, Germany*
[b] *Institut für Entwicklungsbiologie, Universität zu Köln, Gyrhofstr. 17, D-50931 Köln, Germany*
[c] *Institut für Pflanzengenetik und Kulturpflanzenforschung, Corrensstrasse 3, D-06466 Gatersleben, Germany*
[d] *Lehrstuhl für Biochemie, Friedrich-Alexander Universität Erlangen-Nürnberg, Staudtstrasse 5, D-91058 Erlangen, Germany*
[e] *Institut für Genetik, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, D-40225 Düsseldorf, Germany*

Received 14 June 2005; revised 15 July 2005; accepted 16 July 2005

Available online 8 August 2005

Edited by Micheal Sussman

Abstract A population of 9471 stable activation-tagged lines was generated by transposable element mediated activation tagging mutagenesis in Arabidopsis (TAMARA) using the maize *En/Spm* transposon system. Based on DNA gel blot and flanking sequence analysis, this population contains approximately 6000 independent transposon insertions. A greenhouse-based screen identified six dominant or semi-dominant activation tagged mutants with obvious developmental alterations, among these a new *pistillata* mutant allele. In addition, a subset of 1500 lines was screened by a HPLC based high-throughput method for dominant activation tagged mutants with enhanced contents of phenolic compounds. One dominant activation tagged mutant (*hpc1-1D*) was isolated showing accumulation of a particular compound due to the upregulation of an R2R3-MYB transcription factor.

Keywords: Transposon; Activation tagging; Dominant mutation; Shoot development; Phenylpropanoid metabolism

1. Introduction

The screening for loss-of-function mutants has been a primary tool for dissecting genetic pathways in many organisms, including Arabidopsis. However, a limitation of this approach is that redundantly acting genes can hardly be identified. Sequencing of the Arabidopsis genome revealed that about 70% of the genome is made up of duplicated loci [1]. To overcome this type of genetic redundancy, gain-of-function mutants can be studied. Gain-of-function phenotypes can either be caused by mutations in the coding region that lead to constitutive activation of the corresponding protein, as is the case for dominant ethylene response mutants [2] or by activation tagging that leads to altered levels or patterns of gene expression.

A transfer DNA (T-DNA) vector, that possesses four copies of an enhancer element of the constitutively active promoter of the Cauliflower mosaic virus (CaMV) 35S gene [3] was constructed by Walden and colleagues [4]. These enhancers can cause transcriptional activation of adjacent genes, and, because activated genes will be associated with a T-DNA insertion, the method is called activation tagging. In Arabidopsis, large collections of such T-DNA lines have been described [5,6], which contain numerous gain-of-function alleles.

T-DNA insertions, however, are often complex and characterized by multiple inverted or tandem copies or truncated T-DNA inserts, which often make a molecular analysis difficult [7,8]. One possibility to overcome this problem would be the use of single copy transposon insertions as the activation tags. In addition, transposon insertions can be remobilized germinally, thereby producing revertants that can confirm the phenotypic consequences of the insertion.

The enhancer–inhibitor (*En–I*), also known as suppressor–mutator (*Spm–dSpm*), system of maize is an efficient tool for heterologous transposon tagging in Arabidopsis [9–13]. *En/Spm* elements transpose at a high frequency to unlinked locations, without a known bias towards specific regions in the genome.

The use of positive and negative selectable markers has further improved transposon technology [12,13]. The system makes use of the positive selectable marker *bar* [14] conferring resistance to the herbicide Basta, as well as the negative selectable marker *SU1* [15] that converts the pro-herbicide R7402 (DuPont) into the herbicide sulfonylurea. Both markers are carried on a single construct: the *bar* gene is placed inside the nonautonomous *I/dSpm* element, the *SU1* gene and an immobile source of transposase (an *En/Spm* element that lacks the terminal-inverted repeats) in the adjacent T-DNA. Application of R7402 allows selection against the transposase source and subsequent application of Basta allows selection for stable transposed *I/dSpm* elements [12,13]. If, in addition, a tetramer of the CaMV 35S enhancer is placed inside the *I/dSpm* element, the system can also be used to generate activation tagged inserts [13].

*Corresponding authors. Fax: +49 221 470 5164.
E-mail addresses: Anja.Schneider@uni-koeln.de (A. Schneider), ruediger.simon@uni-duesseldorf.de (R. Simon), ui.fluegge@uni-koeln.de (U.-I. Flügge), werr@uni-koeln.dex (W. Werr).

[1] These authors contributed equally to this work.

Abbreviations: HPLC, high performance liquid chromatography; T-DNA, transfer DNA; CaMV, Cauliflower mosaic virus; *En–I*, enhancer–inhibitor; *Spm–dSpm*, suppressor–mutator

0014-5793/$30.00
doi:10.1016/j.febslet.2005.07.030

In the present study, we have used the two-element transposon system to generate an activation tagged population in Arabidopsis (transposable element mediated activation tagging mutagenesis in Arabidopsis, TAMARA), which will become publicly available through the Nottingham Arabidopsis Stock Centre. This population was used to identify and isolate new Arabidopsis gain-of-function alleles affected in shoot development. In addition, an high performance liquid chromatography (HPLC)-based screening was established to identify activation tagged mutants with altered contents of phenolic substances. A first characterization of such a mutant (*hpc1-1D*) is described.

2. Materials and methods

2.1. TAMARA construct

The counterselectable marker gene *SSU-SU1* was isolated from the plasmid SLJ8241 [12,15] as a *Eco*RI + *Bam*HI fragment and ligated into the T-DNA vector pGPTV [16]. Replacement of the *Eco*RI site by an *AscI site* resulted in plasmid pGPTV-SSU. Plasmid SLJ7713 [12] was digested with *Sma*I + *Xho*I to isolate the transposase gene, which was then inserted between *Bam*HI and *Xho*I sites of pRT-Ω/Not/Asc [17] to yield pRT-TNP. Subsequently, the transposase cassette with the 35S promoter- and terminator sequences was excised from pRT-TNP using the flanking *Asc*I.

