*The Plant Journal* (2005) **44**, 62–75 doi: 10.1111/j.1365-313X.2005.02510.x

# Metabolic engineering of proanthocyanidins by ectopic expression of transcription factors in *Arabidopsis thaliana*

**Shashi B. Sharma and Richard A. Dixon***
*Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA*

Received 3 May 2005; revised 28 June 2005; accepted 29 June 2005.
*For correspondence (fax +1 580 224 6601; e-mail radixon@noble.org).

**Summary**

**Genetic transformation of *Arabidopsis thaliana* with the Arabidopsis *TT2* MYB transcription factor resulted in ectopic expression of the *BANYULS* gene, encoding anthocyanidin reductase, *AHA10* encoding a P-type proton-pump and *TT12* encoding a transporter involved in proanthocyanidin biosynthesis. When coupled with constitutive expression of *PAP1*, a positive regulator of anthocyanin biosynthesis, *TT2* expression in Arabidopsis led to accumulation of proanthocyanidins, but only in a subset of cells in which the *BANYULS* promoter is naturally expressed. Ectopic expression of the maize *Lc* MYC transcription factor weakly induced *AHA10* but did not induce *BANYULS*, *TT12* or accumulation of proanthocyanidins. However, high-level combined expression of *TT2*, *PAP1* and *Lc* resulted in proanthocyanidin synthesis throughout young leaves and cotyledons, followed by death of the plants 1 to 2 weeks after germination. We discuss these results in relation to engineering proanthocyanidins to improve the quality of food and forage plants.**

**Keywords: flavonoid, anthocyanin, condensed tannin, MYB transcription factor, bHLH transcription factor, Lc, PAP1.**

## Introduction

Proanthocyanidins (PAs, also called condensed tannins) are dimers or higher oligomers of flavan-3-ol units and are prominent compounds in the seed coats, leaves, fruits, flowers and bark of many plant species (Marles *et al.*, 2003). Proanthocyanidins and their monomeric building blocks, such as epicatechin, catechin and epigallocatechin, are powerful antioxidants and have recently attracted considerable interest owing to their beneficial effects on cardiac health, immunity and longevity (Dixon *et al.*, 2005). The presence of PAs in forage plants protects ruminant animals from pasture bloat due to their ability to bind protein, thereby preventing its rapid fermentation in the rumen and consequent formation of excessive methane and foam (Aerts *et al.*, 1999; McMahon *et al.*, 2000). Since protein–PA complexes are resistant to microbial digestion in the rumen, the presence of PAs in forage crops also leads to increased bypass protein (plant protein leaving the rumen without degradation) available for ruminant nutrition (Barry and McNabb, 1999). Many important forage crops such as alfalfa and white clover lack PAs in their leaves and stems but accumulate them in the seed endothelium. A major goal for improving forage legumes such as alfalfa is therefore to switch on the biosynthesis of PAs in the aerial organs.

Much of our recent understanding of the biosynthesis of PAs has arisen from the genetic and biochemical analysis of mutants of the model plant *Arabidopsis thaliana* that fail to accumulate PAs in their seeds and therefore exhibit a transparent testa (tt) phenotype due to loss of brown pigmentation imparted by oxidized PAs (Shirley *et al.*, 1995). Proanthocyanidin biosynthesis branches from the general phenylpropanoid/flavonoid pathway after the synthesis of anthocyanidins, and mutations which block anthocyanidin biosynthesis therefore also lead to a tt phenotype (Figure 1) (Shirley *et al.*, 1995; Xie *et al.*, 2003).

Anthocyanidin reductase, the product of the *BANYULS* (*BAN*) gene (Devic *et al.*, 1999) catalyzes a committed step in PA biosynthesis by converting anthocyanidins into flavan-3-ols such as (–)-epicatechin (Xie *et al.*, 2003), which are possibly sequestered into vacuoles by a member of the multidrug and toxic compound extrusion (MATE) family of secondary transporter-like proteins encoded by *TT12* (Debeaujon *et al.*, 2001). In addition, a member of the glutathione *S*-transferase (GST) family encoded by *TT19* is also

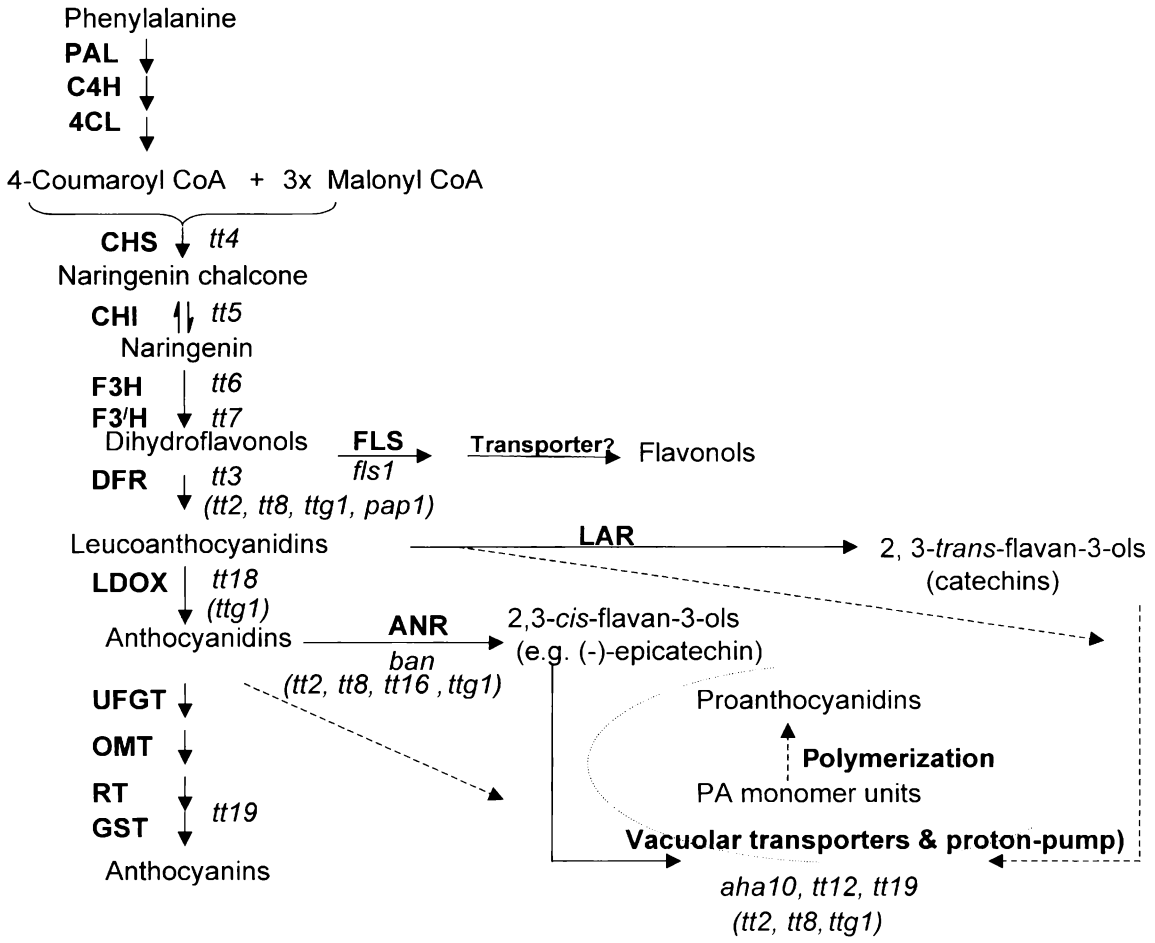


**Figure 1.** The flavonoid biosynthetic pathway leading to the synthesis of anthocyanins, flavonols and PAs.
Solid lines indicate the route supported by biochemical and genetic evidence, whereas dotted lines indicate hypothetical routes. Enzymes are presented in uppercase bold letters, with corresponding genetic loci in Arabidopsis given in lower case italics and mutants of the regulatory genes affecting their expression in Arabidopsis in brackets. AHA10, P-type $H^+$-ATPase; ANR, anthocyanidin reductase; *ban, banyuls*; C4H, cinnamate 4-hydroxylase; CHI, chalcone isomerase; CHS, chalcone synthase, 4CL, 4-coumarate:CoA ligase; DFR, dihydroflavonol reductase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; FLS, flavonol synthase; GST, glutathione *S*-transferase; LAR, leucoanthocyanidin reductase (this enzyme is apparently not present in Arabidopsis); LDOX, leucoanthocyanidin dioxygenase; OMT, *O*-methyltransferase; PAL, phenylalanine ammonia-lyase; *pap1, production of anthocyanin pigment*; RT, rhamnosyl transferase; UFGT, uridine diphosphate glucose-flavonoid glucosyltransferase; *tt*, transparent testa; *ttg1*, transparent testa glabrous.

involved in accumulation of PAs in seed endothelium. TT19 most likely functions as a flavonoid carrier protein (Kitamura *et al.*, 2004), although the exact biochemical functions of TT12 and TT19 remain unknown. A recently characterized proton pumping ATPase, AHA10, is also required for accumulation of PA in Arabidopsis (Baxter *et al.*, 2005). Although predicted to be localized to the plasma membrane, AHA10 might possibly function in vesicular trafficking. Once monomers accumulate in the vacuole, they are polymerized into PAs by as yet unknown enzymatic or non-enzymatic mechanisms (Dixon *et al.*, 2005).

