Expression of the R2R3-MYB Transcription Factor TaMYB14 from *Trifolium arvense* Activates Proanthocyanidin Biosynthesis in the Legumes *Trifolium repens* and *Medicago sativa*[1][W][OA]

Kerry R. Hancock, Vern Collette, Karl Fraser, Margaret Greig, Hong Xue, Kim Richardson, Chris Jones, and Susanne Rasmussen*

AgResearch, Ltd., Palmerston North 4442, New Zealand

Proanthocyanidins (PAs) are oligomeric flavonoids and one group of end products of the phenylpropanoid pathway. PAs have been reported to be beneficial for human and animal health and are particularly important in pastoral agricultural systems for improved animal production and reduced greenhouse gas emissions. However, the main forage legumes grown in these systems, such as *Trifolium repens* and *Medicago sativa*, do not contain any substantial amounts of PAs in leaves. We have identified from the foliar PA-accumulating legume *Trifolium arvense* an R2R3-MYB transcription factor, TaMYB14, and provide evidence that this transcription factor is involved in the regulation of PA biosynthesis in legumes. *TaMYB14* expression is necessary and sufficient to up-regulate late steps of the phenylpropanoid pathway and to induce PA biosynthesis. RNA interference silencing of *TaMYB14* resulted in almost complete cessation of PA biosynthesis in *T. arvense*, whereas *Nicotiana tabacum*, *M. sativa*, and *T. repens* plants constitutively expressing *TaMYB14* synthesized and accumulated PAs in leaves up to 1.8% dry matter. Targeted liquid chromatography-multistage tandem mass spectrometry analysis identified foliar PAs up to degree of polymerization 6 in leaf extracts. Hence, genetically modified *M. sativa* and *T. repens* plants expressing *TaMYB14* provide a viable option for improving animal health and mitigating the negative environmental impacts of pastoral animal production systems.

Proanthocyanidins (PAs), also known as condensed tannins, are polyphenolic plant secondary metabolites that accumulate in a wide range of plant tissues (Dixon et al., 2005). Their function in plants is mainly associated with the ability to bind proteins and metal ions, thereby providing an effective defense against pathogens and herbivores (Gonzalez de Colmenares et al., 1998; Aziz et al., 2005). The beneficial effects of PAs found in green tea (*Camellia sinensis*), grape seed (*Vitis vinifera*), or cranberry (*Vaccinium oxycoccos*) on human health have been widely reported and have resulted in considerable interest in the biosynthesis and chemistry of this compound class (Ahmad et al., 2000; Bagchi et al., 2000; Dufresne and Farnworth, 2001).

PAs also play a crucial role in forage-based animal production systems with high forage legume content. The effects of PAs on grazing mammals can be detrimental when concentrations are very high (above 6%), as found in leaves of some *Lotus* species, leading to reduced forage palatability and digestibility (Aerts et al., 1999). On the other hand, the predominant pasture legume species *Medicago sativa* (lucerne, alfalfa) and *Trifolium repens* (white clover), which contain no or very low levels of PAs in foliar tissues (Marshall et al., 1980; Li et al., 1996), can also have negative effects on animal production. These legumes are very rich in proteins, which are rapidly fermented in the rumen, resulting in gas and foam formation, the cause of a serious condition called pasture bloat that can be lethal (McMahon et al., 2000). In the presence of PAs, excess dietary proteins as well as bacterial enzymes are complexed, and the level of protein degradation in the rumen is significantly reduced. This leads to an increase of protein bypass to the ruminant's gut and improved absorption of essential amino acids, resulting in increased milk and meat production (Aerts et al., 1999; Douglas et al., 1999; McMahon et al., 2000). Moreover, decreased protein degradation in the rumen also decreases methane production and ammonium excretion in urine, which could contribute to significant reductions in emissions of the potent greenhouse gases methane and nitrous oxide from pastures (Woodward et al., 2001; Kingston-Smith and Thomas, 2003; Beauchemin et al., 2007; Smith et al., 2008). Legumes containing moderate amounts of PAs have also been associated with reduced intestinal parasite load

[1] This work was supported by Grasslanz Technologies, Ltd.

