

RESEARCH PAPER

Multiple *cis*-regulatory elements are involved in the complex regulation of the sieve element-specific *MtSEO-F1* promoter from *Medicago truncatula*

M. Bucsenez^{1,*}, B. Rüping^{2,*}, S. Behrens³, R. M. Twyman⁴, G. A. Noll² & D. Prüfer^{1,2}

¹ Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Aachen, Germany

² Institut für Biologie und Biotechnologie der Pflanzen, Westfälische Wilhelms-Universität Münster, Münster, Germany

³ Max Planck Institute for Molecular Genetics, Computational Molecular Biology, Berlin, Germany

⁴ Department of Biological Sciences, University of Warwick, Coventry, UK

Keywords

Companion cell; forisome; immature sieve element; phloem; promoter; reporter gene; sieve element occlusion.

Correspondence

D. Prüfer, Institut für Biologie und Biotechnologie der Pflanzen, Westfälische Wilhelms-Universität Münster, Hindenburgplatz 55, 48143 Münster, Germany.

E-mail: dpruefer@uni-muenster.de

Editor

G. Thiel

*These authors contributed equally to this work.

Received: 30 September 2011; Accepted: 21 November 2011

doi:10.1111/j.1438-8677.2011.00556.x

INTRODUCTION

Long-distance transport in higher plants requires a vascular system comprising two specialised tissues, xylem and phloem, which both originate from the (pro)cambium (Fukuda 2004; Carlsbecker & Helariutta 2005). The regulatory mechanisms controlling early vascular patterning, the differentiation of vascular meristems and the specification of xylem cells have been studied in detail (Fukuda 2004; Sieburth & Deyholos 2006; Baucher *et al.* 2007; Ilegems *et al.* 2010). Much less is known about the specification of phloem cells into sieve elements (SEs) and companion cells (CCs) (Fukuda 2004), which form the functional core unit of sieve tubes and thereby play a central role in photosynthate supply between source and sink tissues throughout the plant (van Bel *et al.* 2002). Among the few transcriptional regulators that have been associated with phloem differentiation (Nishitani *et al.* 2001; Bonke *et al.* 2003; Skirycz *et al.* 2006), the transcription factor APL (ALTERED PHLOEM DEVELOPMENT) appears to have a critical role in the specification of phloem identity. APL is required for the asymmetric cell division of phloem

ABSTRACT

The sieve element occlusion (*SEO*) gene family includes several members that are expressed specifically in immature sieve elements (SEs) in the developing phloem of dicotyledonous plants. To determine how this restricted expression profile is achieved, we analysed the SE-specific *Medicago truncatula* *SEO-F1* promoter (*PMtSEO-F1*) by constructing deletion, substitution and hybrid constructs and testing them in transgenic tobacco plants using green fluorescent protein as a reporter. This revealed four promoter regions, each containing *cis*-regulatory elements that activate transcription in SEs. One of these segments also contained sufficient information to suppress *PMtSEO-F1* transcription in the phloem companion cells (CCs). Subsequent *in silico* analysis revealed several candidate *cis*-regulatory elements that *PMtSEO-F1* shares with other *SEO* promoters. These putative sieve element boxes (PSE boxes) are promising candidates for *cis*-regulatory elements controlling the SE-specific expression of *PMtSEO-F1*.

precursors into SEs and CCs (Bonke *et al.* 2003) and is found in the nuclei of developing SEs and CCs in *Arabidopsis thaliana*. However, no APL target genes or binding motifs have been reported thus far, so a direct link to interaction partners involved in phloem specification remains to be established.

In contrast, there have been numerous studies of phloem-specific and CC-specific genes and their promoters (Brears *et al.* 1991; Yin *et al.* 1997; Hehn & Rohde 1998; Ayre *et al.* 2003; Guo *et al.* 2004; Guan & Zhou 2006; Schneidereit *et al.* 2008; Tsuwamoto & Harada 2010, 2011). One remarkable example is the promoter of the *A. thaliana* sucrose transporter gene *AtSUC2* (Stadler & Sauer 1996), which restricts expression to the CCs of source tissues. *PAAtSUC2* comprises two promoter elements that are essential for CC-specific expression and contain binding sites for HD-ZIP (homeodomain leucine zipper) and DOF (DNA-binding with one finger) transcription factors (Schneidereit *et al.* 2008).

Despite studies of these individual transcription factors and phloem-specific sequence motifs, there is still little known about the regulatory mechanisms controlling early-stage SE/CC complex formation. Genes that are active in the SE/CC

complex during maturation are regulatory targets, such as those encoding callose synthases, which are involved in the formation of sieve pores (Xie *et al.* 2011), and P-proteins (phloem proteins), which are essential for wound sealing in the phloem (Ehlers *et al.* 2000). The synthesis and assembly of P-proteins takes place before the SEs mature and undergo partial autolysis. This is most apparent for the large crystalline P-proteins of the Fabaceae, which are known as forisomes (Wergin & Newcomb 1970; Palevitz & Newcomb 1971). Forisomes are encoded by genes of the recently characterised *sieve element occlusion (SEO)* gene family, which most likely also includes conventional P-protein genes (Rüping *et al.* 2010). We have shown that the forisome genes *MtSEO-F1*, *MtSEO-F2* and *MtSEO-F3* are specifically expressed in immature SEs (Noll *et al.* 2007, 2009). This is also true for *SEO* genes from soybean (*GmSEO-F1*) and *A. thaliana* (*AtSEOa*) (Rüping *et al.* 2010), but expression profiles have not been described for any other genes thus far. By investigating the regulatory control of SE-specific *SEO* genes, we therefore seek new insights into phloem cell type specification during SE development.

Here we describe the identification and characterisation of regulatory regions in the *MtSEO-F1* promoter (formerly described as *Pmtfor1*; Noll *et al.* 2007) that are essential for SE-specific expression, and reveal the presence of additional CC-specific motifs. *In silico* comparisons of the *MtSEO-F1* promoter sequence with other *SEO* promoters and with different phloem-specific promoters reveal several conserved motifs that might also contribute to the specificity of *PMtSEO-F1*.

MATERIAL AND METHODS

In silico identification of the transcriptional start site and TATA box

The putative transcriptional start site and TATA box of the *MtSEO-F1* promoter were predicted using the Transcriptional Start Site Prediction (TSSP) tool from Softberry (<http://www.softberry.com>).

Verification of the transcriptional start site

Total RNA from *M. truncatula* cv. Jemalong A17 leaves was isolated using the NucleoSpin[®] RNA Plant Kit (Macherey-Nagel, Düren, Germany) following the manufacturer's protocol, and the transcriptional start site was verified by 5' RACE. First-strand cDNA was synthesised with SuperScript[™]II (Invitrogen, Karlsruhe, Germany) using an *MtSEO-F1*-specific reverse primer (5'-GAT CTG ATG AGG TTG AAG TC-3') and was purified using the NucleoSpin[®] Extract II Kit (Macherey-Nagel). An oligo(dG) tail was added to the 3'-end using terminal deoxynucleotidyl transferase (NEB, Frankfurt, Germany) and the modified cDNA was amplified with an oligo(dC₁₅) forward primer and an *MtSEO-F1*-specific nested reverse primer (5'-GGG GTG GTT ATC ATC TGA CAA G-3'). The resulting PCR product was sequenced to confirm the transcriptional start site.

Verification of the putative TATA box

Site-directed mutagenesis of plasmid pUC103GUS (Noll *et al.* 2007), containing the *MtSEO-F1* promoter upstream of the

uidA reporter gene encoding β -glucuronidase (GUS), was carried out using the QuikChange[®] Lightning Site-Directed Mutagenesis Kit (Stratagene/Agilent Technologies, Waldbronn, Germany) according to the manufacturer's recommendations, in order to inactivate the putative *PMtSEO-F1* TATA box. Forward primer 5'-AGT GTA AAC TCT TAT GAC TTT AGA GGC CCC CAA TAT GTG GTG GGA GCA CCA ATA CTA C-3' and reverse primer 5'-GTA GTA TTG GTG CTC CCA CCA CAT ATT GGG GGC CTC TAA AGT CAT AAG AGT TTA CAC T-3' were used to replace the sequence 'TATA' with 'CCCC', giving rise to the construct *PMtSEO-F1*_{TATAmut}-GUS.

Tobacco mesophyll protoplasts were isolated and transfected (Negrutiu 1981; Negrutiu *et al.* 1987) with 10 μ g of DNA. Three independent transfections were carried out using the original and mutated constructs. After cultivation for 18 h, protoplasts were collected, frozen in liquid nitrogen and sheared in PBS. Cell debris was pelleted by centrifugation and the protein content of the supernatant was determined as described in Bradford (1976). The GUS activity in 10 μ g of protein extract was determined using 4-methylumbelliferylglucuronide (4-MUG) as the substrate (Jefferson *et al.* 1987).

Preparation of total RNA, and GFP_{ER} expression analysis

We prepared enriched phloem tissue from tobacco (*Nicotiana tabacum* var. Petit Havana SR1) plants as previously described for soybean (Rüping *et al.* 2010). Total phloem RNA was isolated using the NucleoSpin[®] RNA Plant Kit (Macherey-Nagel) following the manufacturer's protocol. First-strand cDNA synthesis with SuperScript[™]II (Invitrogen) was performed using an oligo(dT₂₈) primer following the manufacturer's recommendations. *GFP_{ER}* gene expression was monitored using 1 μ l cDNA for PCR amplification with specific oligonucleotides (forward primer 5'-GAG AAG AAC TTT TCA CTG G-3', reverse primer 5'-CTT TGT ATA GTT CAT CCAT G-3') and the resulting PCR products were verified by sequencing. The *GAPDH* gene was used as an internal control (forward primer 5'-AGC TCA AGG TTA AGG ATG AC-3', reverse primer 5'-TGG CCA AGG GAG CAA GGC AA-3').