The Basta resistance gene *bar* was obtained from vector SLJ0512 [12] as *Bcl*I fragment and transferred into the *Pst*I inside the *dSpm* element in SLJ7648 [12]. The resulting intermediate was digested with *Hind*III to insert the 35S enhancer tetramer isolated as a *Bam*HI and *Bgl*II fragment from pPVICEn4HPT [4]. The final *dSpm* element was excised with *Cla*I and cloned into the *Bam*HI site of pGPTV-SSU after fill-in reactions. To complete construction of the TAMARA plasmid, the *Asc*I fragment from pRT-TNP was inserted into the *Asc*I site of pGPTV-SSU. The TAMARA construct was introduced into wild-type plants (*Arabidopsis thaliana* ecotype Col-0) by vacuum infiltration [18] using *Agrobacterium tumefaciens* strain GV3101.

2.2. Generation of stable activation tagged lines (TAMARA population)

For amplification of TAMARA starter lines, the progeny (T_2) of primary transformants (Basta resistant T_1 plants) were sown at a density of 500 plants per tray (45 × 30 cm). Only lines, which segregated approximately 75% Basta resistant seedlings, were selected as starter lines. Seeds from individual starter lines were harvested as a pool (T_3). To select for transposition events, T_3 plants were sown on soil in a temperature-controlled greenhouse at a density of 2500 plants per tray under a 16 h light/8 h dark cycle. Seedlings were sprayed with a solution of 75 μg L^{-1} R7402 (DuPont) 6 days after germination, and 14 days after germination with a solution containing 150 μg L^{-1} R7402 and 0.5 mL L^{-1} Basta (Hoechst, which contains 200 g L^{-1} glufosinate ammonium). Double resistant plants (transposants) were transferred to 77 well trays to score for developmental phenotypes and to collect seeds of single plants (T_4). In total, seeds of 9471 single plants were obtained and represent the TAMARA population.

2.3. Isolation of genomic DNA, Southern hybridisation and isolation of dSpm-Act flanking sequences

Genomic DNA was isolated as described previously [19]. For Southern blot analysis, about 10 μg of genomic DNA was digested with *Pst*I or *Eco*RI, separated by electrophoresis on 0.8% TBE agarose gels and blotted onto Hybond N^+ nylon membrane (Amersham). For hybridisation, fragments specific for the *bar* gene (1.6 kb *Bam*HI/*Bcl*I from SLJ0512) or the 5′ *dSpm* element (1 kb *Pst*I/*Cla*I from SLJ7648) were labelled in the presence of [α-^{32}P]-dCTP. Isolation of *dSpm*-Act flanking sequences was done as described previously [20] or by inverse PCR. For inverse PCR, genomic DNA (0.5 μg) was digested with four-cutting enzymes. Fragments were phenol/chloroform extracted and circularized overnight at 16 °C in a total volume of 300 μL with 5 U of T4 DNA ligase. After phenol/chloroform extraction, the ligation mix was resuspended in 10 μL of H_2O. PCR was performed in a total volume of 25 μL containing 1 μL of the above solution, 0.2 U Biotherm polymerase, and 50 pmol of each primer. PCR conditions were: 5 min, 94 °C; 40× [30 s, 94 °C; 45 s, 60 °C; 2 min, 72 °C]; 5 min, 72 °C. Flanking sequences were obtained with primer combinations as follows: 3′*dSpm* – for digestions with Sau3a I or Asn I: 5′-AGTCCATACAAAACGCAATCATAG-3′ and 5′-CTTAGAGTGTCGGCTTATTTCAGT-3′, for Rsa I: 5′-GGACCGACGCTCTTATGTTAAAAG-3′ and 5′-CAGTAAGAGTGTGGGGTTTTGG-3′; 5′*dSpm*-digestions with Sau3a I or Asn I: 5′-GCACGACGGCTGTAGAATAGG-3′ and 5′-CAAGAAGTCAAAACGCTATGTGG-3′, Rsa I: 5′-GCACGACGGCTGTAGAATAGG-3′ and 5′-CGCGCACCTCCAAGTAGC-3′. The resulting PCR products were subcloned into pCR II-TOPO (In Vitrogen). Flanking sequences were compared to the Arabidopsis genome database using the BLAST algorithm [21] at http://www.ncbi.nlm.nih.gov/ to determine the position and orientation of the *dSpm* insertion.

2.4. Extraction of metabolites and HPLC conditions

For the identification of metabolic mutants, a set of 1500 lines of the TAMARA population was screened by HPLC. Approximately 20 seeds per individual T_4 line were grown in a temperature controlled greenhouse and sprayed 7 and 10 days after germination with 1.25 mL L^{-1} Basta (Hoechst). 50 mg of leaf material was collected and extracted repeatedly with methanol. Phenolic constituents were analyzed by a HPLC system as described previously [22], consisting of a photodiode array detector (Waters, Eschborn, Germany) combined with a fluorescence detector (FP-920, JASCO, Groß-Umstadt, Germany). Data acquisition and processing were performed with the Millenium software (Waters, Eschborn, Germany). Spectra from 250 to 400 nm were recorded and chromatograms of absorbance at 280 nm were extracted from the data sets. The wavelength settings for the fluorescence detector were 300 nm for excitation and 400 nm for emission. The identification of major peaks in the UV chromatograms (280 nm) in methanolic extracts of Arabidopsis leaves was done by comparing their retention time and UV light spectra (250–400 nm) with those of known standards.

2.5. Expression analysis

Total RNA from mutant and wild-type leaves was isolated using the RNeasy Plant Kit (Qiagen). Oligo (dT)-primed cDNA from 1 μg of total RNA was synthesized using the SuperScript Reverse Transcriptase system (InVitrogen). 2 μl of the total Reverse Transcriptase reaction was used to perform PCR with gene-specific primers. For the *MYB 51* gene, primers used were 5′-GTGTTGCAAAGCTGAACTAGGGTT-3′ and 5′-TTGTTAACGGAGGAATCAGAG-AAC-3′; for *At1g18560*, primers were 5′-TGTGCTTGATTGGTGGAAGGTAA-3′ and 5′-TTCTAAACCTCCGGCAGAATTATC-3′ and the primers used for *ACTIN 1* were 5′-TAACTCTCCCGCTATGTATGTCGC-3′ and 5′-CCACTGAGCACAATGTTACCGTAC-3′. PCR conditions were 2 min, 94 °C; 26 or 40 cycles of [30 s, 94 °C; 45 s, 60 °C; 2 min, 72 °C]; 5 min, 72 °C.