The regulatory circuits involved in PA biosynthesis in Arabidopsis are complex and require coordinated regulation of genes involved in: (i), the general flavonoid biosynthesis pathway to anthocyanidins (Winkel-Shirley, 2001) (Figure 1); (ii), the PA branch pathway controlling the biosynthesis of the PA subunits, their transport into the vacuole and their subsequent polymerization into PA polymers (Dixon *et al.*, 2005) (Figure 1); and (iii), developmental processes for the generation of specialized cells in which PAs are deposited (Johnson *et al.*, 2002; Nesi *et al.*, 2002; Sagasser *et al.*, 2002). There appear to be multiple flavonoid pathway transcription factors in Arabidopsis, and over-expression of the Arabidopsis orthologs of the maize *C1* gene, *PAP1* or *PAP2*, results in constitutive anthocyanin production in Arabidopsis and tobacco (Borevitz *et al.*, 2000; Tohge *et al.*, 2005). A second class of regulatory genes that modulate late flavonoid biosynthetic genes encoding dihydroflavonol reductase (DFR) and leucoanthocyanidin dioxygenase (LDOX), as well as *BAN* and *TT12*, includes *TT2*, *TT8* and *TTG1* (Figure 1), which encode R2-R3 MYB, MYC/bHLH and WD40 repeat domain proteins, respectively (Nesi *et al.*, 2000, 2001; Walker

*et al.*, 1999). *TT2* is a master regulator as its mutation affects the expression of *DFR*, *LDOX* and *BAN* in developing siliques, whereas its ectopic expression induces *BAN* and *TT8* (the latter in roots at least) and up-regulates *DFR* and *LDOX* (Debeaujon *et al.*, 2003; Nesi *et al.*, 2000, 2001). A third category of transcription factors, which includes TT1, TT16 and TTG2, regulates organ and cell development for PA deposition, in addition to transcription of PA pathway-specific genes (Johnson *et al.*, 2002; Nesi *et al.*, 2002; Sagasser *et al.*, 2002). It is not yet known how the *AHA10* gene is regulated.

Several of the maize *myc* (*R*, *B-Peru*, *Sn*, *Lc*) and *myb* (*C1*, *P*) flavonoid regulatory genes have been tested for their ability to influence accumulation of flavonoids, anthocyanins or PAs when expressed in heterologous plants (Bovy *et al.*, 2002; Bradley *et al.*, 1998; deMajnik *et al.*, 2000; Ray *et al.*, 2003; Robbins *et al.*, 2003). *Lc* expression in alfalfa affected accumulation of anthocyanin and PA; these effects were dependent on high light intensity and low temperature (Ray *et al.*, 2003). Constitutive expression of the maize *Sn myc* gene in *Lotus corniculatus* had subtle effects on anthocyanin accumulation, which were restricted to the leaf mid-rib, leaf base and petiole, but stimulated PA accumulation in leaf tissue known to synthesize PAs (Robbins *et al.*, 2003). *B-Peru* and *C1* stimulated anthocyanin synthesis when co-bombarded into white clover (*Trifolium repens*) or pea (*Pisum sativum*) in transient assays (deMajnik *et al.*, 2000).

Here we explore the use of ectopic expression of *TT2*, *PAP1* and *Lc* in engineering PAs in Arabidopsis. *TT2* was either expressed alone or in combination with Arabidopsis *PAP1* for up-regulation of the anthocyanin pathway necessary for the biosynthesis of PA. Ectopic expression of *TT2* induced *BAN*, *TT12*, and *AHA10*, highlighting the importance of MYB transcription factors in coordinated genetic manipulation of natural product pathways for phenylpropanoids. Ectopic expression of *Lc* only induced *AHA10*, whereas over-expression of *PAP1* did not induce *BAN*, *TT12* or *AHA10*. The expression of *TT8*, *DFR*, *CHS*, *LDOX*, *TTG1*, *TTG2* or *TT19* was shown to be constitutive in the leaf tissue of short-day grown plants. Plants ectopically expressing *TT2* and *PAP1* accumulated PAs with a highly selective tissue-specific expression, reflecting the wild-type expression pattern of the Arabidopsis *BAN* promoter.

## Results

### *Tissue-specific expression of the Arabidopsis* BAN *promoter*

Previous studies have suggested that *BAN* expression in Arabidopsis is limited to the seed endothelium (Debeaujon *et al.*, 2003). To address whether *BAN* expression, and therefore the potential for synthesis of epicatechin and/or PAs, can occur in additional cell types, the 5′-upstream region of the Arabidopsis *BAN* gene [from positions 1152 to 1821 of the intergenic region between the Arabidopsis genes *AT1G61730* and *AT1G61720* (*BAN*), see Figure 2a] was isolated by PCR from genomic DNA, and fused independently to the *gusA* (*uidA*) and *GFP* reporter genes. Constructs were introduced into Arabidopsis and tissues examined by histochemistry (for GUS) or confocal microscopy [for green fluorescent protein (GFP)]. Staining of the *BAN* promoter:GUS (*Pro*$_{BAN}$*:gusA*) transgenic plants with X-gluc reagent revealed expression from the *BAN* 5′-upstream region in the mid-rib and hydathodes of rosette leaves, the ovules in the silique, the petal veins, the peduncle, the cortex of the hypocotyl, the roots, the puffs of root hairs, especially at the junction of the root and hypocotyl, and the stipules at the base of rosette leaves, as shown in Figure 2(b–j). These patterns were observed in four independently transformed lines. The previously reported expression pattern in seeds (Devic *et al.*, 1999) was confirmed by analysis of transgenic Arabidopsis plants expressing the *Pro*$_{BAN}$*:GFP* construct (Figure 2k–m). Overall, the results indicate that expression of *BAN* in Arabidopsis has somewhat broader tissue specificity than previously reported, but nevertheless only occurs in a specific subset of cell types.

### *Effects of constitutive expression of* TT2 *on expression of flavonoid pathway genes*

The open reading frame of the Arabidopsis *TT2* gene was isolated by RT-PCR, and introduced on construct pSB239 (Figure 3) into wild-type Col-0 and the *pap1-D* activation tagged mutant of Col-0 which ectopically over-expresses the anthocyanin pathway (Borevitz *et al.*, 2000) and should therefore contain potential substrates for biosynthesis of PA. Leaves of transgenic plants showing monogenic segregation for resistance conferred by the selectable marker were further analyzed by RT-PCR for the expression profile of the *TT2*, *BAN*, *TT12*, *AHA10* and *PAP1* genes. Lines homozygous for the selectable marker were then analyzed for *TT2*, *BAN*, *TT12*, *AHA10*, *DFR*, *TT19*, *CHS*, *PAP1*, *TT1*, *TT16*, *TTG1* and *TTG2* transcripts by RT-PCR (Figure 4).

Figure 4(a) shows RT-PCR data for individual $T_2$ generation plants, with the numbers before the dash referring to independent transgenic lines generated in the *pap1-D* background using pSB239 or an empty vector. Ectopic expression of the *TT2* transgene is seen in each of the independent pSB239 generated transgenic lines, and *TT2* is clearly not expressed in the leaf tissue of the empty vector controls. *PAP1* is expressed in all lines, since it is under the control of a multiple 35S promoter activation tag in the *pap1-D* line, although its expression level is quite variable. With the exception of line 24-1, each line expressing the *TT2* transgene also shows ectopic expression of *BAN*, which is not expressed in leaves of the empty vector controls. *TT12*, encoding a potential transporter for PA monomers

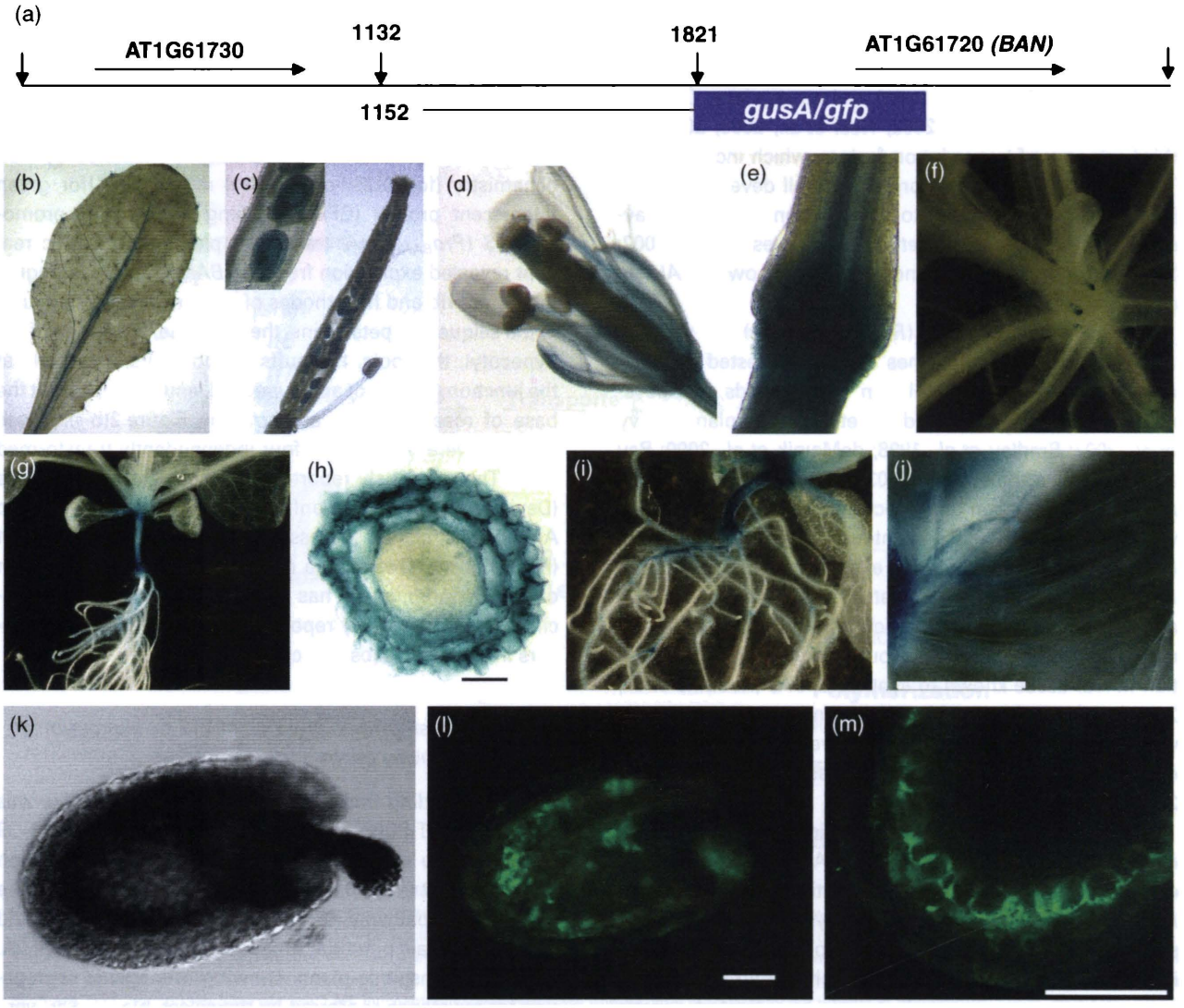


**Figure 2.** Cell-specific expression of *BAN* in Arabidopsis.
(a) The Arabidopsis *BAN* 5′-upstream region used in the present work. Positions 1132 and 1821 are the translation termination site of *AT1G61730* and the translation initiation site of the *BAN* gene, respectively.
(b–j) Cell-specific expression of the *BAN* gene revealed by $Pro_{BAN}$:*gusA* with X-gluc staining: (b) mid-rib and hydathodes of rosette leaves; (c) ovules in the silique; (d) petal veins; (e) peduncle; (f) stipules at the base of rosette leaves; (g) cortex of the hypocotyls; (h) cross-section of the hypocotyl close to the root junction; (i) roots and puffs of root hairs at the junction of root and hypocotyls; (j) magnified view of root hairs.
(k–m) Cell-specific expression of a $Pro_{BAN}$:*GFP* reporter construct in the young seed endothelium: (k) brightfield and (l) corresponding confocal fluorescence image of young seed; (m) cell-specific *BAN* promoter expression in the seed endothelium layer.
Bars represent 50 μm in panels (l) and (m), 100 μm in panel (h) and 200 μm in panel (j)

(Debeaujon *et al.*, 2001), is expressed in some, but not all, of the *TT2* transgenic lines. Expression of *TT12* may require higher levels of *TT2* expression than does that of *BAN*.