* Corresponding author; e-mail susanne.rasmussen@agresearch.co.nz.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Susanne Rasmussen (susanne.rasmussen@agresearch.co.nz).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.112.195420

in ruminants (Aerts et al., 1999). The production of *M. sativa* and *T. repens* cultivars with moderate amounts of foliar PAs (2%–4%), therefore, is of considerable interest to the pastoral agricultural industry (Reed, 1995; Aerts et al., 1999; Smith et al., 2008).

PAs are derived from the phenylpropanoid/flavonoid pathway (for review, see Holton and Cornish, 1995; Winkel-Shirley, 2001; Dixon et al., 2005), sharing most precursors with flavonols and anthocyanins. Molecular genetic studies of Arabidopsis (*Arabidopsis thaliana*) *transparent testa* (*tt*) and *tannin-deficient seed* mutants have allowed the identification of a number of structural and regulatory genes (Shirley et al., 1995; Winkel-Shirley, 2001; Abrahams et al., 2003; Lepiniec et al., 2006; Fig. 1). Briefly, starting from the common precursor *p*-coumaric acid, the chalcone naringenin is formed by chalcone synthase (CHS). Naringenin is subsequently hydroxylated by flavanone 3-hydroxylase (F3H) to dihydrokaempferol, which can be hydroxylated at the B-ring by F3′H and F3′5′H to dihydroquercetin and dihydromyricetin. These dihydroflavonols are either converted by flavonol synthase (FLS) to the flavonols kaempferol, quercetin, and myricetin or reduced by dihydroflavonol 4-reductase (DFR) to the corresponding leucoanthocyanidins (flavan-3,4-diols). A further reduction by leucoanthocyanidin reductase (LAR) yields the monomeric PA units, transflavan-3-ols (afzelechin, catechin, and gallocatechin). This step seems not to be involved in PA biosynthesis in Arabidopsis, as its PAs do not contain any transflavan-3-ols but appear to be made up exclusively of the cis-flavan-3-ol epicatechin. LAR was first isolated from the legume *Desmodium uncinatum*, which accumulates PAs based on both the trans- and cis-flavan-3-ols (Tanner et al., 2003). Leucoanthocyanidins can also be converted by anthocyanidin synthase (ANS; syn. leucoanthocyanidin deoxygenase) to anthocyanidins (pelargonidin, cyanidin, and delphinidin), which are either glycosylated to the stable, colored anthocyanins or reduced by anthocyanidin reductase (ANR) to form the second set of monomeric PA units, cis-flavan-3-ols (epiafzelechin, epicatechin, and epigallocatechin). Most of these genes have also been identified in legumes (for review, see Rasmussen, 2009). Many questions about the transport into vacuoles and polymerization of these monomeric PA units remain open (Zhao et al., 2010), but recently, a UDP-Glc transferase that glycosylates epicatechin to the corresponding glucoside was identified in the model legume *Medicago truncatula* (Pang et al., 2008). Epicatechin 3-*O*-glucoside acts as a substrate for the vacuolar multidrug and toxic compound extrusion (MATE) transporters, which also transport glycosylated anthocyanidins (Marinova et al., 2007; Zhao and Dixon, 2009).

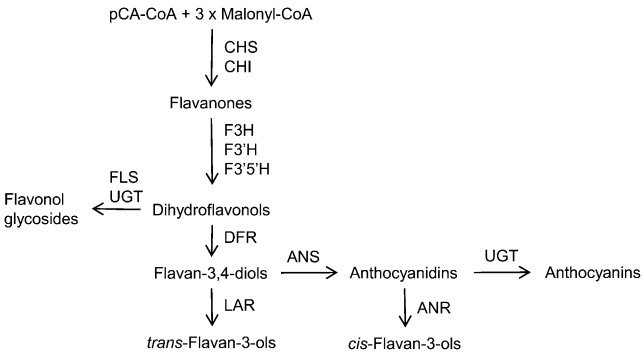

Figure 1. Schematic presentation of the flavonoid pathway. Enzymes are as follows: CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid 3′,5′-hydroxylase; FLS, flavonol synthase; UGT, UDP-glycosyltransferase; DFR, dihydroflavonol reductase; LAR, leucoanthocyanidin reductase; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase.