Construction of promoter deletions

All promoter fragments were prepared by PCR using the plasmid pBSGFP_{ER}, containing *PMtSEO-F1* (Noll *et al.* 2007; PC1), as the template. The sequences of all primers are shown in Table S1. Truncated promoter constructs were amplified using a common reverse primer and different forward primers (PC2: pc2fw/pc2rev; PC3: pc3fw/pc2rev; PC4: pc4fw/pc2rev). The resulting PCR products were digested with *KpnI* and *XhoI* and inserted into the corresponding restriction sites of pBSGFP_{ER} (Noll *et al.* 2007) to obtain the constructs *PMtSEO-F1*₃₄₇-GFP_{ER} (PC2), *PMtSEO-F1*₂₄₄-GFP_{ER} (PC3) and *PMtSEO-F1*₁₃₆-GFP_{ER} (PC4).

Hybrid constructs combining *PMtSEO-F1* and the minimal *Cauliflower mosaic virus* (CaMV) 35S promoter were produced by amplifying *PMtSEO-F1* with different forward primers and a common reverse primer (pc5rev), containing a 46-bp minimal sequence from the CaMV 35S promoter (PC5: pc5fw/pc5rev; PC6: pc3fw/pc5rev; PC7: pc4fw/pc5rev).

PCR products were digested with *KpnI* and *XhoI* and inserted into the corresponding sites of pBSGFP_{ER}, resulting in constructs *PMtSEO-F1*₁₀₂₁₋₄₁P35S_{min}-GFP_{ER} (PC5), *PMtSEO-F1*₂₄₄₋₄₁P35S_{min}-GFP_{ER} (PC6) and *PMtSEO-F1*₁₃₆₋₄₁P35S_{min}-GFP_{ER} (PC7), respectively.

The internal deletion *PMtSEO-F1*_{Δ-244-136}-GFP_{ER} (PC9) was obtained by amplifying two promoter fragments using the primer combinations pc5fw/pc9rev and pc9fw/pc2rev. The products were digested with *PstI*, ligated and then digested with *KpnI* and *XhoI* allowing insertion into the corresponding sites in pBSGFP_{ER}. The same strategy was used for *PMtSEO-F1*_{Δ-244-41}-GFP_{ER} (PC11) with the primer pairs pc5fw/pc9rev and pc11fw/pc2rev, for *PMtSEO-F1*_{Δ-347-41}-GFP_{ER} (PC13) with the primer sets pc5fw/pc13rev and pc11fw/pc2rev, and for *PMtSEO-F1*_{Δ-347Δ-244-136}-GFP_{ER} (PC15) with the primer combination pc2fw/pc9rev and pc9fw/pc2rev. Spacers introduced in the constructs carrying internal deletions were amplified from the *lacZ* gene in plasmid pUC18 with the length corresponding to the deleted *PMtSEO-F1* fragment. The primer combination pc8fw/pc8rev was used to amplify a 108-bp spacer for constructs PC8 and PC14. A 203-bp spacer for constructs PC10 and PC16 was amplified with the primer set pc10fw/pc10rev and amplification with the primers pc12fw/pc12rev generated a 306-bp spacer used for PC12. The PCR products were digested with *PstI* and inserted into the *PstI* site of the corresponding construct without spacers, giving rise to *PMtSEO-F1*_{Δ-244-136+Spc}-GFP_{ER} (PC8), *PMtSEO-F1*_{Δ-244-41+Spc}-GFP_{ER} (PC10), *PMtSEO-F1*_{Δ-347-41+Spc}-GFP_{ER} (PC12) and *PMtSEO-F1*_{Δ-244-136+Spc}-GFP_{ER} (PC14), respectively.

To insert *PMtSEO-F1* fragments into *PatSUC2*, a 939-bp *AtSUC2* promoter sequence was amplified from *A. thaliana* cv. Col-0 genomic DNA isolated according to established protocols (Doyle & Doyle 1990). The primers patsuc2fw/patsuc2rev were used to insert a *PstI* site into the *PatSUC2* sequence. The resulting PCR product was digested with *KpnI/XhoI* and inserted into the corresponding sites of pBSGFP_{ER} giving rise to *PatSUC2*_{PstI}-GFP_{ER}. For the control cassette *PatSUC2*_{Spc}-GFP_{ER} (PC16), the 203-bp spacer described above was introduced into *PatSUC2*_{PstI}-GFP_{ER} via the *PstI* sites. The hybrid constructs *PatSUC2*_{*PMtSEO-F1*Δ-244-41}-GFP_{ER} (PC17), *PatSUC2*_{*PMtSEO-F1*Δ-244-136}-GFP_{ER} (PC18), *PatSUC2*_{*PMtSEO-F1*Δ-136-41}-GFP_{ER} (PC19), *PatSUC2*_{*PMtSEO-F1*Δ-244-189+Spc}-GFP_{ER} (PC20), *PatSUC2*_{*PMtSEO-F1*Δ-209-154+Spc}-GFP_{ER} (PC21) and *PatSUC2*_{*PMtSEO-F1*Δ-174-114+Spc}-GFP_{ER} (PC22) were generated by amplifying the *PMtSEO-F1* sequence with specific primers (PC17: pc17fw/pc17rev; PC18: pc17fw/pc18rev; PC19: pc9fw/pc17rev; PC20: pc17fw/pc20rev; PC21: pc21fw/pc21rev; PC22: pc22fw/pc22rev), digesting the PCR products with *PstI* and inserting them into the *PstI* site of *PatSUC2*_{PstI}-GFP_{ER}. The *lacZ* spacer sequences of PC20, PC21 and PC22 were introduced using the corresponding reverse primers. All the constructs described above were sequenced for verification and inserted into the *KpnI/HindIII* sites of the binary vector pBIN19 (Bevan 1984).

Plant transformation and analysis

The binary vectors described above were transferred into *Agrobacterium tumefaciens* strain LBA4404 (Hoekema *et al.* 1983). Tobacco plants were transformed using the leaf disc method

(Horsch *et al.* 1986) and five independent transgenic lines were regenerated for each construct. GFP_{ER} fluorescence in longitudinal and cross-sections through the petioles was monitored with confocal laser scanning microscopy (CLSM) using a Leica TCS SP5 X (Wetzlar, Germany), with excitation and emission wavelengths of 488 and 500–600 nm, respectively. Sieve plates were stained with 0.01% aniline blue following established protocols (Thompson & Wolniak 2008) and were visualised using CLSM as above, with excitation and emission wavelengths of 364 and 470–530 nm, respectively.

Motif discovery

We screened the literature for *cis*-regulatory elements involved in CC-specific expression, collecting candidate sequences to use as search strings with the DNA-pattern tool available from RSAT (<http://rsat.ulb.ac.be/rsat/>). Elements shorter than 8 bp required exact matches whereas for longer sequences we allowed up to 15% nucleotide substitution. Additional known transcription factor binding sites were identified using the PLACE Web Signal Scan (Prestridge 1991; Higo *et al.* 1999). We also performed *de novo* motif discovery in a number of SE-specific *SEO* promoters using the Amadeus/Allegro software package (Linhart *et al.* 2008) with repeat-masked *A. thaliana* promoters available on the Amadeus/Allegro webpage (<http://acgt.cs.tau.ac.il/amadeus/>) as a background set. Amadeus/Allegro was applied to a target set comprising 13 different *SEO* promoters from various dicotyledonous plants, all of which are known to be SE-specific (Noll *et al.* 2007, 2009; Rüping *et al.* 2010 and unpublished data) and which we therefore consider to be co-regulated. Using the Amadeus/Allegro default parameters, but adjusting the range from –2500 to +200, we searched for motifs 6–12 bp in length. We discarded those with fewer than 12 hits in the analysed promoters. Based on the Amadeus/Allegro output matrices, we used the RSAT matrix scan to localise motifs in the promoter sequences using the default parameters (Turatsinze *et al.* 2008) and designated them as 'putative sieve element' (PSE) boxes.

RESULTS AND DISCUSSION

Determination of the *PMtSEO-F1* transcriptional start site and the TATA box

A 1094-bp promoter sequence from *MtSEO-F1* has previously been shown to drive gene expression specifically in immature SEs, but no detailed promoter characterisation has been carried out (Noll *et al.* 2007). We therefore started our characterisation by locating the transcriptional start site (TSS) and TATA box. *In silico* analysis using the TSSP software predicted the location of the TSS 73 bp upstream of the translational initiation codon (ATG). We verified the predicted TSS by amplifying the 5'-end of the corresponding cDNA using 5' RACE. This generated a PCR product ending 73 bp upstream of the translational initiation codon, which was consistent with the *in silico* prediction. We defined this TSS as position +1 for subsequent description of the promoter sequence (Figure S1).

A putative TATA box with the sequence 'TATAAA' was identified with the TSSP software at position -31 (Figure S1).