2.6. Construction of 35S-MYB51 plants

The full-length cDNA for *MYB51* was amplified using 5′-CACCATGGTGCGGACACCGTGTTGCAAAGC-3′ and 5′-TCATCCAAAA-TAGTTATCAATTTCGTC-3′ as described above and subcloned into pENTR/D TOPO vector (InVitrogene). Subsequently, the cDNA was transferred into the GATEWAY destination vector pGWB2 by an LR Clonase reaction (InVitrogene). The pGWB2 vector containing the CAMV 35S promotor, was provided by Dr. Tuyoshi Nakagawa (Shimane University, Japan). The 35S-MYB51 construct was introduced into wild-type plants (*A. thaliana* ecotype Col-0) by vacuum infiltration [18] using *Agrobacterium tumefaciens* strain GV3101. Of the resulting three Kanamycin resistant lines, one line overexpressed the *MYB51* gene and was further investigated by HPLC analysis.

3. Results

3.1. Generation and establishment of an activation tagged population (TAMARA)

All elements required for the TAMARA system were integrated into a single T-DNA and introduced into Arabidopsis

(ecotype Col-0) via *Agrobacterium* vacuum infiltration. The T-DNA contains three essential components (Fig. 1A): (i) a modified defective *Spm* element (*dSpm*-Act) that carries four copies of the CaMV35S transcriptional enhancer element (base pairs −420 to −90) and a Basta resistance gene (*bar*) to select for T-DNA integration and transposon insertions; (ii) a transposase gene under transcriptional control of the CaMV35S promoter; and (iii) the *SU1* gene from *Streptomyces griseolus* that confers sensitivity to the proherbicide R7402 [15], allowing to select against the presence of the T-DNA. We transformed the TAMARA T-DNA into Arabidopsis and selected 67 Basta-resistant primary transformants. Of these 67 primary transformants, 57 lines were selected as starter lines, because they segregated 75% Basta resistance, indicating a single-locus T-DNA insertion, and expressed the *SU1* gene, i.e., the phenotype was darker green and bushier than wild-type [13]. In a pilot experiment, the frequency of transposants and the frequency of independent transpositions were determined in the progeny of each starter line. Approximately, 1000 seedlings per starter line were sown in the greenhouse and the double selected survivors counted. The frequency of transposants varied from 0.1% to 2.4%. To estimate the frequency of independent transpositions, genomic DNA of individual transposants from 23 starter lines were analyzed by Southern blotting. In 226 plants, a total of 245 *dSpm*-Act elements at 137 locations could be identified in the genome. Thus, more than 90% of all transposants carried a single *dSpm*-Act insertion. The frequency of independent insertions (unique insertions/total of insertions) varied from 5% to 89% between starter lines, with an average of 56% (137/245). For the large-scale production of transposants (TAMARA population), starter lines were selected that showed an independent transposition frequency of at least 63%. Progeny of the T_2 generation plants were germinated on soil and stable transposants were selected as described in material and methods. From a total of 1.6 million seedlings, 9471 double resistant plants, which are calculated to contain 6000 independent insertions, were isolated.

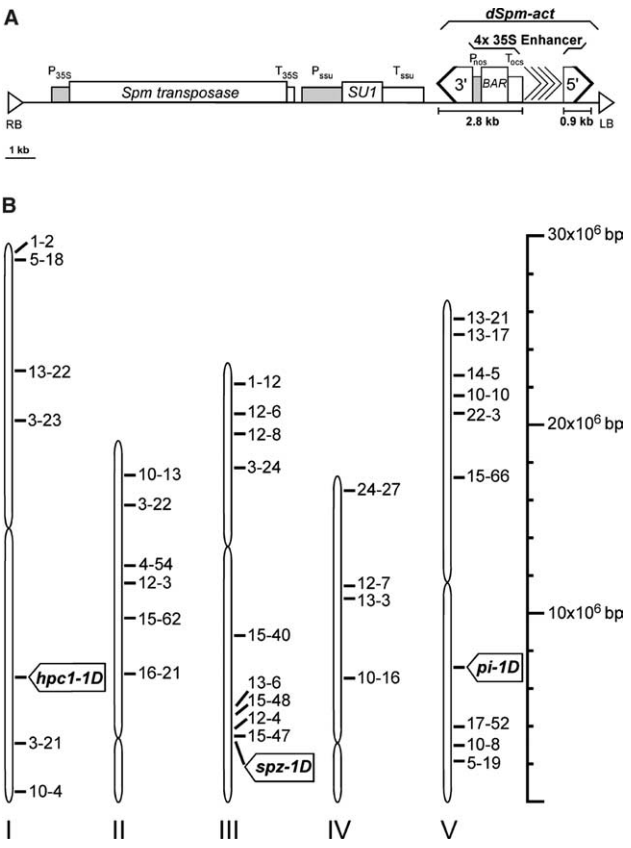

Fig. 1. Schematic representation of the TAMARA construct (A) and of the Arabidopsis chromosomes (1n) carrying transposed *dSpm*-Act elements at distributed positions (B). (A) Expression of the *Spm Transposase* is controlled by the CaMV 35S promoter and terminator (P_{35S}/T_{35S}). Expression of the counterselectable marker gene *SU1* is controlled by the *Rubisco Small Subunit* promoter and terminator (P_{SSU}/T_{SSU}). Expression of the *Phosphinotricine Transferase* is controlled by the *Nopalin Synthase* promoter and the *Octopine Synthase* terminator (P_{nos}/T_{ocs}). The 4 × 35S enhancer and the *dSpm*-Act element are indicated. (B) The insertion sites of 34 randomly chosen *dSpm*-Act elements and of *dSpm*-Act elements in *hpc1-1D*, *spz-1D* and *pi-1D* are indicated.