Temporal and spatial regulation of these various parts of the flavonoid pathway is complex, and a large number of transcription factors (TFs) that regulate the expression and tissue specificity of anthocyanin and PA synthesis were identified in Arabidopsis and a number of other plants. A ternary TF complex of TT2-TT8-TTG1 involving an R2R3-MYB, a basic helix-loop-helix (bHLH), and a WD40 repeat protein regulates the expression of *AtTT3* (DFR), *AtTT18* (ANS), *AtBAN* (ANR), and *AtTT12* (MATE) and is essential for PA biosynthesis and accumulation in Arabidopsis seed coats (Nesi et al., 2001; Baudry et al., 2004; Lepiniec et al., 2006). Several R2R3-MYB TFs have been isolated recently from grapevine (*VvMYBPA1*, *VvMYBPA2*, and *VvMYB5*) and shown to be essential for the expression of PA biosynthetic genes and the accumulation of PAs in grapevine (Bogs et al., 2007; Deluc et al., 2008; Terrier et al., 2009). Other species in which MYB TFs have been reported to be crucial for PA biosynthesis include *Brassica napus*, *Diospyrios kaki* (persimmon), and *Populus tremuloides* (Wei et al., 2007; Akagi et al., 2009; Mellway et al., 2009), indicating that MYB TFs are common regulators of PA biosynthesis. The first legume R2R3-MYB TFs reported to regulate PA biosynthetic genes are three *Lotus japonicus* TT2 orthologs, which, if coexpressed with *AtTT8*, *AtTTG1*, and *LjANR* in transient Arabidopsis leaf cell assays, activate the promoter of *LjANR* (Yoshida et al., 2008). Constitutive expression of *LjTT2a* results in the production and accumulation of PAs and anthocyanins in Arabidopsis seedlings and induces the expression of CHS, DFR, ANS, ANR, and a glutathione *S*-transferase predicted to be involved in PA monomer transport (Yoshida et al., 2010).

In common with the forage legume *M. sativa*, *T. repens* does not accumulate significant levels of PAs in foliar tissues; relatively high levels of PAs, however, are found in *M. sativa* and *T. repens* seed coats and clover flower petals (Goplen et al., 1980; Koupai-Abyazani et al., 1993; Foo et al., 2000). Despite the importance of foliar PAs in these forage species, no convincing strategies of engineering sufficient amounts of PAs into the leaves of these species have been reported. Attempts to induce PA biosynthesis into *M.*

sativa leaves using ectopic coexpression of *PAP1*, a MYB TF responsible for regulating most of the steps in the phenylpropanoid pathway except the biosynthesis of PAs from Arabidopsis (Borevitz et al., 2000) and *MtANR* (Xie et al., 2006), or ectopic expression of *LcMYC*, a bHLH TF from *Zea mays* (Ray et al., 2003), have had only limited success so far. It has been suggested that the activation potential of flavonoid-regulating TFs might be species specific, as induction of anthocyanin biosynthesis was not seen in *M. sativa* when overexpressing *PAP1* from Arabidopsis but required the expression of Legume Anthocyanin Production1 (*LAP1*), a *PAP1* ortholog from *M. sativa* (Peel et al., 2009). The recently identified MYB TF MtPAR (Verdier et al., 2012) has been shown to regulate PA biosynthesis in seed coats of wild-type *M. truncatula* plants and, if constitutively expressed, also induced PA accumulation in *M. truncatula* hairy roots and to a minor degree in *M. sativa* shoots (Verdier et al., 2012).