Experimental verification of this predicted TATA box was carried out through converting the sequence 'TATA' to 'CCCC' using site-directed mutagenesis. The native and mutated promoter sequences were placed upstream of the *uidA* (GUS) gene and the activity of each promoter was determined by measuring GUS activity in transfected tobacco protoplasts (Table 1). We found that GUS activity in protoplasts transfected with the mutated promoter construct was 72% lower than in those transfected with the unmodified construct, confirming the importance of the modified sequence and indicating the predicted TATA box was genuine.

Because there are few descriptions of SE-specific promoters and, to our knowledge, no data concerning SE-specific *cis*-regulatory elements, we decided to localise potential *cis*-regulatory elements in the *PMtSEO-F1* sequence by truncation analysis using an ER-tagged version of *GFP* as a reporter gene.

Analysis of truncated *PMtSEO-F1* fragments and substitutions

In order to localise *cis*-regulatory elements in the *PMtSEO-F1* sequence, we dissected the promoter by creating truncated versions from the 5'-end as well as substitutions of the core promoter and 5'-UTR. The different promoter versions were placed upstream of the *GFP_{ER}* reporter gene, introduced into tobacco plants and analysed by confocal laser scanning microscopy (CLSM). We analysed at least five independent lines per construct by checking several leaf petioles from each line and observed the same expression pattern in each group of lines representing the same construct in all cases. The 1094-bp 'full length' *MtSEO-F1* promoter served as a control (promoter construct 1, PC1; Fig. 1A). In cross-sections of PC1 petioles, fluorescence was detected in the phloem parts of the vascular bundle (Fig. 1B). Longitudinal sections of petioles showed fluorescent sieve elements orientated in their typical pipeline-like assembly (Fig. 1C) and aniline blue staining of the sieve plates confirmed the strict SE-specific expression (Fig. 1D). Immature SEs were identified by the presence of large vacuoles (Fig. 1D), but fluorescence was also detected in a large number of mature SEs (Fig. 1E), which was recognised by the parietal organisation of the ER and the lack of vacuoles. The fluorescence in mature SEs is due to the use of ER-tagged GFP in this study, which allowed us to distinguish the exact cell type in which reporter gene expression occurs. As a co-effect of GFP retention in the ER and its half-life of approximately 18 h (Ruijter *et al.* 2003), fluorescence persists in mature SEs, even though no protein biosynthesis is possible. Therefore, mature fluorescent SEs are always visible when ER-tagged reporter genes are expressed in immature SEs.

Next, three incremental 5' deletions of *PMtSEO-F1* were generated, resulting in the constructs PC2 (*PMtSEO-F1*₋₃₄₇-*GFP_{ER}*), PC3 (*PMtSEO-F1*₋₂₄₄-*GFP_{ER}*) and PC4 (*PMtSEO-F1*₋₁₃₆-*GFP_{ER}*; Fig. 1A). Tobacco plants expressing construct PC2 displayed an identical fluorescence pattern to those expressing PC1 (Fig. 1F–H). The next construct (PC3) still directed expression to immature SEs, although a more intense laser was necessary to visualise the GFP signals, indicating a lower expression level (Fig. 1I). The cross-section of PC3 (Fig. 1I) additionally displays xylem autofluorescence, which can also be observed in wild-type tobacco petioles (Fig. 1J). Because the aniline blue staining procedure resulted in a further disturbance of fluorescence, sieve plates were not stained in these plants, but the structures can nevertheless be seen in the *GFP_{ER}* fluorescence micrograph of an SE (Fig. 1K and L). For the final construct in the series, PC4, no fluorescence could be detected using CLSM.

Parallel to these 5' truncations, we generated a second series of constructs by replacing the 114-bp *PMtSEO-F1*₋₄₁₋₊₇₃ fragment (representing the 41-bp core promoter and the 5'-UTR) with the 46-bp minimal CaMV 35S promoter, which is reported to show no visible activity in transgenic plants (Benfey *et al.* 1990). At the same time, different 5' truncations were introduced, resulting in the hybrid constructs PC5 (*PMtSEO-F1*₋₁₀₂₁₋₄₁P35S_{min}-*GFP_{ER}*), PC6 (*PMtSEO-F1*₋₂₄₄₋₄₁P35S_{min}-*GFP_{ER}*) and PC7 (*PMtSEO-F1*₋₁₃₆₋₄₁P35S_{min}-*GFP_{ER}*), which are shown in Fig. 1A. The analysis of PC5 with CLSM revealed a fluorescence pattern identical to that generated by the native promoter construct PC1 (Fig. 1B–D), showing strong expression in immature SEs (Fig. 1M–O). The next construct in the series, PC6, yielded the same fluorescence pattern, although again a more intense laser was necessary to obtain visible fluorescence signals (Fig. 1P). The shortest construct (PC7) did not produce detectable fluorescence and shared this pattern with PC4, which carried the same upstream sequence but the *PMtSEO-F1*₋₄₁₋₊₇₃ fragment.

Taken together, the expression patterns generated by constructs containing the minimal CaMV 35S promoter (PC5, PC6 and PC7) did not differ from their corresponding native promoter versions (PC1, PC3 and PC4). We can therefore assume that the *MtSEO-F1* core promoter as well as the 5'-UTR do not contain *cis*-regulatory elements that are obligatory for SE-specific expression. Instead, *cis*-regulatory element(s) for transcriptional activation in SEs (designated SE_{act}) must be located between positions -244 and -41, as this part proved to be sufficient for SE-specific expression in construct PC6. Furthermore, the absence of fluorescence in plants carrying the most severely truncated constructs (PC4 and PC7) suggests that the SE_{act} is lost in these constructs and can therefore be localised between positions -244 and

Table 1. Site-directed mutagenesis of the putative *PMtSEO-F1* TATA box. GUS activity in protein extracts from transfected tobacco protoplasts containing either the native (*PMtSEO-F1*-GUS) or the mutated (*PMtSEO-F1*_{TATAmut}-GUS) promoter sequence upstream of the *uidA* (GUS) reporter gene was measured in three independent 4-MUG assays (excitation 365 nm, emission 455 nm).

construct name	fluorescence intensity			mean value	standard deviation	relative activity (%)
	sample 1	sample 2	sample 3			
<i>PMtSEO-F1</i> -GUS	17,004	18,025	15,083	16,704	1220	100
<i>PMtSEO-F1</i> _{TATAmut} -GUS	4810	3739	5595	4715	761	28

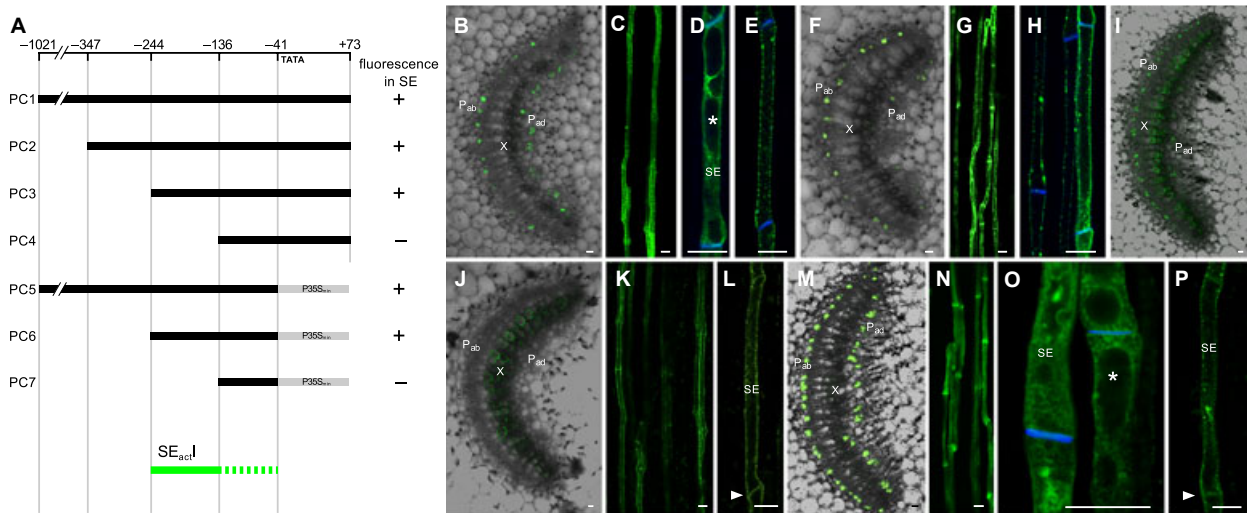


Fig. 1. Analysis of *PMtSEO-F1* truncation and substitution constructs. A: Scheme of *PMtSEO-F1* constructs (PC) with sequential 5' truncations and substituted *PMtSEO-F1*_{-41→+73} fragment. *PMtSEO-F1* fragments are represented by black bars; the CaMV 35S minimal promoter is shown in grey; the preliminary localisation of a *cis*-regulatory element within *PMtSEO-F1*-mediating SE-specific expression (*SE_{act}I*) is shown in green (alternative dashed). PC1: *PMtSEO-F1*-GFP_{ER}, PC2: *PMtSEO-F1*₋₃₄₇-GFP_{ER}, PC3: *PMtSEO-F1*₋₂₄₄-GFP_{ER}, PC4: *PMtSEO-F1*₋₁₃₆-GFP_{ER}, PC5: *PMtSEO-F1*₋₁₀₂₁₋₄₁P35S_{min}-GFP_{ER}, PC6: *PMtSEO-F1*₋₂₄₄₋₄₁P35S_{min}-GFP_{ER}, PC7: *PMtSEO-F1*₋₁₃₆₋₄₁P35S_{min}-GFP_{ER}; (B–P) GFP_{ER} fluorescence in tobacco plants. Cross-sections through leaf petioles are shown as overlays of GFP_{ER} fluorescence and transmitted light (B, F, I, J, M), longitudinal petiole sections are shown in (C–E, G, H, K, L, N–P); (B–E) PC1; (F–H) PC2; (I, K, L) PC3; (M–O) PC5; (P) PC6. Vacuoles are indicated with asterisks. Sieve plates are stained with aniline blue (in D, E, H, O) or indicated with arrowheads (in L, P). P_{ab}, abaxial phloem; P_{ad}, adaxial phloem; X, xylem; scale bars = 25 µm.