3.2. Distribution of transposed dSpm-Act elements

In order to confirm the frequency of independent transposition of at least 63% in the TAMARA population and to determine the distribution of transposed *dSpm*-Act elements in the genome, genomic DNA flanking the transposon insertion in 50 randomly chosen lines that originated from 13 different starter lines were isolated and sequenced (Fig. 1B). These lines represented 34 individual insertion events and thus confirmed the estimated frequency of at least 63% unlinked transposition events (34/50 = 68%). *dSpm*-Act insertions were found on all of the five Arabidopsis chromosomes with no apparent preference for certain regions, except for a potential hotspot on the long arm of chromosome III (Fig. 1B). In addition, centromeric regions were devoid of *dSpm*-Act insertions. Taken these data together, the TAMARA population represents a novel population of approximately 6000 independent activation tagged lines.

3.3. Identification of dominant developmental mutants

The T_3 generation of the population is expected to be hemizygous for the *dSpm*-Act element. This generation was screened for dominant mutations resulting in developmental alterations, and led to the isolation of eight individuals. To test if the respective mutations were caused by the *dSpm*-Act element, backcrosses against wild-type were performed. The progenies were analysed for mutant phenotypes and Basta resistance conferred by the *dSpm*-Act element. Two of the mutations were not dominantly inherited, and the mutant phenotype did not cosegregate with the Basta resistance. The remaining six mutations cosegregated with a *dSpm*-Act element, five were dominant mutations and one mutation was semi-dominant. Considering the frequency of independent insertions (at least 63%) within the population, approximately 0.1% of all plants showed a visible dominant developmental phenotype.

3.4. Phenotypic characterization of developmental mutants

One of the dominant mutants – *pi-1D* – showed a homeotic transformation of floral organs. The apical tip of outer whorl organs exhibited a partial conversion of sepals to petals (Fig. 2A). Cells at the base and margins showed sepals identity while cells in the upper central region were coloured white, characteristic of anthocyanine biosynthesis in petals. Southern analysis of genomic DNA of *pi-1D* mutants initially

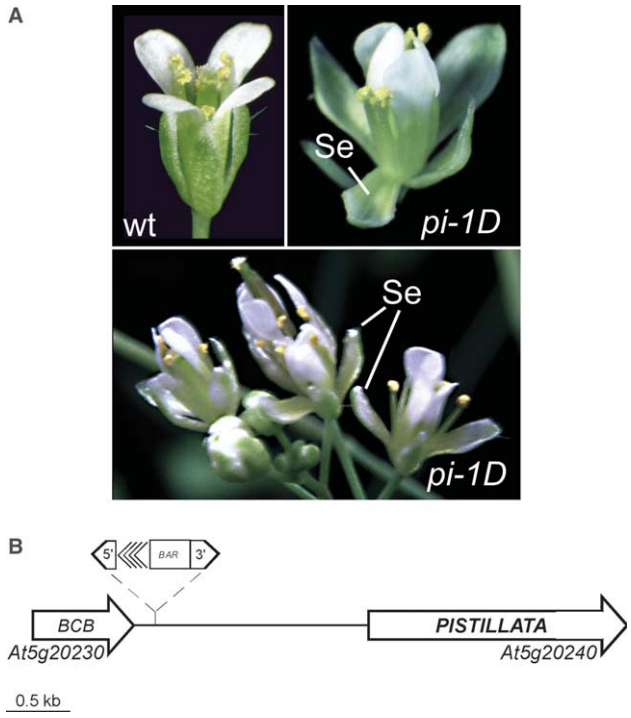

Fig. 2. Characterization of the *pi-1D* mutant. (A) Flower phenotype of *pi-1D* in comparison to the wild-type flower phenotype. Note the petaloid character of the sepals (Se). (B) Schematic representation of the localization and orientation of *dSpm*-Act in *pi-1D*.

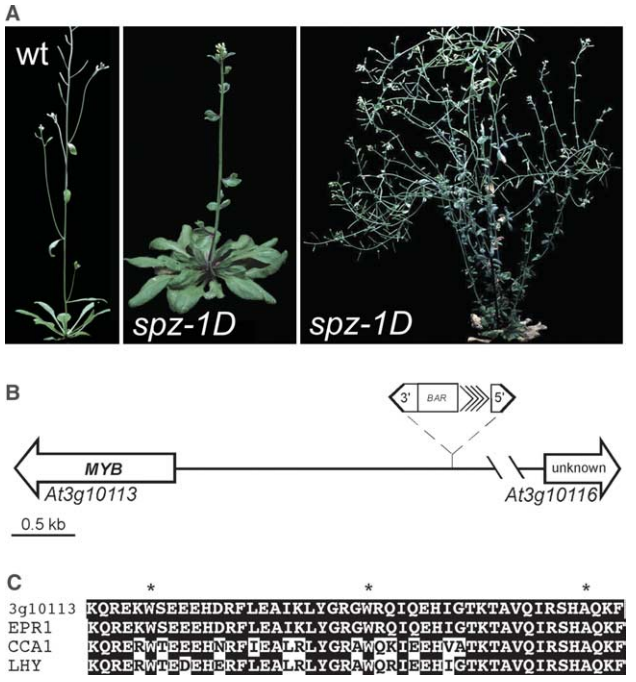

Fig. 3. Characterization of the *spz-1D* mutant. (A) Arabidopsis lines were grown under a 16 h/8 h light/dark cycle and pictures were taken from six weeks old plants (left and middle panel) and from a 16-week old plant (right panel). (B) Schematic representation of the localisation and orientation of *dSpm*-Act in *spz-1D*. (C) Comparison of the predicted MYB domain of the product of *At3g10113* (amino acids 63–109) and those of EPR1 (amino acids 48–94), CCA1 (amino acids 22–68) and LHY (amino acids 22–68). Identical amino acid residues are boxed, asterisks represent the position of the conserved Trp and Ala residues in the MYB domain. The accession numbers of the sequences are EPR1 (AB115696), CCA1 (U28422) and LHY (AJ006404).

revealed the presence of two *dSpm*-Act insertions (data not shown). One element cosegregated with the mutant phenotype, indicating a tight linkage. Isolation of genomic DNA flanking the *dSpm*-Act insertions revealed that a *dSpm*-Act element was inserted approximately 1.7 kb upstream of *At5g20240*, encoding the MADS-box transcription factor PISTILLATA (PI) (Fig. 2B). The *PI* gene is required for the specification of petals and stamens together with *APETALA3*. The same phenotype as observed for *pi-1D* has been reported for overexpression of *PI* from the constitutive 35S promoter [23], The *pi-1D* therefore is reminiscent of the *35S-PI* overexpression supporting that the partial homeotic transformation of first whorl sepals into petaloid organs is due to ectopic *PI* gene activity.