Our major goal for this study was to identify and isolate a TF regulating PA biosynthesis in foliar tissues of *Trifolium* species and to test if this TF would induce PA accumulation in leaves of our species of interest, *T. repens*. The clover genus *Trifolium* is one of the largest genera in the family Leguminosae, with approximately 255 species (Ellison et al., 2006). Only two *Trifolium* species, namely *Trifolium arvense* (rabbit's foot clover) and *Trifolium affine* (*Trifolium preslianum*), are known to accumulate significant levels of PAs in leaves (Fay and Dale, 1993). We used *T. arvense* complementary DNA (cDNA) libraries to identify R2R3-MYB TFs specific for the regulation of PA biosynthesis in legume leaves and isolated and functionally characterized the R2R3-MYB TF *TaMYB14* as well as homologs from several *Trifolium* and *Medicago* species. We silenced *TaMYB14* in *T. arvense* and determined that PA accumulation in leaves of transformed plants was almost completely blocked. Heterologous constitutive expression under the control of the cauliflower mosaic virus (CaMV) *35S* promoter of *TaMYB14* was used to test its effects on PA biosynthesis in *Nicotiana tabacum* (tobacco), *M. sativa*, and *T. repens*; PA monomers and oligomers were identified using liquid chromatography-multistage tandem mass spectrometry (LC-MS^n) analysis and compared with known grape seed PA standards. The activation of structural and transporter PA genes in *T. repens* transformants was analyzed using reverse transcriptase (RT)-quantitative PCR (qPCR).

RESULTS

Isolation and Identification of *TaMYB14*, an R2R3- MYB TF from *T. arvense*

To isolate R2R3-MYB TFs potentially involved in PA biosynthesis in *T. arvense*, one of the two *Trifolium* species known to accumulate PAs in foliar tissues (Jones et al., 1976), we initially searched in-house (*T. repens* EST CS35, Pastoral Genomics NR) and public legume databases (http://compbio.dfci.harvard.edu/tgi/plant.html and http://www.kazusa.or.jp/e/resources/index.html; GenBank accession nos. CO511688–CO517334) for sequences containing the highly conserved R2R3-MYB domain (Stracke et al., 2001; Miyake et al., 2003). We designed degenerate primers (all primer sequences used in this study are listed in Supplemental Table S1) to amplify putative R2R3-MYB TFs by RT-PCR of mRNA isolated from immature and mature *T. arvense* leaf tissues, cloned the PCR products, and randomly selected 200 clones for sequencing. Those amplification products likely to be coding for R2R3-MYB TFs were translated and analyzed using internal National Center for Biotechnology Information TBLASTX analysis against the Database of Arabidopsis Transcription Factors (http://datf.cbi.pku.edu.cn/index.php); more than 50% of the sequences did not return any significant hits to known R2R3-MYB protein sequences. The remaining were identified as sequences containing significant similarity to R2R3-MYB TFs related to abiotic stress responses, phytohormone stimuli, stomatal movement, epidermal cell growth, meristem control, and the down-regulation of caffeic acid *O*-methyltransferase (data not shown).

Two partial *T. arvense* cDNAs coded for protein sequences that fell within the MYB clades NO8 and NO9 (Stracke et al., 2001; Jiang et al., 2004), whose members include those known to activate PA or anthocyanin biosynthesis, and were used to design primers to obtain full-length sequences. The full-length *TaMYB14* (GenBank accession no. JN049641) contains a 942-bp open reading frame coding for a 314-amino acid protein. Analysis of the deduced amino acid sequence showed that TaMYB14 contains an N-terminal R2R3 repeat with the bHLH-binding motif [D/E]Lx2[R/K]x3Lx6Lx3R (Fig. 2) typical for anthocyanin- and PA-regulating MYB TFs (Grotewold et al., 2000; Baudry et al., 2004; Zimmermann et al., 2004; Dubos et al., 2010; Hichri et al., 2011). It also contains a nine-residue consensus sequence, V[I/V]R[T/P][K/R]A[I/L/V][R/K]C, in the highly variable C-terminal region (Fig. 2) that is also present in the PA-regulating R2R3-MYB TFs (for GenBank accession numbers, see Fig. 2 legend) AtTT2 (Nesi et al., 2001), LjTT2a, -b, and -c (Yoshida et al., 2008), PtrMYB134 (Mellway et al., 2009), and VvMYBPA2 (Terrier et al., 2009). The previously described C-terminal motif DEDWLR (Stracke et al., 2001; Takos et al., 2006) found in AtTT2 is absent from TaMYB14 as well as from all other aligned sequences described to be involved in PA regulation. Also absent is the motif KPRPR[S/T][F/L] associated with anthocyanin regulation (Quattrocchio et al., 1999; Borevitz et al., 2000; Stracke et al., 2001; Peel et al., 2009). TaMYB14 aligns most closely with R2R3-MYB TFs of the newly defined subgroup 32, which includes the PA-regulating R2R3-MYB TFs PtrMYB134, AtTT2, and VvMYBPA2 (Wilkins et al., 2009). These features suggest that TaMYB14 is involved in regulating PA biosynthesis in *T. arvense*.