Figure 4(b,c) shows a more extended data set for a number of homozygous $T_3$ transgenic plants, and null segregants, in the Col-0 or *pap1-D* backgrounds grown under short days to promote synthesis of anthocyanins. Again, a clear relationship exists between expression of *TT2* and expression of *BAN* and *TT12*. *AHA10* transcripts were also measured in these lines, and are clearly co-expressed with *BAN* and *TT12*. However, other genes necessary for PA biosynthesis, namely *TT3* (encoding dihydroflavonol reductase, DFR), *LDOX* (encoding leucoanthocyanidin reductase, also known as anthocyanidin synthase), *TT19* (encoding a putative GST involved in monomer transport) (Kitamura *et al.*, 2004), *TT4* (encoding chalcone synthase, CHS) and the transcription factors TTG1, TTG2 and TT8, were constitutively expressed and unaffected by ectopic expression of *TT2* in plants grown under short-day conditions. Note that these genes were also weakly expressed in the *tt2* mutant line. Expression of some of these genes was slightly enhanced in the *pap1-D* background. The TT16 transcription

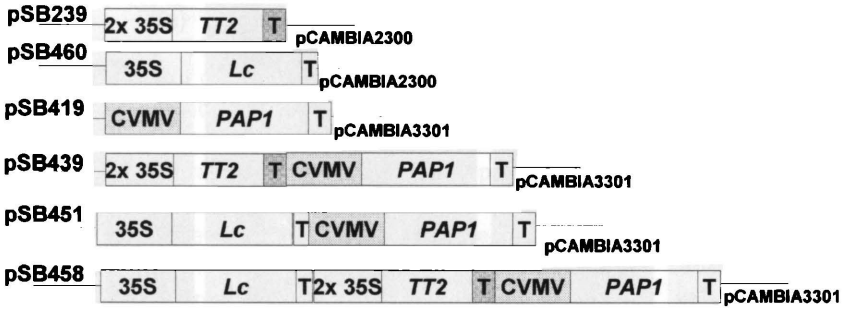


**Figure 3.** Plasmid constructs.
The name of the plasmid on the right-hand side represents the backbone plasmid used to insert the various genes resulting in the new plasmid named on the left-hand side. 35S, 2 × 35S, CVMV, *TT2*, *Lc*, *PAP1* and T inside boxes represent the 35S CaMV promoter, double 35S CaMV promoter and Cassava Mosaic Virus promoter, *TT2*, *Lc* and *PAP1* coding regions, and 35S (shaded) or *nopaline synthase* transcription terminator, respectively.

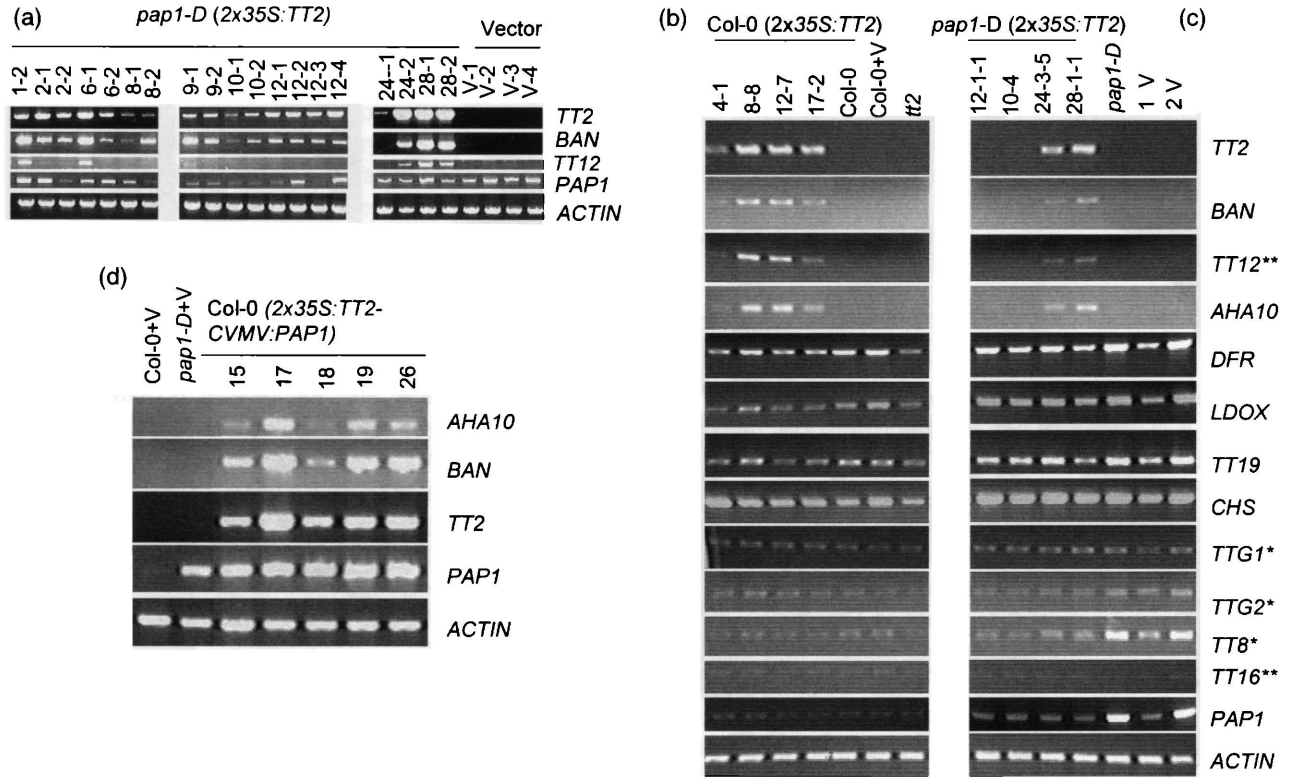


**Figure 4.** Transcript levels in leaves of Col-0 and *pap1-D* Arabidopsis expressing *TT2*, as determined by RT-PCR.
(a) *pap1-D* transformed with Arabidopsis *TT2* (pSB239) or empty vector (pCAMBIA2300). Plants were $T_2$ generation, and RT-PCR was for 30 cycles. First numbers refer to independent 2 × *35S:TT2* transgenic *pap1-D* lines generated using pSB239 and second numbers to individual segregating progeny derived from the same $T_1$ plants. V-1 to V-4 are independent *pap1-D* lines transformed with an empty vector.
(b, c) Homozygous 2 × *35S:TT2* transgenic Arabidopsis and null segregants in Col-0 (b) or *pap1-D* (c) backgrounds. Transcripts were amplified for 21 cycles; genes which showed low amplification at 21 cycles were amplified for 23 cycles (*) or 30 cycles (**). Col + V, empty vector control in Col-0 background; 1V and 2V, empty vector controls in the *pap1-D* background showing no purple pigmentation or reduced pigmentation in their leaves, respectively.
(d) Col-0 transformed with 2 × *35S:TT2* and CVMV:*PAP1* (pSB439). Plants were $T_1$ generation, and RT-PCR was for 21 cycles. Numbers refer to independent transgenic lines.

factor was very weakly expressed in all lines, independently of PAP1 expression. The three transcription factor genes *TTG1*, *TTG2*, and *TT8*, were expressed at lower levels than the genes for the biosynthetic enzymes, requiring additional rounds of PCR amplification for their visualization (Figure 4b,c). The TT1 transcription factor was not expressed in the leaves of plants co-expressing TT2 and PAP1 (data not shown).