The identified semidominant mutant, *spätzünder-1D* (*spz-1D*), showed a late flowering phenotype under long day conditions, e.g., *spz-1D* produced 45–50 rosette leaves before flowering compared to 10–12 rosette leaves in wild-type plants (Fig. 3A). As it is the case for other late flowering mutants [24], many second and higher order inflorescences were found in *spz-1D* (Fig. 3A) and the life cycle was prolonged. The hypocotyl in *spz-1D* was significantly elongated (19.7 ± 1.4 mm in *spz-1D* versus 10.9 ± 0.9 mm in wild-type; 14 days after germination). Southern analysis of genomic DNA isolated from *spz-1D* revealed the presence of only one *dSpm*-Act insertion (data not shown). Outcrosses of *spz-1D* did not segregate for Basta-resistance, flowered with 15–17 rosette leaves and the hypocotyl-length was 14.1 ± 0.9 mm. These observed semidominant phenotypes indicated that the original *spz-1D* mutant was isolated in the homozygous state, which suggests that the transposition event took place in the T_1 generation. The *dSpm*-Act element was inserted on chromosome 3, approximately 2.1 kb upstream of *At3g10113* (Fig. 3B). *At3g10113* is highly homologous to *EPR1* (*At1g18330*) (99% DNA sequence identity), which was isolated in a screen for *EARLY-PHYTOCHROME-RESPONSIVE* genes [25]. EPR1 is related to CCA1 and LHY1, both components of the circadian clock and transcription factors harbouring a single MYB domain [26,27]. The conserved MYB domain of the products of *At3g10113*, *EPR1*, *CCA1* and *LHY1* is shown in Fig. 3C. Except for the MYB domain, EPR1 and its homolog do not share any similarities with CCA1 and LHY1.

Two further examples of dominant, heritable mutations resulting in adult defects are *shrubby-1D* (*shb-1D*) and *wirbel-1D* (*wib-1D*). The *shb-1D* mutant is characterized by a mild dwarfism with a bushy phenotype and many rosette-borne inflorescences, shorter internodes and a lesser fertility due to the lack of mature pollen. The *wib-1D* mutant possessed, all over, rolled rosette leaves and midribs; the flower organs, however, were unaffected (data not shown).

3.5. Identification of dominant metabolic mutants

As mentioned above, the frequency of developmental dominant mutant phenotypes was 0.1% within the TAMARA population. To identify additional dominant mutants, an HPLC-based metabolic screen for mutants with altered phenolic compounds was established. Different phenylpropanoids are formed in response to various stresses such as wounding, high light, UV radiation, or pathogen attack (reviewed in

[28]) and the use of genetics to dissect the phenylpropanoid pathways is well documented (reviewed in [29]).

In wild-type plants, grown under controlled conditions, the phenolic constituents in methanol extracts of leaves were separated by HPLC and detected by UV and fluorescence. A characteristic profile of the wild-type phenolic constituents, monitored by absorbance at 280 nm is shown in Fig. 4. The eight detected peaks correspond to at least four different phenolic compounds. Peaks 3, 6 and 7 were identified as kaempherol derivatives, based on their diode array spectrum with two maxima at 265 and 348 nm (Fig. 4). Peak 2 was identified as sinapoylglucose and peak 8 was identified as sinapoylmalate. Peaks 1, 4 and 5 could not be assigned yet.

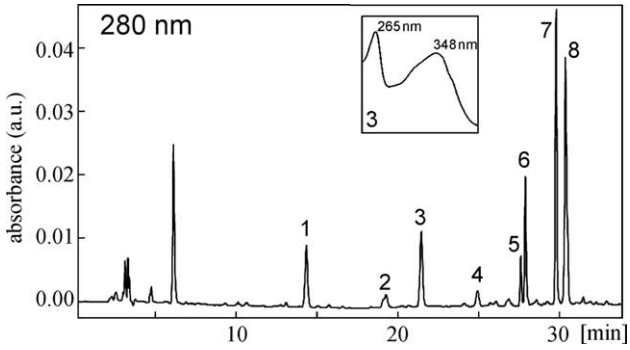

Fig. 4. HPLC-profile of phenolic constituents in methanolic extracts of Arabidopsis wild-type leaves. Extracts were monitored by absorbance at 280 nm. Peaks 3, 6 and 7 were identified as kaempherol derivatives, inset shows the spectrum (250–400 nm) of peak 3 eluting at 21.5 min. Peaks 2 and 8 were identified as sinapoylglucose and sinapoylmalate. Peaks 1, 4, and 5 are not yet identified.

A subset of the TAMARA population (1500 lines) was analysed as described above and 10 individuals with altered HPLC profiles were isolated. Outcross experiments revealed that in eight individuals the altered HPLC profile phenotype was not inheritable. The remaining two individuals showed the same HPLC profile (see below), were derived from the same starter line and possessed the same insertion as revealed by Southern blot analysis (data not shown), suggesting that the responsible mutation was caused by the same transposition event.

3.6. Characterization of hpc1-1D

The *h*igh *p*henolic *c*ompound1-*1D*ominant (*hpc1-1D*) mutant showed a particular accumulation of only one compound, namely the not yet identified substance eluting at 13.6 min (Fig. 5A). The diode array spectrum revealed only one maximum at 280 nm for this compound (Fig. 5B). Comparing the integrated peak areas obtained from wild-type and *hpc1-1D* plants on a fresh weight basis, the content of this compound was increased 6.3-fold (0.214 ± 0.086 in *hpc1-1D* versus 0.034 ± 0.016 in wild-type). The content of other constituents remained unchanged.