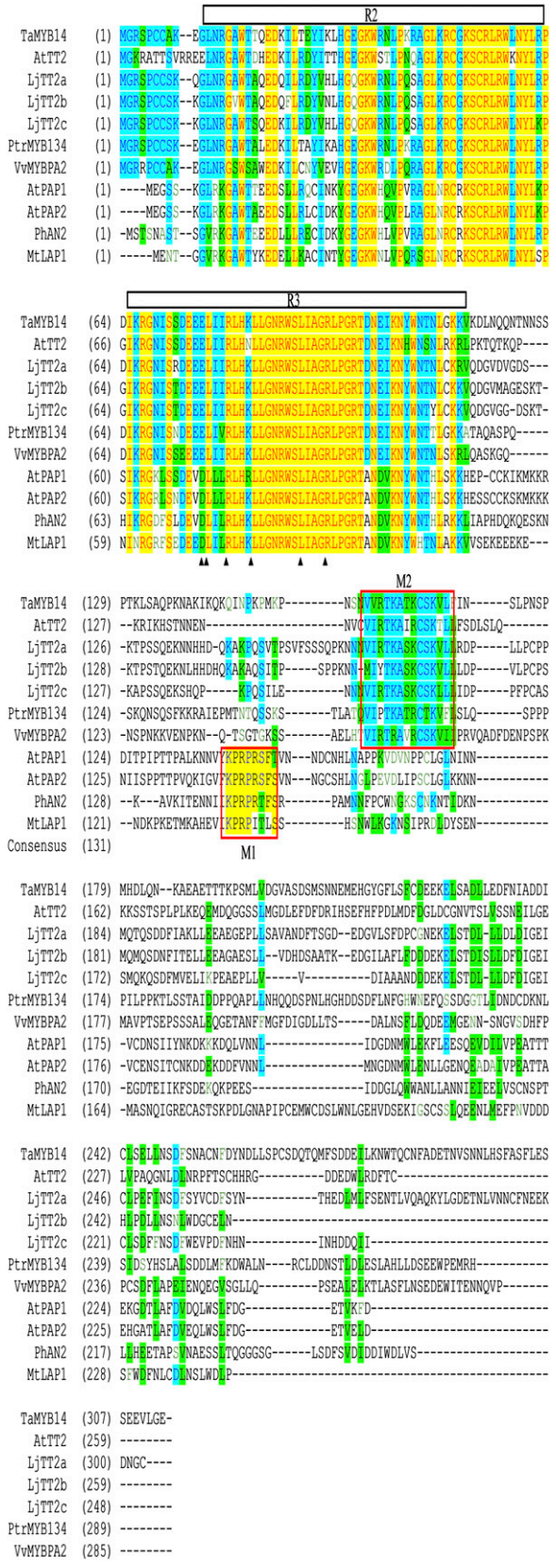

Figure 2. Amino acid alignment of TaMYB14 with known R2R3-MYB regulators of PA and anthocyanin biosynthesis from other species. Black arrowheads identify the amino acid residues for the motif that specifies interaction with bHLH cofactors. M1 is the KPRPRS/T motif identified in anthocyanin-related MYB TFs. M2 is the consensus motif V[V/I]xT[K/R]Ax[K/R]C[S/T]K for PA-related R2R3-MYBs from subgroup 32 (Wilkins et al., 2010). The GenBank accession numbers are as follows: TaMYB14, JN049641; AtTT2, Q9FJA2; LjTT2a, AB300033; LjTT2b, AB300034; LjTT2c, AB300035; PtrMYB134, XP_002308528.1; VvMYBPA2, ACK56131; AtPAP1, CAB09230; AtPAP2, NP176813; PhAN2, AAF66727; and MtLAP1, FJ199998.