-136. It is also possible that enhancer elements rather than spatial control elements are missing in the context of PC4 and PC7. In order to test the overlapping promoter segments potentially containing *SE_{act}* sequences (-244 to -41 and -244 to -136) while avoiding the loss of putative upstream enhancers, we decided to delete the promoter sequences within the full-length *PMtSEO-F1* sequence for verification.

Analysis of *PMtSEO-F1* internal deletion constructs

We created a series of *PMtSEO-F1* internal deletion constructs by removing internal segments in the context of the full-length promoter, and for each removed segment, creating one construct in which the flanking sequences became contiguous and another in which the deleted segment was replaced with an equal-sized but non-specific region of the *lacZ* gene to maintain spacing and preserve the activity of any *cis*-regulatory elements with a distance dependency. The constructs PC8 (*PMtSEO-F1*_{Δ-244-136+Spc}-GFP_{ER}) and PC9 (*PMtSEO-F1*_{Δ-244-136}-GFP_{ER}) each lacked the segment from position -244 to -136 (Fig. 2A), which, as discussed above, contains a putative *SE_{act}* element (Fig. 1A; an overview of all constructs is provided in Figure S2). Both PC8 and PC9 generated fluorescence in immature SEs (Fig. 2B–D) identical to the pattern generated by the full-length promoter (Fig. 1B–D). Because SE specificity was not abolished in PC8 and PC9, we extended the internal deletion towards position -41. The resulting constructs PC10 (*PMtSEO-F1*_{Δ-244-41+Spc}-GFP_{ER}) and PC11 (*PMtSEO-F1*_{Δ-244-41}-GFP_{ER}) lacked the segment from position -244 to -41 (Fig. 2A), which had proven sufficient for SE-specific expression in PC6 (Fig. 1A). In cross-sections of petioles expressing PC10 and PC11, fluorescence was still restricted to the phloem (Fig. 2E), and typical SE

assemblies were visible in longitudinal sections (Fig. 2F). However, the fluorescence was also present in the adjacent CCs, which were characterised by their close association with SEs, their shape and the lack of sieve plates (Fig. 2G).

We decided to focus first on the localisation of *SE_{act}* elements before addressing questions concerning the unexpected activity of the PC10 and PC11 promoter constructs in CCs. The direct comparison of PC6 (Fig. 1A, I, K and L) and PC10/11 (Fig. 2A and E–G), which share no common *PMtSEO-F1* fragments, shows that the *MtSEO-F1* promoter contains at least two independent *SE_{act}* elements. One element, located between positions -244 and -41, activates expression in the SEs of plants transformed with construct PC6, and this is hereafter described as *SE_{act}I*. A second element (designated *SE_{act}II*) activates expression in the SEs of plants transformed with constructs PC10/11 and is probably located upstream of position -244 (Fig. 2A). In order to narrow down the location of *SE_{act}II* within this upstream region, we extended the internal deletion of PC10/11 upstream to -347 and obtained the constructs PC12 (*PMtSEO-F1*_{Δ-347-41+Spc}-GFP_{ER}) and PC13 (*PMtSEO-F1*_{Δ-347-41}-GFP_{ER}; Fig. 2A). In cross-sections of PC12 and PC13 petioles, fluorescence was still restricted to the phloem (Fig. 2H), but detailed micrographs of longitudinal sections showed that only CCs exhibited fluorescence (Fig. 2I). CCs were observed adjacent to non-fluorescent SEs, and were identified by specific staining of the sieve plates with aniline blue (Fig. 2J). Because SE-specific expression was abolished in PC12/13, but present in PC10/11, we concluded that *SE_{act}II* is located between positions -347 and -244.

To determine whether *SE_{act}II* is sufficient for SE-specific expression, we then created the constructs PC14 (*PMtSEO-F1*_{-347Δ-244-136+Spc}-GFP_{ER}) and PC15 (*PMtSEO-F1*_{-347Δ-244-136}-

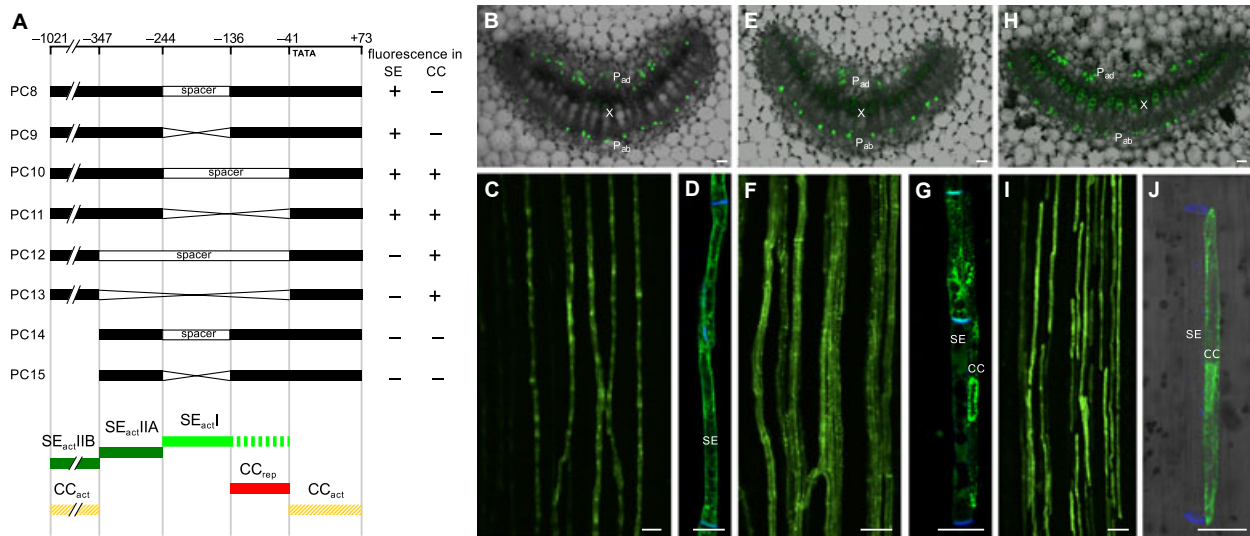


Fig. 2. Analysis of *PMtSEO-F1* constructs with internal deletions. **A:** Schematic map of *PMtSEO-F1* constructs with internal deletions and preliminary model of SE_{act} elements (green, alternative dashed), a CC_{rep} element (red) and CC_{act} elements (alternative localisations hatched yellow) within *PMtSEO-F1*. PC8: *PMtSEO-F1* $_{\Delta-244-136+SpC}$ -GFP_{ER}, PC9: *PMtSEO-F1* $_{\Delta-244-136}$ -GFP_{ER}, PC10: *PMtSEO-F1* $_{\Delta-244-41+SpC}$ -GFP_{ER}, PC11: *PMtSEO-F1* $_{\Delta-244-41}$ -GFP_{ER}, PC12: *PMtSEO-F1* $_{\Delta-347-41+SpC}$ -GFP_{ER}, PC13: *PMtSEO-F1* $_{\Delta-347-41}$ -GFP_{ER}, PC14: *PMtSEO-F1* $_{-347\Delta-244-136+SpC}$ -GFP_{ER}, PC15: *PMtSEO-F1* $_{-347\Delta-244-136}$ -GFP_{ER}; (B–J) GFP_{ER} fluorescence of transgenic tobacco plants was analysed in cross-sections through leaf petioles, shown as overlay of GFP_{ER} fluorescence and transmitted light (B, E, H). GFP_{ER} fluorescence of longitudinal petiole sections (C, D, F, G, I); GFP_{ER} fluorescence of longitudinal petiole section shown as overlay with transmitted light (J); (B–D) PC9; (E–G) PC11; (H–J) PC13. Sieve plates are stained with aniline blue (in D, G and J). P_{ab} , abaxial phloem; P_{ad} , adaxial phloem; X, xylem; scale bars = 25 μ m (B, C, E, F, H, I) and 12.5 μ m (D, G, J).

GFP_{ER}; Fig. 2A), which were similar to construct P2 in terms of 5' truncation (Fig. 1A) but also carried an internal deletion between positions -244 and -136 similar to PC8/9 (Fig. 2A). Interestingly, PC14 and PC15 produced no detectable fluorescence, indicating that the region between -347 and -244 is not sufficient on its own for SE-specific expression (we named this region $SE_{act}IIA$; Fig. 2A). A comparison between constructs PC14/15 (no SE expression) and PC8/9 (SE-specific activation), suggested the presence of a functional $SE_{act}II$ element upstream of position -347 (we named this region $SE_{act}IIB$; Fig. 2A). Even so, $SE_{act}IIB$ alone was also not sufficient to achieve expression in SEs, or constructs PC12/13 would have generated SE-localised fluorescence. We therefore conclude that $SE_{act}IIA$ and $SE_{act}IIB$ must work in a cooperative manner to confer SE-specific expression. Drawing the data together, SE-specific *MtSEO-F1* promoter activity appears to be controlled by the $SE_{act}I$ element (located between positions -244 and -41) as well as by the combination of elements $SE_{act}IIA$ and $SE_{act}IIB$ located in the adjacent upstream promoter regions between positions -1021 and -347, and -347 and -244, respectively (Fig. 2A).