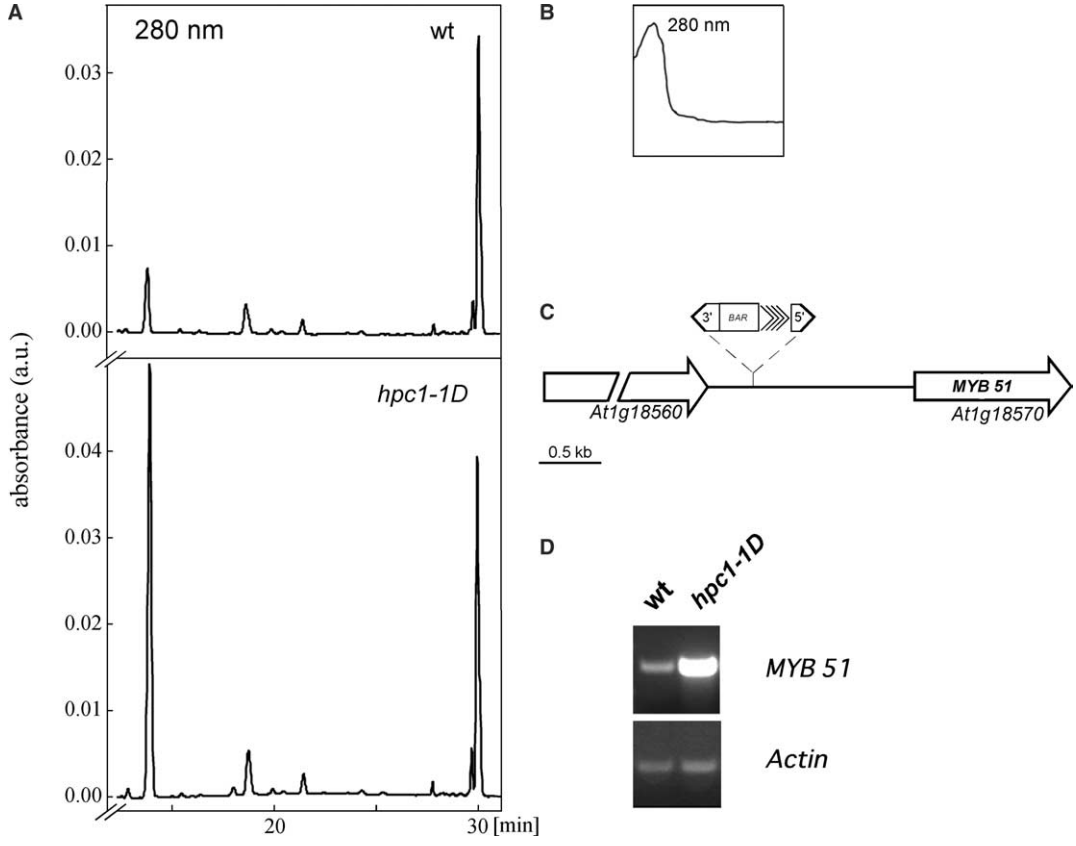

Fig. 5. Characterization of the *hpc1-1D* mutant. (A) HPLC-profile based phenotype of *hpc1-1D* in comparison to a wild-type HPLC-profile. Note the increased peak of a substance eluting at 13.6 min. (B) Spectrum of the peak eluting at 13.6 min. (C) Schematic representation of the localization and orientation of *dSpm*-Act in *hpc1-1D*. (D) RT-PCR analysis of the same plants as in (A) using *MYB 51* specific primers and *ACTIN1* specific primers as control. PCR conditions were set at 26 cycles.

The *hpc1* phenotype is transmitted to subsequent generations with a 1:1 or 3:1 segregation ratio in backcrosses to wild-type or in selfings, respectively, thus indicating a dominant mode of inheritance. Furthermore, the *hpc1* phenotype cosegregated with the *dSpm*-Act element (data not shown), demonstrating that the *hpc1-1D* mutant represents an activation tagged line. In *hpc1-1D* the *dSpm*-Act element was inserted on chromosome I, approximately 1.4 kb upstream of *At1g18570* (Fig. 5C). The product of *At1g18570*, MYB51, belongs to the large family of R2R3-MYB transcription factors, which are characterized by two related helix–turn–helix motifs [30]. A further hint that an increased transcription of *MYB51* is indeed responsible for the observed *hpc1* phenotype came from RT-PCR experiments. As shown in Fig. 5D, an upregulation of *MYB51* was detected in *hpc1-1D* mutants compared to wild-type plants. In contrast, the upstream located *At1g18560* gene was not expressed, neither in wild-type nor in mutant plants (data not shown). To confirm that the observed phenotype was due to gene activation, a construct was made in which the *MYB51* coding sequence was placed under control of the CaMV 35S promoter and transformed into wild-type plants. Wild-type plants and *35S-MYB51* plants were analysed by HPLC as described above. The content of the compound (eluting at 13.6 min), based on the integrated peak areas, was increased 11.3-fold (0.512 ± 0.06 in *35S-MYB51* plants versus 0.045 ± 0.022 in wild-type) in the overexpressors, thus confirming that the upregulation of *MYB51* was responsible for the *hpc1* phenotype.

4. Discussion

In the present study, we describe a two-element transposon system that was used to create a novel activation tagging population (TAMARA) in Arabidopsis. The TAMARA population consist of 9471 lines harbouring approximately 6000 independent *dSpm*-Act elements. This observed frequency of independent transposition events is in accordance with previously published data [12,13]. In contrast to these earlier reports [13], we did not observe an inverse correlation between the overall frequency of recovered transposants and the frequency of independent transpositions.

Furthermore, the low copy number (90% of TAMARA lines contained a single copy *dSpm*-Act element) and the uniform distribution of *dSpm*-Act elements throughout the genome (Fig. 1) are comparable with the previous described two-element transposon system [13].