### *Effects of constitutive expression of* TT2 *on flavonoid and PA accumulation*

Transgenic $T_1$ and $T_2$ plants expressing *TT2* in the *pap1-D* background were grown under short-day conditions and then stained with dimethylaminocinnamaldehyde (DMACA) reagent to indicate the localization of PAs. Blue staining was only observed in those plants expressing *PAP1* and strongly

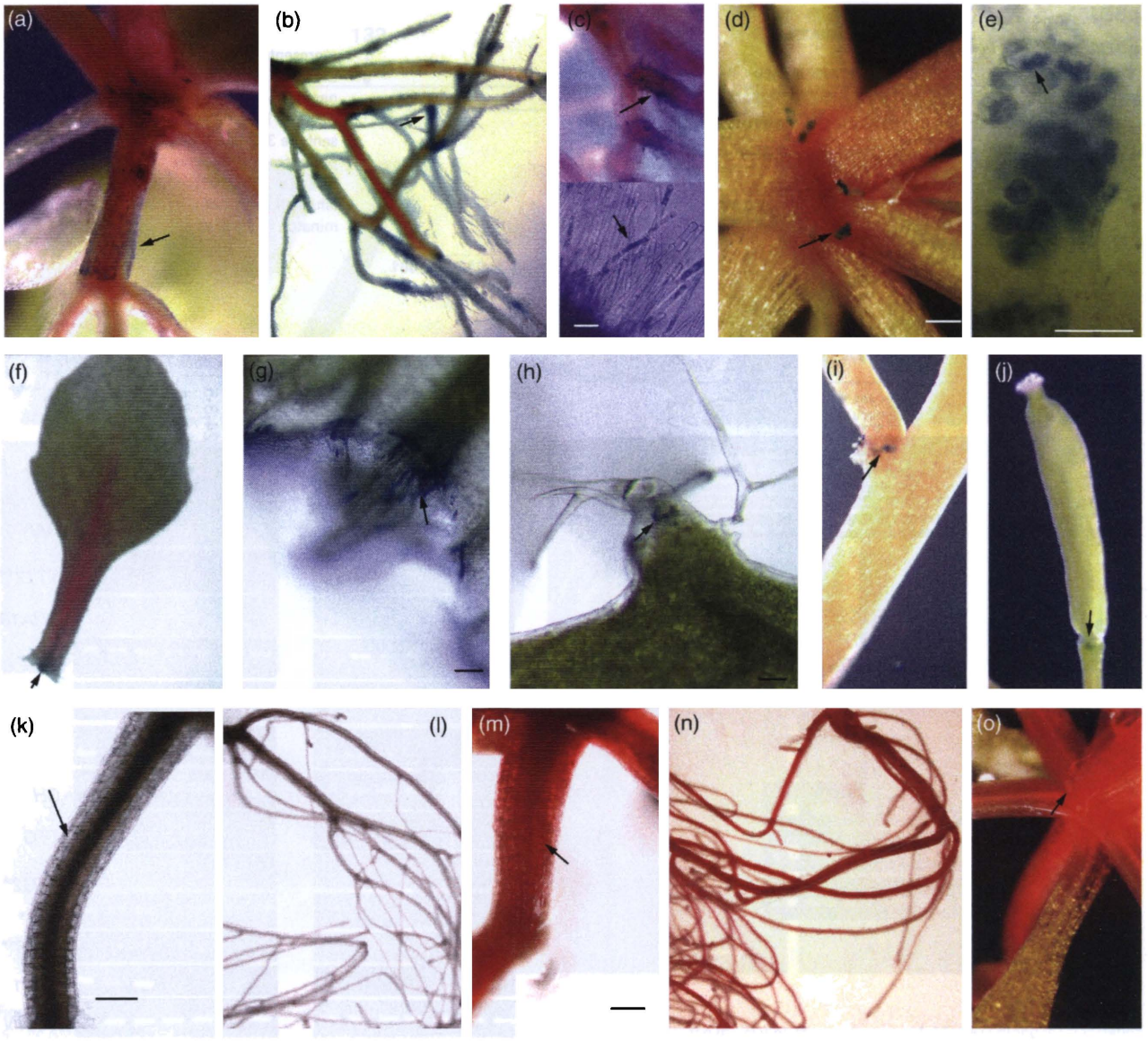


**Figure 5.** DMACA staining to show localization of PAs in *pap1-D* plants expressing *TT2* (line 28), and in Col-0 and *pap1-D* control lines. Additional lines (2, 6, 12 and 24) exhibited an identical staining pattern.
Figures show: (a) outer cortex of hypocotyls; (b) roots; (c) root hairs at the junction of the primary and secondary roots (insert shows a magnified view to visualize root hairs); (d) stipules at the base of rosette leaves; (e) cells of stipule with blue staining (magnified view); (f) mid-rib veins in the petiole; (g) mid-rib veins in the petiole, magnified view; (h) cell layers at the base of terminal trichomes of hydathodes of rosette/cauline leaves; (i) primary and secondary branch junctions; (j) peduncle of 3–4-day-old siliques; (k) outer cortex of hypocotyls of Col-0; (l) Col-0 roots; (m) outer cortex of hypocotyls of *pap1-D*; (n) *pap1-D* roots; (o) *pap1-D* stipules. Blue staining (indicated by black arrows) reveals PA monomers, oligomers or polymers. Arrows (in k, m and o) indicate lack of staining in control lines in areas that stain in plants co-expressing *TT2* and *PAP1*. Scale bars are 50 μm (panels c, e), 100 μ m (g, h) and 200 μm (panels d, k, m).

expressing *TT2* (lines 2, 6, 12, 24 and 28 in Figure 4a). Furthermore, staining was not found constitutively throughout the plant, in spite of the constitutive expression of *TT2* and *PAP1* in these lines. Rather, the pattern of staining (Figure 5) reflected the pattern of expression of the *BAN* promoter shown in Figure 2. Thus, blue DMACA staining was observed in the outer cortex of hypocotyls, in some lateral roots, in root hairs at the junction of the primary and secondary roots, in stipules at the base of rosette leaves, in primary and secondary branch junctions, in mid-rib veins in the petiole, in cell layers at the base of terminal trichomes of hydathodes on rosette/cauline leaves and in peduncles of 3–4-day-old siliques. Thus, specific cell types are pre-programmed for accumulation of PAs in Arabidopsis.

Expression of *PAP1* appeared to be highest in untransformed *pap1-D* or the empty vector line (2V) with the higher anthocyanin pigmentation (Figure 4c), suggesting that genomic incorporation of the additional 35S promoter sequence driving the *TT2* transgene might cause partial silencing of *PAP1* expression, itself driven by multiple 35S

enhancers. Consistent with this, a significant number of $T_1$ *pap1-D* plants transformed with pSB239 lost their purple leaf pigmentation. Out of 372 $T_1$ antibiotic-resistant plants, 327 had a wild-type visible phenotype, 26 (7%) had a purple mid-rib and petiole and only 19 (5%) had pigmented leaves. Out of 44 kanamycin-resistant ($Km^R$) $T_1$ *pap1-D* plants transformed with pCAMBIA2300, two (4.5%) had purple leaves and 13 (30%) showed purple pigmentation in the mid-rib and petiole. These results could either be due to silencing of the *PAP1* gene or to metabolism of the anthocyanidins produced. In support of the former possibility, different progenies of *TT2* expressing plants showed varying degrees of *PAP1* expression (see lines 12-1 to 12-4 in Figure 4a). Furthermore, *pap1-D* lines homozygous for 2 × *35S:TT2* showed little pigmentation on the mid-rib or petiole and had dramatically reduced *PAP1* expression compared with $T_1$ heterozygous plants. The homozygous plants also failed to show any DMACA staining. In addition to highlighting problems with gene silencing on incorporation of additional 35S promoter copies into *pap1-D* plants, these observations also indicate that high expression of both *PAP1* and *TT2* was essential for accumulation of PA in these lines.

To overcome the problem of transcriptional silencing associated with the introduction of additional 35S promoters into the *pap1-D* background, we constructed pSB439 (Figure 3), in which *TT2* and *PAP1* are separately driven by 2 × 35S Cauliflower mosaic virus (CaMV) and Cassava mosaic virus (CVMV) promoters, respectively. $T_1$ Col-0 plants transformed with pSB439 were screened for *TT2* and *PAP1* transcript levels by RT-PCR (Figure 4d). *PAP1* expression was much more consistent in these lines than in the *pap1-D* mutant transformed with 2 × *35S:TT2*. Plants 117-15, 117-17, 117-18 and 117-19 were analyzed by TLC for (–)-epicatechin and PA levels (Figure 6a) and by methanol extraction and spectrophotometry for anthocyanin levels (Figure 6c). The lines exhibited anthocyanin levels that were intermediate between those of *pap1-D* and Col-0 transformed with the empty vector (Figure 6c), and accumulated low levels of PA monomer units (epicatechin) and a small amount of polymeric DMACA-reactive material (but only in the two lines with the highest epicatechin levels, i.e. 117-17 and 117-18). $T_2$ plants (four to five plants from each line) which were segregating monogenically for glufosinate resistance were analyzed for their DMACA staining patterns; these were identical to those observed in Figure 5 for plants expressing *TT2* in the *pap1-D* background (data not shown). The overall low level of epicatechin and potential polymers is consistent with accumulation in a limited number of cell types.

To better understand the basis for the tissue specificity of accumulation of PA in relation to the expression patterns of *TT2* (constitutive under 35S promoter control) and *BAN* (induced by *TT2* expression), we bombarded leaves of wild-type Arabidopsis and Col-0 expressing *TT2* with gold particles harboring the $Pro_{BAN}$*:gusA* construct used for stable transformation in Figure 2. The results (Figure S1) indicated that *BAN* promoter expression was only observed in plants expressing *TT2*, and that it occurred generally throughout the young developing leaf rather than only in the specific cell types observed in Figure 5 to accumulate PAs following co-expression of *TT2* and *PAP1*. The frequency of blue spots on the young and expanding leaves of *TT2*-expressing plants bombarded with the $Pro_{BAN}$*:gusA* construct was similar to that observed following bombardment of wild-type leaves with *35S:gusA* (Figure S1). Interestingly, we did not find GUS expression in fully grown leaves of *TT2*-expressing plants bombarded with the $Pro_{BAN}$*:gusA* construct, consistent with the observation that *BAN* expression is limited to cells expressing *TTG1*, such as in young leaves but not in mature leaves (Baudry *et al.*, 2004). Overall, these results confirm that constitutive expression of *TT2* induces constitutive expression of *BAN*, and that the lack of PA accumulation throughout the leaf in plants expressing *TT2* and *PAP1* is not the result of restricted tissue specificity of *BAN* expression.

### *Constitutive expression of* Lc *induces anthocyanins but not PAs*

Several independent transgenic plants were generated with constitutive expression of *Lc* (Figure 7a). RT-PCR analysis (23 cycles) revealed low *AHA10* expression in these plants (Figure 7a). Expression of *BAN* and *TT12* was either very weak or not detectable after 30 cycles of RT-PCR, although *BAN* transcripts were observed after 40 cycles (data not shown). *Lc* transgenic plants did not show a constitutive purple leaf phenotype like *pap1-D*. However, leaves became purple about 6 weeks after sowing, when the plants started flowering (Figure 8a). Flowers, sepals, siliques and roots of *Lc* transgenics showed anthocyanin accumulation similar to that of *pap1-D* (Figure 8c–h), but no DMACA staining was observed, either constitutively or in the specific cell types that stain with DMACA in plants co-expressing *TT2* and *PAP1* (data not shown). Plants transformed with pSB451 for co-expression of *PAP1* and *Lc* showed anthocyanin accumulation in leaves and roots, but again there was no constitutive or tissue-specific DMACA staining (data not shown). Thus, although *Lc* induces anthocyanin synthesis it is insufficient to lead to accumulation of PA in Arabidopsis, even when co-expressed with *PAP1*.