Although activity of the native *MtSEO-F1* promoter is restricted to immature SEs (PC1), some of the internal deletion constructs activated expression additionally (PC10/11) or exclusively (PC12/13) in CCs, suggesting an activation/repression mechanism for CCs in the native promoter that is de-repressed in constructs PC10 through to PC13. The *cis*-regulatory elements required for transcriptional activation in CCs (CC_{act}) are still present in constructs PC10 through to PC13 and might therefore be located downstream of position -41 or upstream of position -347 in the promoter (Fig. 2A). In contrast, one or more *cis*-regulatory elements medi-

ating transcriptional repression in CCs (CC_{rep}) must be present in all promoter constructs except PC10 through to PC13, and must therefore be located between positions -136 and -41 (Fig. 2A). This region also contains $SE_{act}I$, which offers the possibility that $SE_{act}I$ and CC_{rep} could be independent, overlapping or even identical.

Analysis of *PMtSEO-F1*-*PAAtSUC2* hybrid constructs

In order to verify the activity and location of the CC_{rep} in an ectopic context and to confirm the position of $SE_{act}I$, we generated hybrid promoter constructs using the CC-specific *AtSUC2* promoter from *A. thaliana*. The internal deletion experiments described above suggested that the -244 to -41 segment of *PMtSEO-F1* influences cell type specificity within the phloem by controlling both SE activation and CC repression (Fig. 2A). We therefore inserted this fragment into the CC-specific *AtSUC2* promoter, ensuring that the distance to the TATA box was identical to that in its native promoter. A control construct was generated with a 203-bp *lacZ*-derived spacer inserted into the *PAAtSUC2* sequence at the same position. These constructs, PC16 (*PAAtSUC2*_{SpC}-GFP_{ER}) and PC17 (*PAAtSUC2*_{*PMtSEO-F1*-244-41}-GFP_{ER}), are shown in Fig. 3A.

In transgenic plants expressing the PC16 control construct, no fluorescence was observed in apical leaves, but petiole cross-sections of medial leaves revealed fluorescence in phloem tissue (Fig. 3B). Longitudinal sections of these petioles showed that the fluorescence was restricted to CCs (Fig. 3C), as confirmed by the aniline blue staining of sieve plates in adjacent, non-fluorescent SEs (Fig. 3D). This expression pattern was identical to that reported for the native *PAAtSUC2* sequence (Truernit & Sauer 1995; Stadler & Sauer

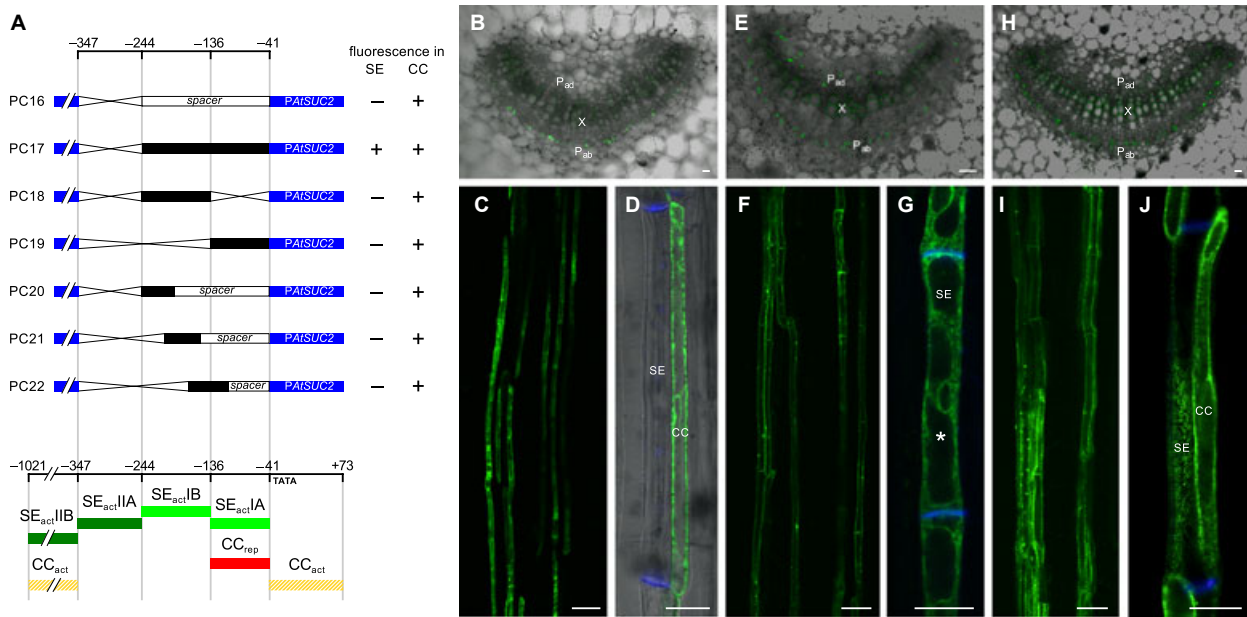


Fig. 3. Analysis of *PmtSEO-F1/PatSUC2* hybrid constructs. A: Schematic map of *PmtSEO-F1/PatSUC2* hybrid constructs where *PmtSEO-F1* fragments are represented by black bars and *PatSUC2* fragments are shown in blue. The model of SE_{act} elements (green), a CC_{rep} element (red) and CC_{act} elements (alternative localisations hatched yellow) within *PmtSEO-F1* is presented at the bottom. PC16: *PatSUC2*_{SPC}-GFP_{ER}, PC17: *PatSUC2*_{*PmtSEO-F1*-244-41}-GFP_{ER}, PC18: *PatSUC2*_{*PmtSEO-F1*-244-136}-GFP_{ER}, PC19: *PatSUC2*_{*PmtSEO-F1*-136-41}-GFP_{ER}, PC20: *PatSUC2*_{*PmtSEO-F1*-244-189+5pc}-GFP_{ER}, PC21: *PatSUC2*_{*PmtSEO-F1*-209-154+5pc}-GFP_{ER}, PC22: *PatSUC2*_{*PmtSEO-F1*-174-114+5pc}-GFP_{ER}; (B–J) GFP_{ER} fluorescence of transgenic tobacco plants was analysed in cross-sections through leaf petioles, shown as overlay of GFP_{ER} fluorescence and transmitted light (B, E, H), and in longitudinal petiole sections (C, D, F, G, I, J); (B–D) GFP_{ER} fluorescence in a medial leaf petiole from a transgenic tobacco plant expressing PC16, where D shows an overlay of GFP_{ER} fluorescence and transmitted light; (E–G) GFP_{ER} fluorescence in an apical leaf petiole from a transgenic tobacco plant expressing PC17; (H–J) GFP_{ER} fluorescence in a medial leaf petiole from a transgenic tobacco plant expressing PC17; sieve plates are stained with aniline blue (in D, G and J). P_{ab} , abaxial phloem; P_{ad} , adaxial phloem; X, xylem; scale bars = 25 μ m (B, C, E, F, H, I) and 12.5 μ m (D, G, J).

1996; Wright *et al.* 2003). In transgenic plants expressing the hybrid promoter (PC17), phloem-specific fluorescence was observed in apical and medial leaves, as seen in petiole cross-sections (Fig. 3E and H). The cross sections also display xylem autofluorescence, which has already been reported by Imlau *et al.* (1999). Longitudinal sections of apical leaf petioles showed fluorescence exclusively in SEs, as demonstrated by the typical end-to-end connections (Fig. 3F), aniline blue-stained sieve plates and the presence of vacuoles (Fig. 3G). In medial leaves of PC17 plants, longitudinal sections revealed fluorescence in both CCs and SEs (Fig. 3I). This is clearly shown in the aniline blue-stained sections, in which fluorescence is present in SEs with stained sieve plates as well as in the adjacent CCs (Fig. 3J).

The fact that the *PmtSEO-F1* segment -244 to -41 made the CC-specific *AtSUC2* promoter ectopically active in immature SEs confirmed the presence of the $SE_{act}I$ element in this region. However, as the hybrid construct PC17 was still expressed in CCs, it is clear that the CC_{rep} sequence located between positions -136 and -41 was not functional in the heterologous background of the *AtSUC2* promoter. One potential explanation is that the CC-specific transcriptional activation conferred by the *AtSUC2* promoter overwhelms the *PmtSEO-F1* CC_{rep} , perhaps simply because more CC_{act} sequences are present in the *AtSUC2* promoter resulting in a bias towards activation (an additive model). Alternatively, because transcriptional activators bound to CC_{act} sequences

in the *AtSUC2* promoter interact with the putative transcriptional repressor that binds CC_{rep} , its silencing activity could be quashed (an epistatic model). A further potential explanation is that the transcriptional repression mediated by CC_{rep} is strictly distance-dependent, *i.e.*, effective silencing would require the presence of both positive and negative regulatory proteins on the promoter sequence and precise spacing between the *cis*-regulatory elements involved – a prerequisite that may not be satisfied in the hybrid promoter. Finally, the structures of the transcriptional regulators binding to *PatSUC2* and *PmtSEO-F1* may be incompatible, *i.e.*, the CC_{rep} activity of the PC17 hybrid promoter cannot be achieved because the repressor that binds this sequence does not interact with activators binding elsewhere in the *AtSUC2* promoter, and is therefore excluded from the transcription complex.