Six developmental mutants out of 6000 inserts and one metabolic mutant out of approximately 950 inserts were isolated. With respect to dominant aerial morphological phenotypes an activation efficiency of 0.1% was determined, that compares well to the frequency of one dominant morphological mutant in 1000 lines reported for the T-DNA transformed activation tagged Arabidopsis population [5]. In a different type of gain-of-function screen using a Ds element carrying a CaMV 35S promoter, four dominant mutants were found among 1100 lines analyzed [31]. Surprisingly, the frequency of dominant aerial morphological mutants was 1% in the previous two-element transposon system [13] suggesting that activation tagging is more efficient with transposons than with T-DNA. The discrepancy of the activation efficiencies between the two systems (TAMARA, described in this study and in [13]) could be best explained by the difference in length of the terminal sequences used to construct the mobile elements. The use of shorter terminal sequences [13] might prevent methylation and hence silencing of the mobile element and, therefore, the activation efficiency could be higher. DNA methylation is an intrinsic property of transposons and, e.g., it has been shown that the binding of TnpA protein to the subterminal repetitive region of *En-1* is dependent on methylation (reviewed in [32]). However, the estimation of activation efficiencies needs to be put on a broader experimental base and thus awaits further studies.

The success of the TAMARA population was demonstrated by the functional characterization of the *Dornröschen* (*drn-D*) mutant [33] and further validated by the isolation of a new dominant *pistillata-1D* allele (Fig. 2). Except for *drn-D* and *pi-1D*, three additional developmental mutants and *spz-1D* were identified. The late flowering phenotype of the semi-dominant *spz-1D* mutant (Fig. 3) is likely caused by an upregulation of an *EPR1-homolog* gene (*At3g10113*). *EPR1* encodes a transcription factor containing a single MYB domain with high similarity to the Arabidopsis CCA1 and LHY MYB domains (Fig. 3C). Interestingly, overexpression of *EPR1*, *CCA1* and *LHY* resulted in delayed flowering [25,27,34], whereby CCA1 and LHY are components of the central circadian oscillator and EPR1 contributes to and fine-tunes circadian output pathways. The *EPR1-homolog* gene most probably originated from a gene duplication event. Whether the product of *At3g10113* itself is a component of the circadian oscillator or simple interferes with EPR1 in the *spz-1D* background remains to be determined.

In a second approach, the TAMARA population was used in a metabolic screen for increased levels of phenylpropanoid-derived compounds. For this, a robust HPLC-based screening method was developed, which ultimately led to the isolation of a dominant mutant, *hpc1-1D*. Arabidopsis leaves accumulate flavonols, especially kaempferols, the sinapate ester sinapoylmalate and to a lesser degree its precursor sinapoylglucose (Fig. 4). In contrast to the large number of mutations that affect the accumulation of pigmented phenylpropanoids, fewer mutations are known to affect the biosynthesis of other phenolic compounds (hydroxycinnamic acid-derived). Examples for the latter class are mutations that reduce sinapoylmalate accumulation in Arabidopsis [35]. In *hpc1-1D*, the contents of kaempferol derivatives and sinapate esters are not altered (Fig. 5), suggesting that the increased compound belongs to another group of secondary metabolites. Interestingly, the gene responsible for the observed *hpc1-1D* phenotype, belongs to the R2R3-type At*MYB* genes (Fig. 5). R2R3-type *MYB* genes have been shown to regulate phenylpropanoid metabolism in Arabidopsis and other plant species. ZmMYBC1 from *Zea mays*, the first MYB-related protein found in plants [36] and PhMYBAN2 from *Petunia hybrida* [37] are known to control phenylpropanoid metabolism. In Arabidopsis, the overexpression of At*MYB75*/*PAP1* (*Production of Anthocyanin Pigment1*) and At*MYB90*/*PAP2* results in an accumulation of lignin, hydroxycinnamic esters and flavonoids, including anthocyanins, throughout development [38] and AtMYB4 represses the synthesis of sinapoylmalate in leaves [39]. Together, the analysis of the *hpc1-1D* mutant and the functional characterisation of the responsible gene will contribute to the understanding of the complex regulatory

network of phenylpropanoid metabolism. Furthermore, it is envisaged that the use of the TAMARA population will enable the identification of further novel interesting traits.

Acknowledgements: We thank Kerstin Kunze for technical assistance and Barbara Hess and Ingmar Bürstel as undergraduate students, participating in this project. This work was supported by grants from the Deutsche Forschungsgemeinschaft (WE 1262/3-1 to W.W and SI 677/1-1 to R.S.) and the European Union (QLK1-CT-2001-01080 to U-I.F. and U.S.).