### *Combined expression of* TT2, PAP1 *and* Lc

$T_1$ transgenic plants transformed with pSB458 for constitutive co-expression of *PAP1*, *Lc* and *TT2* exhibited growth abnormalities after germination on MS plates supplemented with glufosinate ammonium (6 mg $l^{-1}$). About 5% of the antibiotic-resistant plants showed yellowing, necrosis and death (Figure 8i–n). A frequency of 1.5% lethals has been

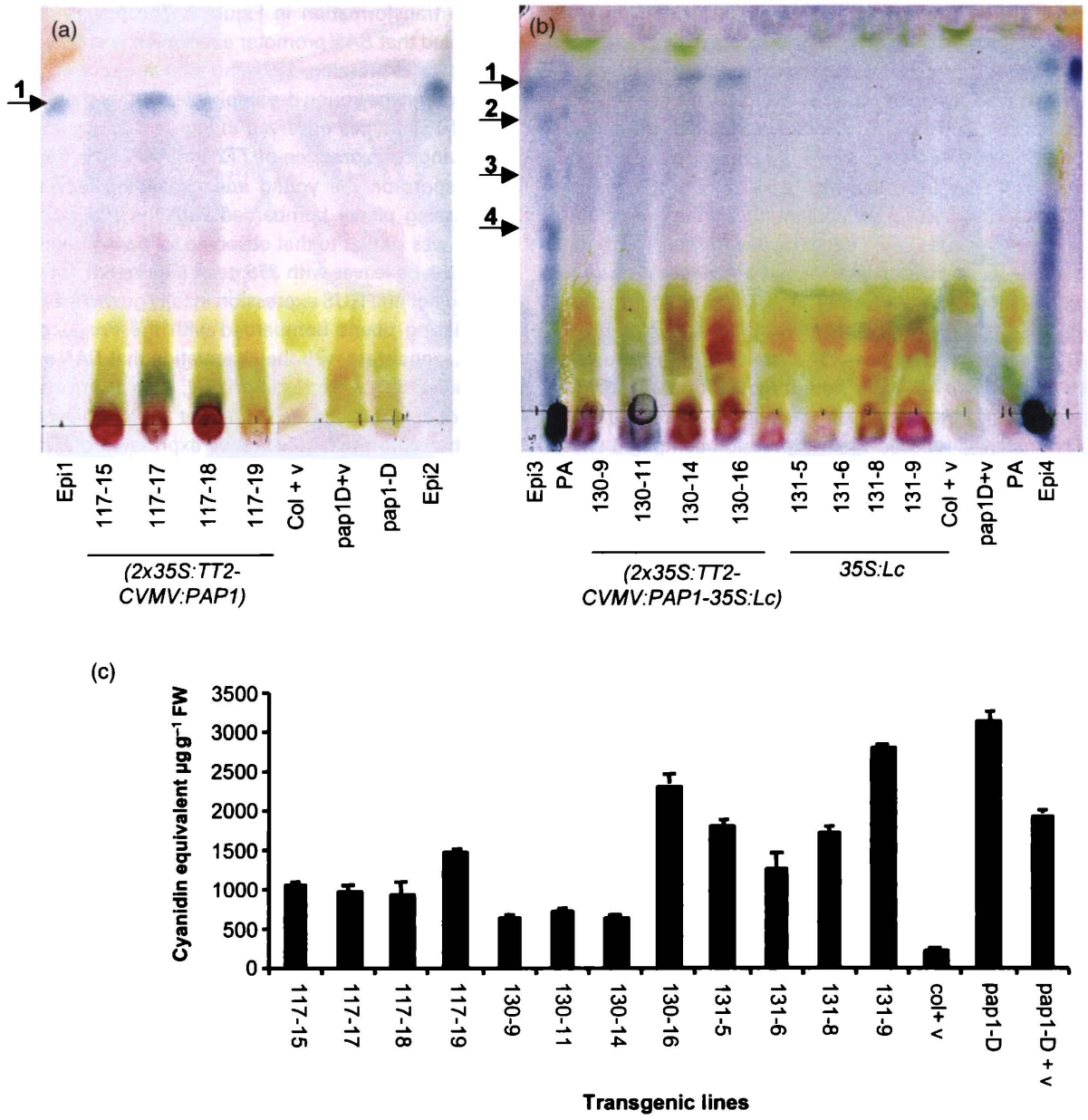


**Figure 6.** Anthocyanin, epicatechin and PA levels in Col-0 plants constitutively expressing combinations of the PAP1, TT2 and Lc transcription factors.
(a) Thin layer chromatogram showing epicatechin and PAs in leaf material (150 mg fresh weight) from *pap1-D*, *pap1-D* transformed with empty vector (pap1-D – v), Col-0 transformed with empty vector (Col + v), and four independent lines co-expressing *PAP1* and *TT2* (117-15 to 117-19) from construct pSB439. Epi1, 1 μg (–)-epicatechin standard; Epi2, 2 μg epicatechin.
(b) Thin layer chromatogram showing epicatechin and PAs in leaf material (200 mg fresh weight) from *pap1-D*, *pap1-D* transformed with empty vector (pap1-D + v), Col-0 transformed with empty vector (Col + v), four independent lines expressing Lc alone (131-5, -6, -8, -9) from construct pSB460, and four independent lines co-expressing *PAP1*, *TT2* and *Lc* (130-9, -11, -14, -16) from construct pSB458. PA, proanthocyanidins isolated from 5 mg leaf tissue of *Desmodium uncinatum*: Epi3, 0.5 μg epicatechin; Epi4, 2.5 μg epicatechin. Black arrows preceded by 1, 2, 3 or 4 indicate flavan-3-ol monomers (epicatechin), dimers, trimers and tetramers respectively.
(c) Anthocyanin levels in the above transgenic lines and their empty vector controls. Results are the average and standard deviation of three analyses.

reported previously following transformation in Arabidopsis; however, most of these mutations were recessive, and no dominant or semidominant lethals were found (Budziszewski *et al.*, 2001). Therefore, the dominant lethals observed among $T_1$ plants generated by transformation with pSB458 can be attributed to simultaneous expression of *TT2*, *PAP1* and *Lc*. When these necrotic plants were stained with DMACA, a significant number had blue-staining material in leaves and roots (Figure 8o). This staining was observed throughout the leaf, and was not localized to the specific cell

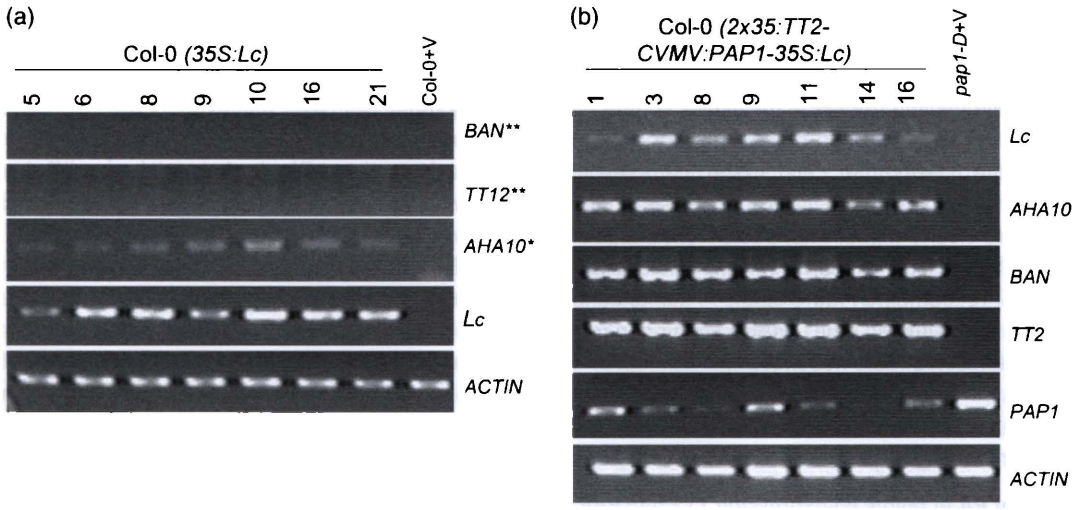


**Figure 7.** Transcript levels in leaves of Arabidopsis expressing *Lc* alone, or co-expressing *Lc*, *TT2* and *PAP1*, as determined by RT-PCR.
(a) *Lc*, *AHA10*, *TT12* and *BAN* transcript levels in $T_1$ Col-0 lines transformed with pSB460. *ACTIN* (internal control) and *Lc* were amplified for 21 cycles, *AHA10* was amplified for 23 cycles (*), and *BAN* and *TT12* were amplified for 30 cycles (**). *BAN* transcripts were visible after amplification for 40 cycles (data not shown). (b) *PAP1*, *TT2*, *Lc*, *BAN* and *AHA10* transcript levels in leaves of selected independent surviving transgenic plants following transformation with pSB458 (21 cycles of amplification).