Nevertheless, the *AtSUC2* promoter can serve as a tool to narrow down the position of $SE_{act}I$, taking advantage of the relatively high fluorescence intensity of the hybrid constructs in the absence of native *PmtSEO-F1* fragments (such as $SE_{act}I$ and $SE_{act}IIB$) that could potentially interfere with data interpretation. Constructs PC6 (Fig. 1A) and PC17 (Fig. 3A) prove that $SE_{act}I$ lies between positions -244 and -41. We dissected this sequence at position -136 and inserted the two fragments separately into the *AtSUC2* promoter, resulting in the constructs PC18 (*PatSUC2*_{*PmtSEO-F1*-244-136}-GFP_{ER}) and PC19 (*PatSUC2*_{*PmtSEO-F1*-136-41}-GFP_{ER}; Fig. 3A). Again,

control constructs with nonspecific *lacZ*-derived spacer sequences of similar length were created (data not shown), leading to an expression pattern identical to that of PC16 (Fig. 3B–D) and the native *AtSUC2* promoter. The analysis of PC18 and PC19 transgenic plants revealed a CC-specific expression pattern identical to that of PC16 plants (Fig. 3B–D), but there was no fluorescence in SEs (data not shown). These results indicate that $SE_{act}I$ cannot be localised to the upstream fragment (between positions -244 and -136) or the downstream fragment (between positions -136 and -41; Fig. 3A). This could mean that $SE_{act}I$ spans position -136 thus becoming incomplete and non-functional in both PC18 and PC19. Alternatively, $SE_{act}I$ may be located between positions -244 and -136 but its activity is strictly distance-dependent relative to the TATA box, which is not maintained in construct PC18. Finally, it is possible that $SE_{act}I$ comprises two or more motifs in different fragments, both required for correct SE-specific expression, as is the case for $SE_{act}IIA$ and $SE_{act}IIB$.

In order to address all three questions at once, we generated the three linker-scanning *PAtSUC2_{PMtSEO-F1}* hybrid constructs PC20 (*PAtSUC2_{PMtSEO-F1-244-189+Spc}-GFP_{ER}*), PC21 (*PAtSUC2_{PMtSEO-F1-209-154+Spc}-GFP_{ER}*) and PC22 (*PAtSUC2_{PMtSEO-F1-174-114+Spc}-GFP_{ER}*), which included *lacZ*-derived spacer sequences to maintain the correct distance between the *PMtSEO-F1* fragments and the TATA box (Fig. 3A). As already observed for PC18 and PC19, all three constructs PC20, PC21 and PC22 showed an identical expression pattern to the control construct PC16 (Fig. 3B–D) lacking any SE-specific expression.

Taken together, the inserted overlapping *PMtSEO-F1* fragments in PC20 to PC22 cover the *PMtSEO-F1* fragment present in PC18 while keeping distance dependency relative to the TATA box (Fig. 3A), but do not mediate SE specificity.

The possibility that $SE_{act}I$ spans position -136 is refuted by the results observed in plants transformed with construct PC22, which does not activate expression in SEs. We can further deduce that a model including distance-dependent binding sites would have to be even more complex and would require more than one motif, with distance dependency relative to the TATA box. Therefore, the third proposed mechanism is favoured as it represents the simplest model of two distance-independent $SE_{act}I$ motifs. Consequently, $SE_{act}I$ is assumed to be composed of $SE_{act}IA$ (located between positions -136 and -41) and $SE_{act}IB$ (located between positions -244 and -136), both of which must be present to confer SE-specific expression (Fig. 3A).

The model of *PMtSEO-F1* organisation

Based on the different expression patterns of the various truncated, internally deleted and hybrid *PMtSEO-F1* constructs, we were able to develop a model for the localisation of *cis*-regulatory elements regulating cell specificity in the *MtSEO-F1* promoter (Fig. 4A). *PMtSEO-F1* appears to have a rather complex organisation with at least four different SE_{act} elements, $SE_{act}IA$, $SE_{act}IB$, $SE_{act}IIA$ and $SE_{act}IIB$, which are distributed over the promoter sequence (Fig. 4A). None of these *cis*-regulatory elements is sufficient on its own to achieve SE-specific expression, but $SE_{act}IA$ can act in concert with $SE_{act}IB$ to achieve SE-specific expression, as can $SE_{act}IIA$ plus $SE_{act}IIB$, suggesting a cooperative binding mechanism. The *MtSEO-F1* promoter also contains a combination of positive and negative *cis*-regulatory elements that prevent CC-specific expression. A CC_{act} element may be located upstream of position -347 or within the *PMtSEO-F1*₋₄₁₋₊₇₃ fragment, whereas a CC_{rep} element lies between positions -136 and -41, the same region that contains $SE_{act}IA$ (Fig. 4A).

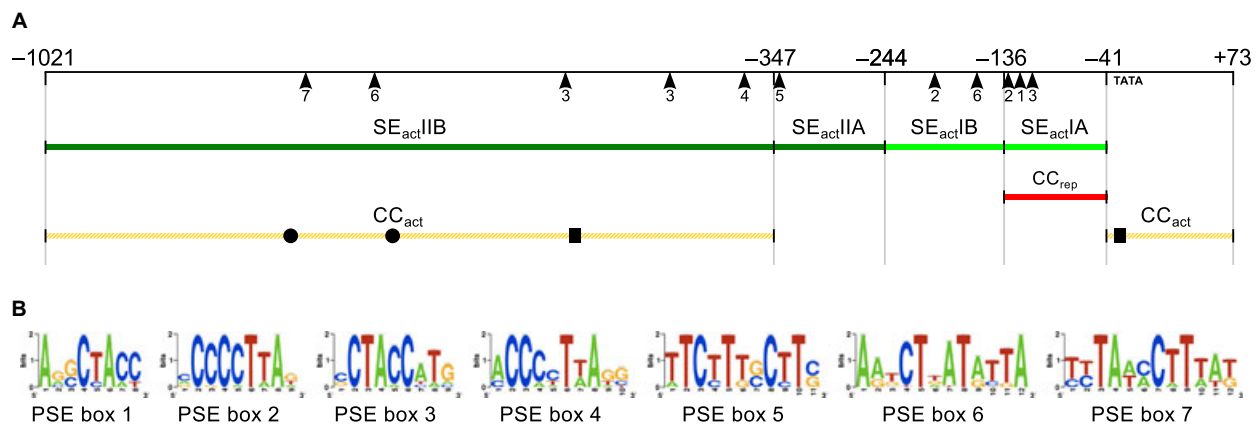


Fig. 4. *Cis*-regulatory elements in the *MtSEO-F1* promoter sequence. A: The model of *PMtSEO-F1* organisation shows promoter regions that are assumed to contain *cis*-regulatory elements activating SE-specific expression (SE_{act} , indicated in green) and CC-specific expression (CC_{act} , alternative localisations are hatched yellow) as well as one region containing a *cis*-regulatory element that represses expression in CCs (CC_{rep} , indicated in red). The positions of PSE boxes identified *de novo* by comparing SE-specific *SEO* promoters (see also B) are indicated with arrowheads displaying the number of the corresponding PSE box. Potential *cis*-regulatory elements resembling those previously reported to be involved in phloem-specific expression are shown as black circles (element 2-like boxes) and black boxes (EER element-like boxes) within the CC_{act} fragments. B: Sequence logos of PSE boxes that were generated *de novo* through the comparative analysis of all known SE-specific *SEO* promoter sequences using the Amadeus/Allegro software. The corresponding P-values are: PSE box 1: 4.1E-10, PSE box 2: 2.2E-12, PSE box 3: 4.1E-12, PSE box 4: 5.7E-14, PSE box 5: 6.7E-13, PSE box 6: 4.7E-15 and PSE box 7: 5.7E-15.

Cis-regulatory elements for activation and repression in CCs

The expression pattern conferred by the full-length *MtSEO-F1* promoter matches the observation that forisome synthesis takes place solely in immature SEs (Wergin & Newcomb 1970). SE development involves the degeneration of the nucleus and ribosomes, so forisome synthesis in differentiated SEs is no longer possible (Sjölund 1997). The ability to induce additional forisome synthesis in CCs and then transport the proteins into adjacent SEs would be the only way to replace forisomes after wounding. However, this mechanism seems rather unlikely because the forisome reaction is reversible and their native conformation is restored when SEs recover and return to physiological conditions (Knoblauch *et al.* 2001). Interestingly, forisomes are rarely detected in CCs (Wergin *et al.* 1975), and it seems likely that such events represent occasional malfunctions of the repression mechanism, resulting in leaky expression. This agrees with our finding that *PMtSEO-F1* activity in CCs can be observed when the putative CC_{rep} sequence is deleted. In such situations transcriptional activation is achieved by CC_{act} elements in the *PMtSEO-F1* promoter, located upstream of position -347 or downstream of position -41.