Identification of *TaMYB14* Homologs from *Trifolium* Species

Primers designed to the start and stop region of *TaMYB14* were used to amplify full-length homologs from genomic DNA isolated from the *Trifolium* species *T. arvense*, *T. affine*, *T. repens*, and *T. occidentale* and the *Medicago* species *M. sativa* and *M. truncatula*. All of the corresponding full-length genomic *MYB14* sequences had three exons and two introns of varying sizes (data not shown), except for *MtMYB14-1* (JN157820) and *MsMYB14-3* (JN225514), which contained the first intron only. Two PCR products were amplified from *T. arvense* genomic DNA, one of which corresponded to the expressed sequence *TaMYB14*; the other corresponded to an unidentified isoform/allelic variant of *TaMYB14*, designated *TaMYB14_2* (JN052773); expression of this variant was not detected in any tissue of the *T. arvense* accession AZ2925 used in this study. Alignment of the two open reading frames and protein sequences revealed the presence of several deletions and insertions in TaMYB14_2 compared with TaMYB14 (Supplemental Fig. S1). We also identified two MYB14 isoforms/allelic variants in *T. affine* (JN117921, JN117922), *T. occidentale* (JN117926, JN117927), and *M. truncatula* (JN157820, JN157821), whereas *M. sativa* (JN120176, JN120177, JN225514) and *T. repens* (JN117923, JN117924, JN117925) appeared to contain three isoforms. The translated sequences of the identified R2R3-MYB14 TFs shared 95% similarity with TaMYB14, with the majority of amino acid differences being located in the 3′ region downstream of the conserved MYB domain (Supplemental Fig. S1). All of the sequences contained the bHLH-binding and nine-residue consensus sequences as described above.

A phylogenetic tree was constructed with the amino acid sequences of other plant R2R3-MYB TFs involved in the regulation of PAs. Related MYB proteins from *Gossypium hirsutum* (Cedroni et al., 2003), whose functions have not yet been confirmed, were also included. The sequences were first aligned using the multiple sequence alignment program MAFFT (Katoh et al., 2009), and a maximum-likelihood tree was constructed using PhyML from the program SeaView version 4 (Gouy et al., 2010). The tree identifies two broad phylogenetic groups for R2R3-MYBs involved in PA biosynthesis (Fig. 3). Group 1 consists of MYB14 TFs from *Trifolium* and *Medicago* genera isolated in this study, MtPAR recently reported by Verdier et al. (2012), AtTT2 (Nesi et al., 2001), LjTT2a, -b, and -c (Yoshida et al., 2008), VvMYBPA2 (Terrier