**Figure 8.** Visible phenotypes of *Lc* and *Lc:TT2:PAP1* transgenic Arabidopsis.
(a, b) Visible phenotype of Col-0 expressing *35S:Lc* compared with empty vector control.
(c–h) DMACA staining of Col-0 expressing *35S:Lc* compared with *pap1-D* and empty vector controls.
(i–p) Phenotypes of Col-0 co-expressing *PAP1*, *TT2* and *Lc*. Visible phenotypes of transgenic plants are indicated by white arrows: normal (i, l); yellowing and necrotic (j, m); dying (k, n); DMACA-stained dying plant showing PAs in leaves (o); 3-month-old extreme dwarf plant (p). Bar = 5.0 mm.

types that accumulate PAs following co-expression of *TT2* and *PAP1*. Most of these plants died when transferred to soil. Some of the plants that survived were extremely dwarf with a purple phenotype (Figure 8p). Many of the transgenic plants which survived in soil were confirmed by RT-PCR analysis to co-express *TT2*, *PAP1* and *Lc* (Figure 7b). Four $T_1$ plants (nos 9, 11, 14, 16) showing high expression of *BAN* and *AHA10* (Figure 7b) were analyzed for accumulation of

anthocyanin and PA monomer in leaves (Figure 6b,c). Levels of anthocyanin in these plants were higher than in Col-0 but lower than in *pap1-D*. The plants accumulated epicatechin and flavan-3-ol dimers and trimers in the leaves (Figure 6b). The dimers and trimers, although present in very low concentrations (about 0.5–10 μg $g^{-1}$ leaf tissue; compare the levels in leaves from the positive control plant *Desmodium uncinatum*), were clearly higher in extracts from plants co-expressing *PAP1*, *TT2* and *Lc* than from plants expressing only *PAP1* and *TT2* (compare Figure 6a,b). $T_2$ plants generated from these lines showed blue DMACA staining in roots and stipules as expected (data not shown).

## Discussion

### *Accumulation of PA in plants co-expressing* TT2 *and* PAP1 *mirrors expression of* BAN *promoter in the wild-type background*

Previously, *BAN* expression in Arabidopsis has been reported as being primarily localized to the endothelial layer of the seed coat (Devic *et al.*, 1999), associated with the accumulation of PAs in this tissue. Our data reveal a somewhat broader expression pattern for the *BAN* promoter, to include a number of cell types in the roots and above-ground organs of the plant. In addition to the major expression in the seed endothelium, these regions included the mid-rib and hydathodes of rosette leaves, the petal veins, the peduncle, the cortex of the hypocotyl, some secondary roots, the puffs of root hairs at the junction of root and hypocotyls and the stipules at the base of rosette leaves. We could not, however, demonstrate staining for PAs in any of these cell types in wild-type Arabidopsis plants. It is possible that the levels of PAs were too low to be detected by DMACA staining, or that the main function of *BAN* in these tissues is the production of non-polymerized flavan-3-ol-derived materials rather than PAs. Previous studies of *BAN* expression using the GUS reporter in PA-accumulating cells of the seed endothelium have helped to define the *cis* elements in the *BAN* regulatory region (Debeaujon *et al.*, 2003). This work did not report expression of *BAN* outside the seed coat. Previous studies investigating *BAN* expression in different organs of Arabidopsis plants used RT-PCR analysis (Debeaujon *et al.*, 2003; Devic *et al.*, 1999). In these experiments, roots from 4-day-old seedlings or tissues from older flowering plants (siliques, flower, bud, stems and leaves) were used, where *BAN* expression in a limited number of cell types in a particular organ could have been diluted by the non-expressing predominant tissue.

Constitutive expression of *TT2* in aerial organs of Arabidopsis leads to *BAN* promoter activity (Baudry *et al.*, 2004), although only in young leaves which also express *TTG1*. It is believed that a ternary complex between TT2 and two other transcription factors, TT8 and TTG1, is necessary for correct expression of *BAN* in the Arabidopsis seed endothelium (Baudry *et al.*, 2004). The same might also be the case in young leaves, explaining the limited accumulation of PAs to only those cell types that express *BAN* (as visualized in our promoter–reporter study) and *TTG1*/*TT8*. TTG1 is a WD40 repeat protein which regulates trichome differentiation, root hair formation, production of seed mucilage and biosynthesis of PA and anthocyanin (Walker *et al.*, 1999). Expression of *BAN* near trichomes at hydathodes and in root hairs is consistent with co-expression with *TTG1*. However, although we did not perform in-depth analysis of the cell-specific expression of *BAN* within the aerial parts of plants expressing *TT2*, this appears to be constitutive based on transient expression of the $Pro_{BAN}$:*gusA* construct bombarded into young leaves of *TT2*-expressing, but not control, plants. The fact that we could only see PA staining in *TT2* transgenic lines also expressing *PAP1* suggests that anthocyanidin supply is limiting for PA formation in the various Arabidopsis cell types that naturally express *BAN*.

### *Bottlenecks for introducing PAs into Arabidopsis foliar tissue*

Leaves of plants co-expressing *TT2* and *PAP1*, with or without *Lc*, accumulated low levels of flavan-3-ols, epicatechin and PAs. The fact that PAs only appeared to accumulate in those cells in which the *BAN* promoter was naturally expressed indicates that specific cell types are programmed for synthesis and accumulation of PAs, and that co-expression of *CHS*, *DFR*, *LDOX*, *BAN*, *AHA10*, *TT12* and *TT19*, plus any other as yet known or unidentified genes that might be up-regulated by *TT2* and *PAP1*, is of itself insufficient to permit accumulation of PA throughout the plant.

It is perhaps surprising that so many of the flavonoid pathway genes were constitutively expressed in leaves of wild-type Arabidopsis grown under short-day conditions. Under short-day conditions the stems of wild-type plants had increased anthocyanin pigmentation, and the pap1-D phenotype was a deeper purple than under long-day conditions (data not shown). Detailed studies of the effect of short-day photoperiods on expression of genes involved in anthocyanin and PA biosynthesis have been lacking. However, *PAP1* and several genes of the phenylpropanoid pathway are under circadian control with daily changes in the level of gene expression of more than twofold (Harmer *et al.*, 2000). A constitutive background expression level for genes of the anthocyanin pathway in plants grown under short-day photoperiods is therefore not surprising.

Wild-type Col-0, and transgenic plants expressing *PAP1* and *Lc* alone or in combination, failed to accumulate flavan-3-ols, consistent with the lack of significant expression of *BAN* transcripts. Transgenic plants expressing PAP1 fused to the EAR-motif repression domain from SUPERMAN have suppressed *BAN* expression in immature siliques (Matsui

*et al.*, 2004), indicating that PAP1 can bind to the regulatory region of the *BAN* gene. However, the lack of significant accumulation of *BAN* transcript in transgenic Arabidopsis expressing *PAP1* and/or *Lc* indicates that, although these transcription factors may bind to the *BAN* promoter, additional factors are required to induce significant transcription.

We previously reported (Xie *et al.*, 2003) that ectopic expression of Arabidopsis *BAN* in Arabidopsis leaves resulted in accumulation of PAs as measured by the butanol-HCl method (Dalzell and Kerven, 1998). However, leaves of these lines did not show an overall blue coloration when stained with DMACA, suggesting that the material being detected by the butanol-HCl reagent was primarily flavan-3-ol monomers and possibly lower-order oligomers. Clearly, co-expression of additional factors regulated via TT2 is important to allow accumulation of higher polymers.

Procyanidin, (+)-catechin and anthocyanins are markedly increased by treatment of grape cell cultures with salicylic acid (SA) (Obinata *et al.*, 2003), and many genes involved in seed development and metabolism are regulated by plant hormones (McCourt, 1999) and *N*-acylethanolamines (NAEs) (Chapman *et al.*, 1999), which also induce phenylpropanoid biosynthesis in cell cultures (Tripathy *et al.*, 2003). However, we were unable to show activation of $Pro_{BAN}$:*gusA* expression in 2–3-week-old transgenic Arabidopsis plants following exposure to a range of concentrations of auxin (NAA), cytokinin (BAP), SA, jasmonic acid, abscisic acid, gibberellic acid or NAE14:0 (data not shown). Expression of additional transcription factors such as TT8, TTG1, TTG2, TT1 and TT16, previously shown to be involved in biosynthesis of PA in Arabidopsis seeds, may be necessary to overcome the bottleneck in inducing PA biosynthesis throughout foliar tissues.

In the absence of cell-specific metabolite analysis, we cannot conclude whether the low level of flavan-3-ol accumulation in leaf tissue of 2 × *35S:TT2-CVMV:PAP1* or 2 × *35S:TT2-CVMV:PAP1-35S:Lc* transgenic plants represents low constitutive accumulation in the leaves or relatively higher accumulation but only in specific cell types. Due to the constitutive nature of the *BAN* and *PAP1* expression, it is more likely to be the former. In this case, flavan-3-ol accumulation might be limited by rapid glycosylation of anthocyanidin substrates for the ANR (BAN) reaction, or by inefficient transport of flavan-3-ols to vacuoles with resulting degradation or modification in the cytoplasmic environment.

Combined high-level expression of PAP1, Lc and TT2 leads to necrosis and death in a high percentage of plants, associated with accumulation of PAs in leaves and roots. It is possible that co-expression of the three transcription factors specifically switches on 'missing components' of the PA biosynthesis machinery throughout the leaf. However, accumulation of PA in this case could also be a response to cell death initiated by high co-expression of the three transcription factors resulting in biosynthesis of cytotoxic compounds, or general cellular dysregulation.

### *Prospects for the genetic manipulation of PA biosynthesis*

Several attempts have been made to engineer PAs in forage crops with the goal of reducing pasture bloat, the major targets for this trait being alfalfa and clovers. Whereas efforts targeting biosynthetic pathway genes have either been unsuccessful or have given equivocal results (Dixon *et al.*, 2005), two studies using transcription factors have suggested the feasibility of the approach. In alfalfa, although neither the maize *B-Peru* nor *C1* anthocyanin regulatory genes was effective in modulating the pathway, accumulation of PAs, as measured by reactivity with butanol-HCl reagent, was recorded in leaf tissue expressing the maize *Lc* gene (Ray *et al.*, 2003). The *Lc* transgenic alfalfa appeared to require a stress, such as exposure to high light intensity or cold, for appearance of anthocyanin and PA (Ray *et al.*, 2003). In birdsfoot trefoil (*Lotus corniculatus*), which naturally accumulates high levels of PAs in leaf tissues, constitutive expression of the maize Sn transcription factor resulted in increased accumulation of PA in roots, but levels in leaves were unaffected or suppressed (Damiani *et al.*, 1999).

Successful engineering of PAs in forage crops will require constitutive or developmentally controlled as opposed to stress-induced anthocyanin production coupled to co-expression of all the genes necessary for biosynthesis, transport, polymerization and storage of PA monomers. Given the availability of an extensive collection of mutants blocked in the various steps of the PA pathway, a whole genome sequence and facile transformation system, Arabidopsis provides an excellent system for defining bottlenecks in metabolic engineering of PAs and then systematically overcoming them by design strategies for introducing PAs into the leaves of forage or feed crops.