Numerous phloem-specific promoters have been characterised and several *cis*-regulatory elements involved in the activation of phloem-specific genes have been described. Some of these studies distinguished between specific cell types and revealed *cis*-regulatory elements that confer CC-specific expression, but no SE-specific elements have been described (Brears *et al.* 1991; Medberry & Olszewski 1993; Keller & Heierli 1994; Yin & Beachy 1995; Yin *et al.* 1997; Hehn & Rohde 1998; Baumann *et al.* 1999; Ayre *et al.* 2003; Guo *et al.* 2004; Guan & Zhou 2006; Schneidereit *et al.* 2008; Tsuwamoto & Harada 2010, 2011). We used the *PMtSEO-F1* sequence to search against the PLACE database, which collects the sequences of known *cis*-acting elements (Higo *et al.* 1999). We concentrated on motifs reported to confer CC-specific expression and searched for similar sequences in the promoter regions containing putative CC_{act} elements. We identified two copies of the CC-specific evening element-related element (EER) at positions -538 and -32 (Tsuwamoto & Harada 2010; Fig. 4A; Figure S1), although the latter encompasses the putative TATA box and is unlikely to be functional. We also found several copies of an element 2-like box, which contributes to the CC-specific activity of *PAtSUC2* (Schneidereit *et al.* 2008; Figure S1). However, only three copies of this element contained an intact Dof core motif (CTTT), which is thought to be crucial for the CC-specific activity of the *AtSUC2* promoter (Schneidereit *et al.* 2008). Two of those elements were found in the region beyond position -347 (Fig. 4A), and the other was localised to a promoter region that our experiments suggested did not contain a CC_{act} (Figure S1). If one of the element 2-like boxes is active in the context of *PMtSEO-F1*, additional motifs are likely to play an important role, as has been reported for the *AtSUC2* promoter, where an HD-ZIP motif is required for CC-specific activity (Schneidereit *et al.* 2008). Taken together, our data suggest that EER-related elements and the two element 2-like boxes described above are candidates for the putative CC_{act} element in *PMtSEO-F1*.

There is little evidence for the existence of CC_{rep} elements in other genes, although one exception is a 240-bp regulatory region that represses CC-specific activity in larger veins and is located upstream of an *A. thaliana* gene encoding an auxin-responsive transcriptional regulator (McGarry & Ayre 2008). Unfortunately, we could find no significant sequence similarities between this 240-bp region and the *PMtSEO-F1* promoter region containing the putative CC_{rep} (-136 to -41). As stated above, the potential CC_{rep} is located in the same promoter region as SE_{actIA} (Fig. 4A), which means that the two motifs could be independent or overlapping, but could also be the same multifunctional sequence. If there is a single SE_{actIA}/CC_{rep} motif, it may bind with a multifunctional transcription factor that interacts with a SE_{actIB} -binding factor in SEs to mediate SE-specific expression and with CC_{act} -binding factor in CCs to block CC-specific expression. *PMtSEO-F1* harbours at least four SE_{act} elements, so it is also possible that the greater number of binding sites allows the binding of several copies of a SE transcriptional activator, which might overwhelm and abolish the activity of CC_{act} . In some of our internal deletion constructs, one or more SE_{act} elements are deleted therefore shifting the balance more in favour of CC-specific activity.

Cis-regulatory elements mediating SE-specific expression

The *PMtSEO-F1* deletion analysis revealed the presence of four SE_{act} elements located in different promoter segments, each acting as part of a cooperative pair. Each SE_{act} pair (SE_{actI} and SE_{actII}) was independently capable of conferring SE-specific expression, but the elements in each pair (A and B) were mutually dependent (Fig. 4A). Such paired SE_{act} motifs are likely to have evolved by duplication events and it is reasonable to assume they comprise identical or similar sequences. Because of the lack of SE-specific *cis*-regulatory elements in the literature, we initially compared the four promoter segments that contain putative SE_{act} elements to look for repeated motifs (Fig. 4A). Integrating the fact that eukaryotic transcription factors usually bind to degenerate sequences (Latchman 1998; Courey 2001), we found a huge number of degenerate motifs of 6–12 bp in length that are shared between the different SE_{act} segments, making it difficult to speculate which of these motifs might be interesting as candidate binding sites. This limitation can be resolved by adding promoters of orthologous SE-specific genes to the search, resulting in the *de novo* identification of potential binding sites (Wasserman & Sandelin 2004). We therefore expanded our analysis to include additional *SEO* promoters known to be expressed in immature SEs, *i.e.*, *PMtSEO-F2*, *PMtSEO-F3* (Noll *et al.* 2009), *PGmSEO-F1* and *PAtSEOa* (Rüping *et al.* 2010). We used the Amadeus/Allegro software package to compare *PMtSEO-F1* with the abovementioned promoters and eight additional SE-specific *SEO* promoters from various dicotyledonous plants. In this way, we carried out a *de novo* search for sequence motifs that are statistically overrepresented in these *SEO* promoters compared to a background set of *Arabidopsis* promoters. This search yielded seven candidate motifs for SE-specific *cis*-regulatory elements, which we named putative sieve element (PSE) boxes. The sequence logos and P-values are provided in Fig. 4B (exact positions in the *PMtSEO-F1* sequence are shown in

Figure S1). Some PSE boxes occur more than once in the *PMtSEO-F1* sequence (Fig. 4B). This is the case for PSE box 2, which is a candidate for SE_{act}IA and IB, whereas PSE box 3 is present in the promoter segments considered for SE_{act}IA and IIB. PSE box 6 is a candidate motif for SE_{act}IB and IIB (Fig. 4A). Additional PSE boxes may also be candidates because the SE_{act} elements do not necessarily have to be alike. This has been reported for CC-specific *cis*-regulatory elements mediating transcriptional activation in the *AtSUC2* promoter, where DOF and HD-ZIP factors with different binding sites contribute to cell type specificity and only the elimination of both elements can abolish CC-specific expression completely (Schneidereit *et al.* 2008). The exact sequence motifs need to be defined by combinatorial deletion analysis and subsequent site-directed mutagenesis using the native *PMtSEO-F1* sequence, selected *PMtSEO-F1* deletion constructs or the *AtSUC2* promoter as background.

We cannot exclude the possibility that *PMtSEO-F1* is controlled by cryptic sequence motifs that have yet to be identified, but the experimentally identified SE_{act} elements shed light on the regulation of *PMtSEO-F1* activity and the PSE boxes identified *in silico* provide candidate motifs for SE_{act} elements that have not been described for any other promoter.

CONCLUSION

Our experiments have provided new insight into the complex regulation of the *MtSEO-F1* promoter, which involves combinatorial control by multiple *cis*-regulatory elements. Deletion analysis revealed the presence of both positive and negative *cis*-regulatory elements controlling CC-specific gene expres-

sion, even though the activity of the native *MtSEO-F1* promoter is restricted to immature SEs. We found that SE-specific expression is achieved by cooperative transcriptional regulation mediated by at least four SE_{act} elements. PSE boxes that are present in *PMtSEO-F1* and in many other SE-specific *SEO* promoters are excellent candidates for these SE_{act} elements. Dissecting the *cis*-acting regulatory elements involved in SE- and CC-specific gene expression could eventually lead to a better understanding of phloem identity by identifying regulatory differences between phloem cell types.

ACKNOWLEDGEMENTS

We gratefully acknowledge the technical assistance of Ann-Christin Müller, Claudia Hansen and Raphael Soeur. This work was supported by grants from the VolkswagenStiftung and the Fraunhofer Society.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. The *MtSEO-F1* promoter sequence displaying potential *cis*-regulatory elements.

Figure S2. Overview of all *PMtSEO-F1* promoter constructs from Figs 1A, 2A and 3A.

Table S1. List of primer sequences used for the cloning of *PMtSEO-F1* constructs.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