References

[1] The Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–815.
[2] Chang, C., Kwok, S.F., Bleecker, A.B. and Meyerowitz, E.M. (1993) Arabidopsis ethylene-response gene *ETR1*: similarity of product to two-component regulators. Science 262, 539–544.
[3] Odell, J.T., Nagy, F. and Chua, N.H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313, 810–812.
[4] Hayashi, H., Czaja, I., Lubenow, H., Schell, J. and Walden, R. (1992) Activation of a plant gene by T-DNA tagging: auxin-independent growth in vitro. Science 258, 1350–1353.
[5] Weigel, D. et al. (2000) Activation tagging in Arabidopsis. Plant Physiol. 122, 1003–1013.
[6] Ichikawa, T. et al. (2003) Sequence database of 1172 T-DNA insertion sites in Arabidopsis activation-tagging lines that showed phenotypes in T1 generation. Plant J. 36, 421–429.
[7] Gheysen, G., Herman, L., Breyne, P., Gielen, J., Van Montagu, M. and Depicker, A. (1990) Cloning and sequence analysis of truncated T-DNA inserts from *Nicotiana tabacum*. Gene 94, 155–163.
[8] Nacry, P., Camilleri, C., Courtial, B., Caboche, M. and Bouchez, D. (1998) Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. Genetics 149, 641–650.
[9] Aarts, M.G., Corzaan, P., Stiekema, W.J. and Pereira, A. (1995) A two-element enhancer–inhibitor transposon system in *Arabidopsis thaliana*. Mol. Gen. Genet. 247, 555–564.
[10] Wisman, E., Cardon, G.H., Fransz, P. and Saedler, H. (1998) The behaviour of the autonomous maize transposable element *En*/*Spm* in *Arabidopsis thaliana* allows efficient mutagenesis. Plant Mol. Biol. 37, 989–999.
[11] Speulman, E., Metz, P.L., Van Arkel, G., Te Lintel Hekkert, B., Stiekema, W.J. and Pereira, A. (1999) A two-component enhancer–inhibitor transposon mutagenesis system for functional analysis of the Arabidopsis genome. Plant Cell 11, 1853–1866.
[12] Tissier, A.F., Marillonnet, S., Klimyuk, V., Patel, K., Torres, M.A., Murphy, G. and Jones, J.D. (1999) Multiple independent defective suppressor–mutator transposon insertions in Arabidopsis: a tool for functional genomics. Plant Cell 11, 1841–1852.
[13] Marsch-Martinez, N., Greco, R., Van Arkel, G., Herrera-Estrella, L. and Pereira, A. (2002) Activation tagging using the *En–I* maize transposon system in Arabidopsis. Plant Physiol. 129, 1544–1556.
[14] De Block, M., Botterman, J., Van de Wiele, M., Dockx, J., Thoen, C., Gossele, V., Movva, N.R., Thompson, C., Van Montagu, M. and Leemans, J. (1987) Engineering herbicide resistance in plants by expression of a detoxifying enzyme. EMBO J. 6, 2513–2518.
[15] O'Keefe, D.P., Tepperman, J.M., Dean, C., Leto, K.J., Erbes, D.L. and Odell, J.T. (1994) Plant expression of a bacterial cytochrome P450 that catalyzes activation of a sulfonylurea pro-herbicide. Plant Physiol. 105, 473–482.
[16] Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) New plant binary vectors with selectable markers located proximal to the left T-DNA border. Plant Mol. Biol. 20, 1195–1197.
[17] Überlacker, B. and Werr, W. (1996) Vectors with rare-cutter restriction enzymes sites for expression of open reading frames in transgenic plants. Mol. Breeding 2, 293–295.
[18] Bechtold, N., Ellis, J. and Pelletier, G. (1993) In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C.R. Acad. Sci. 316, 1194–1199.
[19] Liu, Y.G., Mitsukawa, N., Oosumi, T. and Whittier, R.F. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. 8, 457–463.
[20] Varotto, C., Pesaresi, P., Meurer, J., Oelmuller, R., Steiner-Lange, S., Salamini, F. and Leister, D. (2000) Disruption of the Arabidopsis photosystem I gene *psaE1* affects photosynthesis and impairs growth. Plant J. 22, 115–124.
[21] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
[22] Mock, H.P., Heller, W., Molina, A., Neubohn, B., Sandermann Jr., H. and Grimm, B. (1999) Expression of uroporphyrinogen decarboxylase or coproporphyrinogen oxidase antisense RNA in tobacco induces pathogen defense responses conferring increased resistance to tobacco mosaic virus. J. Biol. Chem. 274, 4231–4238.
[23] Krizek, B.A. and Meyerowitz, E.M. (1996) The Arabidopsis homeotic genes *APETALA3* and *PISTILLATA* are sufficient to provide the B class organ identity function. Development 122, 11–22.
[24] Koornneef, M., Hanhart, C.J. and Van der Veen, J.H. (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol. Gen. Genet. 229, 57–66.
[25] Kuno, N., Moller, S.G., Shinomura, T., Xu, X., Chua, N.H. and Furuya, M. (2003) The novel MYB protein EARLY-PHYTOCHROME-RESPONSIVE1 is a component of a slave circadian oscillator in Arabidopsis. Plant Cell 15, 2476–2488.
[26] Wang, Z.Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, M.S. and Tobin, E.M. (1997) A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. Plant Cell 9, 491–507.
[27] Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I.A. and Coupland, G. (1998) The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. Cell 93, 1219–1229.
[28] Dixon, R.A. and Paiva, N.L. (1995) Stress-induced phenylpropanoid metabolism. Plant Cell 7, 1085–1097.
[29] Memelink, J. (2005) The use of genetics to dissect plant secondary pathways. Curr. Opin. Plant Biol. 8, 230–235.
[30] Stracke, R., Werber, M. and Weisshaar, B. (2001) The *R2R3-MYB* gene family in *Arabidopsis thaliana*. Curr. Opin. Plant Biol. 4, 447–456.
[31] Wilson, K., Long, D., Swinburne, J. and Coupland, G. (1996) A Dissociation insertion causes a semidominant mutation that increases expression of *TINY*, an Arabidopsis gene related to *APETALA2*. Plant Cell 8, 659–671.
[32] Kunze, R., and Weil, C.F. (2002) In: Mobile DNA II (Craig, N.L., Ed.) The *hAT* and *CACTA* Superfamilies of Plant Transposons (pp. 565–610). Washington, DC.
[33] Kirch, T., Simon, R., Grunewald, M. and Werr, W. (2003) The *DORNRÖSCHEN/ENHANCER OF SHOOT REGENERATION1* gene of Arabidopsis acts in the control of meristem cell fate and lateral organ development. Plant Cell 15, 694–705.
[34] Wang, Z.Y. and Tobin, E.M. (1998) Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*) gene disrupts circadian rhythms and suppresses its own expression. Cell 93, 1207–1217.
[35] Ruegger, M. and Chapple, C. (2001) Mutations that reduce sinapoylmalate accumulation in *Arabidopsis thaliana* define loci with diverse roles in phenylpropanoid metabolism. Genetics 159, 1741–1749.
[36] Paz-Ares, J., Wienand, U., Peterson, P.A. and Saedler, H. (1986) Molecular cloning of the *c1* locus of *Zea mays*: a locus regulating the anthocyanin pathway. EMBO J. 5, 829–833.
[37] Quattrocchio, F., Wing, J., Van der Woude, K., Souer, E., De Vetten, N., Mol, J. and Koes, R. (1999) Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. Plant Cell 11, 1433–1444.
[38] Borevitz, J.O., Xia, Y., Blount, J., Dixon, R.A. and Lamb, C. (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell 12, 2383–2394.
[39] Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B. and Martin, C. (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. EMBO J. 19, 6150–6161.