The need to introduce both upstream (the anthocyanin pathway) and downstream (the PA pathway) genes into leaves of forage crops for PA engineering is potentially problematical, requiring either transformation with multiple genes or a risk of the kinds of pleiotropic effects observed in the present work following expression of multiple transcription factors. Fortunately, mutants and breeding lines of white clover and *Medicago* are available in which significant anthocyanin accumulation occurs naturally in the leaves (Diwan *et al.*, 1994; Pederson, 1995). For example, the 'red spot' is a phenotypic characteristic of the leaves of the model legume *Medicago truncatula*, making this a valuable 'intermediate' species for moving technology from Arabidopsis to forage legumes.

## Experimental procedures

### *Plant materials and growth conditions*

*Arabidopsis thaliana* accessions Columbia (Col-0) and its T-DNA activation tagged mutant *pap1-D* (Borevitz *et al.*, 2000) were used. The *tt2* mutant CS83 was obtained from the ABRC (Columbus, OH,

USA). Plants were grown in soil (Metromix 200; Scotts, Marysville, OH, USA) at 22 to 25°C under either 16 h of light and 8 h of dark (long days) or 10 h of light and 14 h dark (short days). Plants grown aseptically were plated on MS medium with 1.5% (w/v) sucrose solidified with 0.6% (w/v) phytagar, either alone or supplemented with glufosinate ammonium (6 mg $l^{-1}$) (Sigma-Aldrich, St Louis, MO, USA) or kanamycin (50 mg $l^{-1}$). Plates were wrapped with gas-permeable 3 M Micropore surgical tape (3 M Health Care, St. Paul, MN, USA) and grown at 22°C under long days.

### *Plasmid constructs and plant transformation*

Standard recombinant DNA techniques were used for gene cloning (Sambrook *et al.*, 1989). The Arabidopsis *TT2* cDNA was isolated by reverse transcriptase polymerase chain reaction (RT-PCR). Total RNA was isolated from the first three to four newly emerged young siliques using TRI-REAGENT (Molecular Research Center Inc., Cincinnati, OH, USA) according to the manufacturer's instructions. Four micrograms total RNA was reverse transcribed to synthesize first-strand cDNA in a total volume of 20 μl containing 20 mM Tris-HCl pH 8.4, 50 mM KCl, 5 mM $MgCl_2$, 10 mM DTT, 1 mM deoxyribonucleoside triphosphate mixture, 500 ng oligodinucleotides [oligo(dT)] 12–18, 25 units of RNaseOUT (ribonuclease inhibitor; Invitrogen, Carlsbad, CA, USA) and 200 units of Moloney murine leukemia virus reverse transcriptase (SuperScriptII RNAase $H^-$ Reverse Transcriptase Kit; Invitrogen) at 42°C for 1 h. Ten microliters of first-strand cDNA was amplified by PCR using high-fidelity DNA polymerase (PfuTurbo DNA polymerase; Stratagene, La Jolla, CA, USA) and *TT2* gene-specific primers (Table S1) in 100 μl reaction mix according to the manufacturer's protocol. The PCR products were de-salted, digested with *Nco*I and *Xba*I, gel purified (gel purification kit; Qiagen Inc., Valencia, CA, USA) and cloned into *Nco*I- and *Xba*I-digested pRTL2 vector (Restrepo *et al.*, 1990). Recombinant pRTL2 plasmids containing the *TT2* insert were sequenced to verify the *TT2* coding region and insert junctions, and one such plasmid with the correct *TT2* sequence was named pSB207. The *Pst*I fragment of pSB207 carrying the coding region of *TT2* fused to the double CaMV 35S promoter and the CaMV 35S polyadenylation signal was cloned into pCAMBIA2300 (AF234315) to generate pSB239 (Figure 3).

The 669 bp intergenic region upstream from the translation initiation site of the Arabidopsis *BAN* gene (Figure 1) was isolated by PCR using *BAN* promoter-specific ($Pro_{BAN}$) primers (Table S1). PCR products were de-salted, digested with *Hin*dIII and *Nco*I, and the fragments gel purified and directly cloned into pCAMBIA1301 (AF234297) to generate the $Pro_{BAN}$:*gusA* fusion construct pSB159. The *Bam*HI–*Nco*I fragment of pSB159 was cloned into pBluesGFPS65Tsk (Niwa *et al.*, 1999) to generate the $Pro_{BAN}$:*sGFP* construct, which was digested with *Bam*HI–*Sal*I and cloned into the binary vector pCAMBIA2300.

Plasmid pSB460, which contains the 2.2 kb maize *Lc* gene under control of the 35S promoter and NOS terminator for constitutive expression (Ludwig *et al.*, 1990), was constructed by cloning the 3.1 kb *Eco*RI–*Hin*dIII fragment of pSRLc349 (kindly provided by Professor S. Wessler, University of Georgia) into *Eco*R1–*Hin*dIII-cut pCAMBIA2300. The plasmid pSB419 which carries the Arabidopsis *PAP1* gene under the control of the Cassava Mosaic virus promoter ($Pro_{CVMV}$) was constructed by ligating *Hin*dIII–*Bst*EII-cut pCAMBIA3301 with a purified 528 bp *Hin*dIII–*Nco*I fragment carrying $Pro_{CVMV}$ from pILTAB380 (provided by Claude Fauquet, ILTAB, Donald Danforth Plant Science Center, St Louis, MO, USA) and the *Nco*I–*Bst*EII fragment carrying the *PAP1* coding region from plasmid pPAP-TOPO (provided by Z. Wang, Noble Foundation). Plasmid pSB439, for co-expression of *35S*:*TT2* and *PAP1* under control of the Cassava mosaic virus promoter (*CVMV*:*PAP1*), was generated by inserting the *Pst*I fragment of pSB207 carrying the coding region of the *TT2* gene fused to the double 35S promoter and the 35S polyadenylation signal of CaMV into the *Pst*I site of pSB419. For constitutive co-expression of *PAP1* and *Lc*, the 3.1 kb *Eco*RI–*Hin*dIII fragment of pSRLc349 was cloned into *Eco*R1–*Hin*dIII-cut pSB419 to generate pSB451. Plasmid pSB458, which carries *TT2*, *Lc* and *PAP1* for constitutive co-expression, was derived from pSB451 by insertion of a *Hin*dIII fragment carrying the coding region of *TT2* fused to the double 35S promoter and the CaMV 35S polyadenylation signal from pSB207. Plasmids pSB239, pSB439, pSB451, pSB458 and pSB460 (Figure 3) were transformed into *Agrobacterium tumefaciens* strain GV3101 (Koncz and Schell, 1986) by electroporation.

Arabidopsis was transformed using the floral dip method following seed sterilization by the liquid or vapor phase methods (Clough and Bent, 1998). Transgenic Arabidopsis lines in the Col-0 background resulting from transformation with pSB451 or pSB458 were selected on MS media with glufosinate (6 mg $l^{-1}$), and those transformed with pSB239 or pSB460 on kanamycin (50 mg $l^{-1}$). *pap1-D* plants transformed with pSB239 were selected on kanamycin (50 mg $l^{-1}$).

Histochemical staining of *gusA* transgenic plants was done as described previously (Jefferson *et al.*, 1987; Stomp, 1992). GFP fluorescence in transgenic Arabidopsis plants was monitored by confocal microscopy.

### *RT-PCR analysis*

Total RNA was isolated from the rosette leaves of 4–5-week-old plants using TRI-REAGENT. Two micrograms of total RNA was used to synthesize first-strand cDNA using Ready-To-Go RT-PCR beads (Amersham Biosciences, Piscataway, NJ, USA) in a total volume of 50 μl according to the manufacturer's instructions. Five microliters of this reaction (equivalent to first-strand cDNA from 200 ng total RNA) was amplified using Taq Polymerase (Ex Taq, TaKaRa, Shiga, Japan or GoTaq, Promega, Madison, WI, USA) and gene-specific primers (Table S1) in a total volume of 35 μl according to the manufacturer's protocols. The cycle conditions were 95°C for 7 min; 21, 23 or 30 cycles at 95°C for 1 min, 55°C for 1 min, 72°C for 2 min, followed by a final extension at 72°C for 5 min.

### *Anthocyanin, epicatechin and PA analysis*

The presence of PAs was monitored by immersing tissues in dimethylaminocinnamaldehyde (DMACA) solution (0.1% DMACA in 6N HCl:95% ethanol, 1:1) (Li *et al.*, 1996). After staining for 5–10 min, plant or tissue samples were washed three times with distilled water and histochemical staining (blue color) was observed under the microscope.

For further analysis of PAs in leaf tissues, leaves were ground in liquid $N_2$ and mixed with 10 volumes of 70% aqueous acetone containing 5.26 mM sodium metabisulphite. The sample was sonicated for 20 min at 20°C, centrifuged at 2630 ***g*** for 10 min, and the supernatant collected. The extraction was repeated three times. Acetone from one-third of the supernatant was dried under nitrogen gas and further extracted with ethyl acetate to partition out the monomers and small oligomers, leaving PA polymers in the aqueous phase. Both fractions were then extracted with hexane (three times) and finally with chloroform. Fractions were dried, dissolved in methanol, and 10 μl samples were spotted onto silica gel 60 F254 aluminum-backed TLC sheets (EMD Chemical Inc. Gibbstown, NJ, USA) that were then developed in ethyl acetate:

formic acid:water [20:1:1 (v/v)]. Dried plates were sprayed with 0.1% DMACA solution.

For analysis of anthocyanin levels, leaf tissues were ground in liquid $N_2$, and 100 mg ground tissue was extracted four times with 1 ml methanol:HCl [99.5:0.5 (v/v)] by shaking at room temperature in the dark. The first three extractions were done for 3–4 h each followed by centrifugation at 2630 ***g*** to collect the supernatant. The fourth extraction was done overnight. Supernatants were pooled and total anthocyanins were quantified by spectrophotometry at 530 nm after subtracting the optical density readings at 600 nm (Abrahams *et al.*, 2002).

## Acknowledgements

This work was supported by Forage Genetics International and the Samuel Roberts Noble Foundation. We thank Drs Stephen Temple and Lahoucine Achnine for critical reading of the manuscript.