REFERENCES

- Ayre B.G., Blair J.E., Turgeon R. (2003) Functional and phylogenetic analyses of a conserved regulatory program in the phloem of minor veins. *Plant Physiology*, **133**, 1229–1239.
- Bauchner M., El Jaziri M., Vandeputte O. (2007) From primary to secondary growth: origin and development of the vascular system. *Journal of Experimental Botany*, **58**, 3485–3501.
- Baumann K., De Paolis A., Costantino P., Gualberti G. (1999) The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the *rolB* oncogene in plants. *The Plant Cell*, **11**, 323–334.
- van Bel A.J., Ehlers K., Knoblauch M. (2002) Sieve elements caught in the act. *Trends in Plant Science*, **7**, 126–132.
- Benfey P.N., Ren L., Chua N.H. (1990) Combinatorial and synergistic properties of CaMV 35S enhancer subdomains. *The EMBO Journal*, **9**, 1685–1696.
- Bevan M. (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Research*, **12**, 8711–8721.
- Bonke M., Thitamadee S., Mähönen A.P., Hauser M., Helariutta Y. (2003) APL regulates vascular tissue identity in *Arabidopsis*. *Nature*, **426**, 181–186.
- Bradford M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72**, 248–254.
- Brears T., Walker E.L., Coruzzi G.M. (1991) A promoter sequence involved in cell-specific expression of the pea glutamine synthetase *GS3A* gene in organs of transgenic tobacco and alfalfa. *The Plant Journal*, **1**, 235–244.
- Carlsbecker A., Helariutta Y. (2005) Phloem and xylem specification: pieces of the puzzle emerge. *Current Opinion in Plant Biology*, **8**, 512–517.
- Courey A.J. (2001) Regulatory transcription factors and *cis*-regulatory regions. In: Locker J. (Ed.), *Transcription factors*. Academic Press, Oxford, UK, pp 17–34.
- Doyle J.J., Doyle J.L. (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.
- Ehlers K., Knoblauch M., van Bel A.J.E. (2000) Ultrastructural features of well-preserved and injured sieve elements: minute clamps keep the phloem transport conduits free for mass flow. *Protoplasma*, **214**, 80–92.
- Fukuda H. (2004) Signals that control plant vascular cell differentiation. *Nature Reviews Molecular Cell Biology*, **5**, 379–391.
- Guan C., Zhou X. (2006) Phloem specific promoter from a satellite associated with a DNA virus. *Virus Research*, **115**, 150–157.
- Guo H., Chen X., Zhang H., Fang R., Yuan Z., Zhang Z., Tian Y. (2004) Characterization and activity enhancement of the phloem-specific pumpkin *PP2* gene promoter. *Transgenic Research*, **13**, 559–566.
- Hehn A., Rohde W. (1998) Characterization of *cis*-acting elements affecting strength and phloem specificity of the coconut foliar decay virus promoter. *Journal of General Virology*, **79**, 1495–1499.
- Higo K., Ugawa Y., Iwamoto M., Korenaga T. (1999) Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research*, **27**, 297–300.
- Hoekema A., Hirsch P.R., Hooykaas P.J.J., Schilperoort R.A. (1983) A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, **303**, 179–180.
- Horsch R.B., Klee H.J., Stachel S., Winans S.C., Nester E.W., Rogers S.G., Fraley R.T. (1986) Analysis of *Agrobacterium tumefaciens* virulence mutants in leaf discs. *Proceedings of the National Academy of Sciences USA*, **83**, 2571–2575.
- Ilegems M., Douet V., Meylan-Bettex M., Uyttewaal M., Brand L., Bowman J.L., Stieger P.A. (2010) Interplay of auxin, KANADI and Class III HD-ZIP transcription factors in vascular tissue formation. *Development*, **137**, 975.
- Imlau A., Truernit E., Sauer N. (1999) Cell-to-cell and long-distance trafficking of the green fluorescent protein in the phloem and symplastic unloading of the protein into sink tissues. *The Plant Cell*, **11**, 309–322.
- Jefferson R.A., Kavanagh T.A., Bevan M.W. (1987) GUS fusions: β -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal*, **6**, 3901–3907.
- Keller B., Heierli D. (1994) Vascular expression of the *gpr1.8* promoter is controlled by three specific regulatory elements and one unspecific acti-

- vating sequence. *Plant Molecular Biology*, **26**, 747–756.
- Knoblauch M., Peters W.S., Ehlers K., van Bel A.J.E. (2001) Reversible calcium-regulated stopcocks in legume sieve tubes. *The Plant Cell*, **13**, 1221.
- Latchman D.S. (1998) *Transcription factors: an overview*. Academic Press, San Diego, CA.
- Linhart C., Halperin Y., Shamir R. (2008) Transcription factor and microRNA motif discovery: the Amadeus platform and a compendium of metazoan target sets. *Genome Research*, **18**, 1180–1189.
- McGarry R., Ayre B. (2008) A DNA element between At4g28630 and At4g28640 confers companion-cell specific expression following the sink-to-source transition in mature minor vein phloem. *Planta*, **228**, 839–849.
- Medberry S.L., Olszewski N.E. (1993) Identification of *cis* elements involved in Commelina yellow mottle virus promoter activity. *The Plant Journal*, **3**, 619–626.
- Negrutiu I. (1981) Improved conditions for large-scale culture, mutagenesis, and selection of haploid protoplasts of *N. plumbaginifolia*. *Zeitschrift für Pflanzenphysiologie*, **104**, 431–442.
- Negrutiu I., Shillito R., Potrykus I., Biasini G., Sala F. (1987) Hybrid genes in the analysis of transformation conditions. *Plant Molecular Biology*, **8**, 363–373.
- Nishitani C., Demura T., Fukuda H. (2001) Primary phloem-specific expression of a *Zinnia elegans* homeobox gene. *Plant and Cell Physiology*, **42**, 1210–1218.
- Noll G.A., Fontanellaz M.E., Rüping B., Ashoub A., van Bel A.J.E., Fischer R., Knoblauch M., Prüfer D. (2007) Spatial and temporal regulation of the forisome gene *for1* in the phloem during plant development. *Plant Molecular Biology*, **65**, 285–294.
- Noll G.A., Rüping B., Ernst A.M., Bucsenetz M., Twyman R.M., Fischer R., Prüfer D. (2009) The promoters of forisome genes *MtSEO2* and *MtSEO3* direct gene expression to immature sieve elements in *Medicago truncatula* and *Nicotiana tabacum*. *Plant Molecular Biology Reporter*, **27**, 526–533.
- Palevitz B.A., Newcomb E.H. (1971) The ultrastructure and development of tubular and crystalline P-phloem: evidence for phloem loading and unloading by SUC2. *Planta*, **196**, 564–570.
- Tsuwamoto R., Harada T. (2010) Identification of a *cis*-regulatory element that acts in companion cell-specific expression of *AtMT2B* promoter through the use of *Brassica* vasculature and gene-gun-mediated transient assay. *Plant and Cell Physiology*, **51**, 80–90.
- Tsuwamoto R., Harada T. (2011) The *Arabidopsis* COR13 promoter contains two *cis*-acting regulatory regions required for transcriptional activity in companion cells. *Plant Cell Reports*, **30**, 1723–1733.
- Turatsinze J., Thomas-Chollier M., Defrance M., van Helden J. (2008) Using RSAT to scan genome sequences for transcription factor binding sites and *cis*-regulatory modules. *Nature Protocols*, **3**, 1578–1588.
- Wasserman W.W., Sandelin A. (2004) Applied bioinformatics for the identification of regulatory elements. *Nature Reviews Genetics*, **5**, 276–287.
- Wergin W.P., Newcomb E.H. (1970) Formation and dispersal of crystalline P-protein in sieve elements of soybean (*Glycine max* L.). *Protoplasma*, **71**, 365–388.
- Wergin W.P., Palevitz B.A., Newcomb E.H. (1975) Structure and development of P-protein in phloem parenchyma and companion cells of legumes. *Tissue and Cell*, **7**, 227–242.
- Wright K.M., Roberts A.G., Martens H.J., Sauer N., Oparka K.J. (2003) Structural and functional vein maturation in developing tobacco leaves in relation to *AtSUC2* promoter activity. *Plant Physiology*, **131**, 1555–1565.
- Xie B., Wang X., Zhu M., Zhang Z., Hong Z. (2011) *CalS7* encodes a callose synthase responsible for callose deposition in the phloem. *The Plant Journal*, **65**, 1–14.
- Yin Y., Beachy R.N. (1995) The regulatory regions of the rice tungro bacilliform virus promoter and interacting nuclear factors in rice (*Oryza sativa* L.). *The Plant Journal*, **7**, 969–980.
- Yin Y., Chen L., Beachy R.N. (1997) Promoter elements required for phloem-specific gene expression from the RTBV promoter in rice. *The Plant Journal*, **12**, 1179–1188.
- protein in the sieve elements of certain papilionaceous legumes. *Protoplasma*, **72**, 399–426.
- Prestridge D.S. (1991) SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Computer Applications in the Biosciences*, **7**, 203–206.
- Ruijter N., Verhees J., van Leeuwen W., van der Krol A. (2003) Evaluation and comparison of the GUS, LUC and GFP reporter system for gene expression studies in plants. *Plant Biology*, **5**, 103–115.
- Rüping B., Ernst A.M., Jekat S.B., Nordzicke S., Reincke A.R., Müller B., Bornberg-Bauer E., Prüfer D., Noll G.A. (2010) Molecular and phylogenetic characterization of the sieve element occlusion gene family in *Fabaceae* and non-*Fabaceae* plants. *BMC Plant Biology*, **10**, 219.
- Schneidererit A., Imlau A., Sauer N. (2008) Conserved *cis*-regulatory elements for DNA-binding-with-one-finger and homeo-domain-leucine-zipper transcription factors regulate companion cell-specific expression of the *Arabidopsis thaliana* *SUCROSE TRANSPORTER 2* gene. *Planta*, **228**, 651–662.
- Sieburth L.E., Deyholos M.K. (2006) Vascular development: the long and winding road. *Current Opinion in Plant Biology*, **9**, 48–54.
- Sjölund R.D. (1997) The phloem sieve element: a river runs through it. *The Plant Cell*, **9**, 1137–1146.
- Skirycz A., Reichelt M., Burow M., Birkemeyer C., Rolcik J., Kopka J., Zanon M.I., Gershenzon J., Strnad M., Szopa J., Mueller-Roeber B., Witt I. (2006) DOF transcription factor AtDof1.1 (OBP2) is part of a regulatory network controlling glucosinolate biosynthesis in Arabidopsis. *The Plant Journal*, **47**, 10–24.
- Stadler R., Sauer N. (1996) The *Arabidopsis thaliana* *AtSUC2* gene is specifically expressed in companion cells. *Botanica Acta*, **109**, 299–306.
- Thompson M.V., Wolniak S.M. (2008) A plasma membrane-anchored fluorescent protein fusion illuminates sieve element plasma membranes in Arabidopsis and tobacco. *Plant Physiology*, **146**, 1599–1610.
- Truernit E., Sauer N. (1995) The promoter of the *Arabidopsis thaliana* *SUC2* sucrose-H⁺ symporter gene directs expression of β -glucuronidase to the