## Supplementary Material

The following supplementary material is available for this article online:

**Figure S1.** Transient expression of *Pro$_{BAN}$:gusA* in leaves of 2 × *35S:TT2*-expressing or wild-type Arabidopsis.

**Table S1** Gene-specific primers used in the present work

## References

**Abrahams, S., Tanner, G.J., Larkin, P.J. and Ashton, A.R.** (2002) Identification and biochemical characterization of mutants in the proanthocyanidin pathway in *Arabidopsis*. *Plant Physiol.* **130**, 561–576.

**Aerts, R.J., Barry, T.N. and McNabb, W.C.** (1999) Polyphenols and agriculture: beneficial effects of proanthocyanidins in forages. *Agric. Ecosyst. Environ.* **75**, 1–12.

**Barry, T.N. and McNabb, W.C.** (1999) The implications of condensed tannins on the nutritive value of temperate forages fed to ruminants. *Br. J. Nutr.* **81**, 263–272.

**Baudry, G., Heim, M.A., Debreucq, B., Caboche, M., Weisshaar, B. and Lepiniec, L.** (2004) TT2, TT8 and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* **39**, 366–380.

**Baxter, I.R., Young, J.C., Armstrong, G., Foster, N., Bogenschutz, N., Cordova, T., Peer, W.A., Hazen, S.P., Murphy, A.S. and Harper, J.F.** (2005) A plasma membrane H−-ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA*, **102**, 2649–2654.

**Borevitz, J., Xia, Y., Blount, J.W., Dixon, R.A. and Lamb, C.** (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell*, **12**, 2383–2393.

**Bovy, A.G., de Vos, R., Kemper, M. *et al.*** (2002) High-flavonol tomatoes resulting form the heterologous expression of the maize transcription factor genes *LC* and *C1*. *Plant Cell*, **14**, 2509–2526.

**Bradley, J.M., Davies, K.M., Deroles, S.C., Bloor, S.J. and Lewis, D.H.** (1998) The maize *Lc* regulatory gene up-regulates the flavonoid biosynthetic pathway of *Petunia*. *Plant J.* **13**, 381–392.

**Budziszewski, G.J., Potter, L.S., Glover, L.W. *et al.*** (2001) Arabidopsis genes essential for seedling viability: isolation of insertional mutants and molecular cloning. *Genetics*, **159**, 1765–1778.

**Chapman, K.D., Venables, B., Markovic, R., Blair, R.W., Jr and Bettinger, C.** (1999) N-Acylethanolamines in seeds. Quantification of molecular species and their degradation upon imbibition. *Plant Physiol.* **120**, 1157–1164.

**Clough, S.J. and Bent, A.F.** (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

**Dalzell, S.A. and Kerven, G.L.** (1998) A rapid method for the measurement of *Leucaena* spp proanthocyanidins by the proanthocyanidin (butanol/HCl) assay. *J. Sci. Food Agric.* **78**, 405–416.

**Damiani, F., Paolocci, F., Cluster, P.D., Arcioni, S., Tanner, G.J., Joseph, R.G., Li, Y.G., de Majnik, J. and Larkin, P.J.** (1999) The maize transcription factor Sn alters proanthocyanidin synthesis in transgenic *Lotus corniculatus* plants. *Aust. J. Plant Physiol.* **26**, 159–169.

**Debeaujon, I., Peeters, A.J.M., Leon-Kloosterziel, K.M. and Korneef, M.** (2001) The *TRANSPARENT TESTA 12* gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *Plant Cell*, **13**, 853–871.

**Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M. and Lepiniec, L.** (2003) Proanthocyanidin-accumulating cells in Arabidopsis testa: regulation of differentiation and role in seed development. *Plant Cell*, **15**, 2514–2531.

**Devic, M., Guilleminot, J., Debeaujon, I., Bechtold, N., Bensaude, E., Koornneef, M., Pelletier, G. and Delseny, M.** (1999) The *BANYULS* gene encodes a DFR-like protein and is a marker of early seed coat development. *Plant J.* **19**, 387–398.

**Diwan, N., Bauchan, G.R. and McIntosh, M.S.** (1994) A core collection for the United States annual *Medicago* germplasm collection. *Crop Sci.* **34**, 279–285.

**Dixon, R.A., Xie, D. and Sharma, S.B.**, (2005) Proanthocyanidins – a final frontier in flavonoid research? *New Phytol.* **165**, 9–28.

**Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A. and Kay, S.A.** (2000) Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science*, **290**, 2110–2113.

**Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W.** (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.

**Johnson, C.S., Kolevski, B. and Smyth, D.R.** (2002) TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor. *Plant Cell*, **14**, 1359–1375.

**Kitamura, S., Shikazono, N. and Tanaka, A.** (2004) TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. *Plant J.* **37**, 104–114.

**Koncz, C. and Schell, J.** (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* **204**, 383–396.

**Li, Y.G., Tanner, G. and Larkin, P.** (1996) The DMACA-HCl protocol and the threshold proanthocyanidin content for bloat safety in forage legumes. *J. Sci. Food Agric.* **70**, 89–101.

**Ludwig, S.R., Bowen, B., Beach, L. and Wessler, S.R.** (1990) A regulatory gene as a novel visible marker for maize transformation. *Science*, **247**, 449–450.

**deMajnik, J., Weinman, J.J., Djordjevic, M.A., Rolfe, B.G., Tanner, G.J., Joseph, R.G. and Larkin, P.J.** (2000) Anthocyanin regulatory gene expression in transgenic white clover can result in an altered pattern of pigmentation. *Aust. J. Plant Physiol.* **27**, 659–667.

**Marles, M.A.S., Ray, H. and Gruber, M.Y.** (2003) New perspectives on proanthocyanidin biochemistry and molecular regulation. *Phytochemistry*, **64**, 357–383.

**Matsui, K., Tanaka, H. and Ohme-Takagi, M.** (2004) Suppression of the biosynthesis of proanthocyanidin in Arabidopsis by a chimeric PAP1 repressor. *Plant Biotechnol. J.* **2**, 487–493.

**McCourt, P.** (1999) Genetic analysis of hormone signaling. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 219–243.

**McMahon, L.R., McAllister, T.A., Berg, B.P., Majak, W., Acharya, S.N., Popp, J.D., Coulman, B.E., Wang, Y. and Cheng, K.J.** (2000) A review of the effects of forage condensed tannins on ruminal fermentation and bloat in grazing cattle. *Can. J. Plant Sci.* **80**, 469–485.

**Nesi, N., Debeaujon, I., Jond, C., Pelletier, G., Caboche, M. and Lepiniec, L.** (2000) The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in *Arabidopsis* siliques. *Plant Cell*, **12**, 1863–1878.

**Nesi, N., Jond, C., Debeaujon, I., Caboche, M. and Lepiniec, L.** (2001) The Arabidopsis *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell*, **13**, 2099–2114.

**Nesi, N., Debeaujon, I., Jond, C., Stewart, A.J., Jenkins, G.I., Caboche, M. and Lepiniec, L.** (2002) The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *Plant Cell*, **14**, 2463–2479.

**Niwa, Y., Hirano, T., Yoshimoto, K., Shimizu, M. and Kobayashi, H.** (1999) Non-invasive quantitative detection and applications of non-toxic, S65T-type green fluorescent protein in living plants. *Plant J.* **18**, 455–463.

**Obinata, N., Yamakawa, T., Takamiya, M., Tanaka, N., Ishimaru, K. and Kodama, T.** (2003) Effect of salicylic acid on the production of procyanidin and anthocyanin in cultured grape cells. *Plant Biotechnol.* **20**, 105–111.

**Pederson, G.A.** (1995) Registration of MSRLM red leaf mark white clover germplasm. *Crop Sci.* **35**, 1235–1236.

**Ray, H., Yu, M., Auser, P., Blahut-Beatty, L., McKersie, B., Bowley, S., Westcott, N., Coulman, B., Lloyd, A. and Gruber, M.Y.** (2003) Expression of anthocyanins and proanthocyanidins after transformation of alfalfa with maize *Lc*. *Plant Physiol.* **132**, 1448–1463.

**Restrepo, M.A., Freed, D.D. and Carrington, J.C.** (1990) Nuclear transport of plant potyviral proteins. *Plant Cell*, **2**, 987–998.

**Robbins, M.P., Paolocci, F., Hughes, J.-W., Turchetti, V., Allison, G., Arcioni, S., Morris, P. and Damiani, F.** (2003) *Sn*, a maize bHLH gene, modulates anthocyanin and condensed tannin pathways in *Lotus corniculatus*. *J. Exp. Bot.* **54**, 239–248.

**Sagasser, M., Lu, G.-H., Hahlbrock, K. and Weisshaar, B.** (2002) *A. thaliana* TRANSPARENT TESTA 1 is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. *Genes Dev.* **16**, 138–149.

**Sambrook, J., Fritsch, E.F. and Maniatis, T.** (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

**Shirley, B.W., Kubasek, W.L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F.M. and Goodman, H.M.** (1995) Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *Plant J.* **8**, 659–671.

**Stomp, A.-M.** (1992) Histochemical localization of β-glucuronidase. In *GUS Protocols: Using the GUS Gene as a Reporter of Gene Expression* (Gallagher, S., ed.). New York: Academic Press, pp. 103–113.

**Tohge, T., Nishiyama, Y., Hirai, M.Y. *et al.*** (2005) Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. *Plant J.* **42**, 218–235.

**Tripathy, S., Kleppinger-Sparace, K., Dixon, R.A. and Chapman, K.D.** (2003) N-acylethanolamine signaling in tobacco is mediated by a membrane-associated, high-affinity binding protein. *Plant Physiol.* **131**, 1781–1791.

**Walker, A.R., Davison, P.A., Bolognesi-Winfield, A.C., James, C.M., Srinivasdan, N., Blundell, T.L., Esch, J.J., Marks, M.D. and Gray, J.C.** (1999) The *TRANSPARENT TESTA GLABRA 1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell*, **11**, 1337–1349.

**Winkel-Shirley, B.** (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* **126**, 485–493.

**Xie, D., Sharma, S.R., Paiva, N.L., Ferreira, D. and Dixon, R.A.** (2003) Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. *Science*, **299**, 396–399.