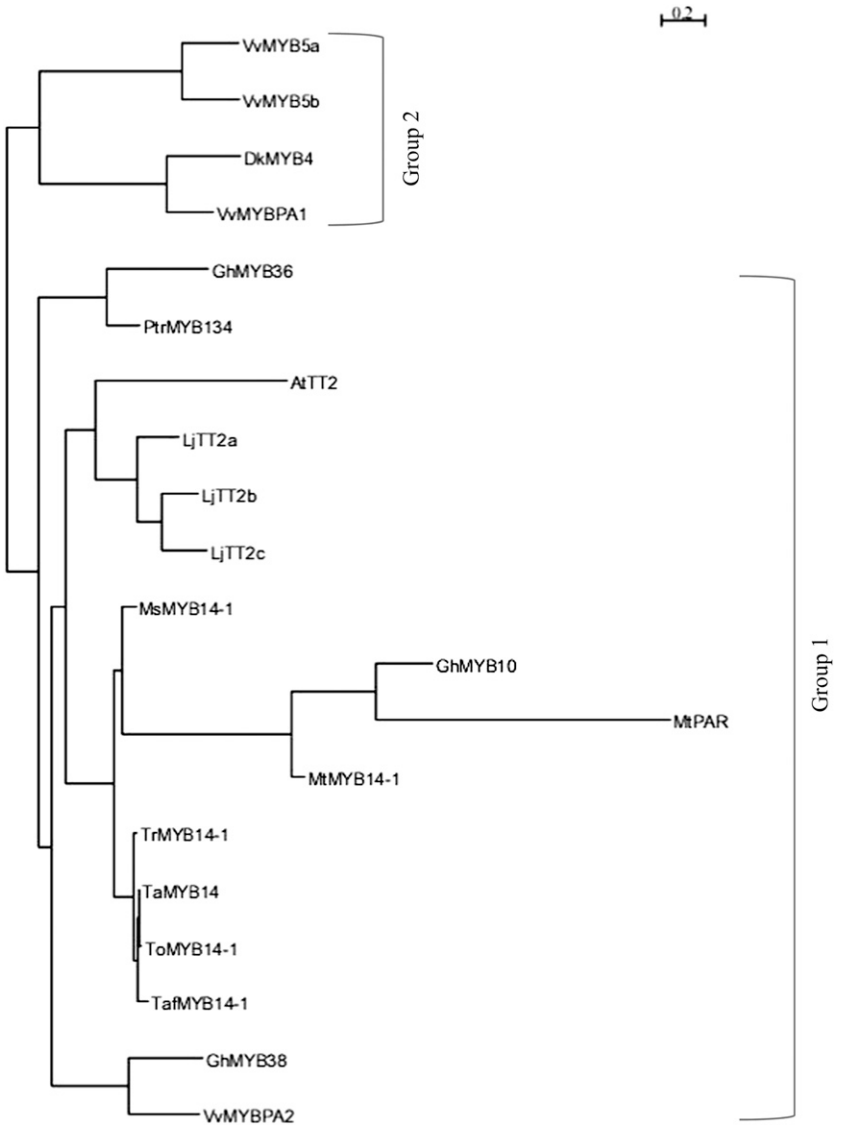

Figure 3. Phylogenetic tree of *Trifolium* and *Medicago* R2R3-MYB14 homologs with selected PA-related R2R3-MYB proteins from other species. The scale bar measures substitutions per site. The length of a branch is the average number of substitutions that occurred at each site of the sequence between the two ends of the branch. Accession numbers are as follows: TaMYB14 (*T. arvense* MYB14; JN049641), AtTT2 (Arabidopsis TT2; Q9FJA2), LjTT2a (*L. japonicus* TT2a; AB300033), LjTT2b (*L. japonicus* TT2b; AB300034), LjTT2c (*L. japonicus* TT2c; AB300035), PtrMYB134 (*P. trichocarpa* MYB134; XP_002308528.1), GhMYB36 (*G. hirsutum* MYB36; AAK19617), GhMYB38 (*G. hirsutum* MYB38; AAK19618), GhMYB10 (*G. hirsutum* MYB10; AAK19615), VvMYBPA1 (*V. vinifera* VvMYBPA1; CAJ90831), VvMYBPA2 (*V. vinifera* VvMYBPA2; ACK56131), DkMYB4 (*D. kaki* DkMYB4; AB503701), TafMYB14-1 (*T. affine* TafMYB14-1; JN117921), TrMYB14-1 (*T. repens* TrMYB14-1; JN117923), ToMYB14-1 (*T. occidentale* ToMYB14-1; JN117926), and MsMYB14-1 (*M. sativa* MsMYB14-1; JN120176).

et al., 2009), PtrMYB134 (Mellway et al., 2009), and GhMYB10, -36, and -38 (GenBank accession nos. AAK19615, AAK19617, and AAK19618, respectively). MYB14 homologs from the *Trifolium* genera have the least number of substitutions between them (as represented by the very short branch lengths) and exhibit greater homology across species than what is observed for the LjTT2 MYB family within the same species. Interestingly, MtPAR, which was shown to activate PA biosynthesis in *Medicago*, is conserved to MtMYB14 sequences only in the MYB domain. Its C-terminal region exhibits considerable variation, and this is reflected in its long branch length in the phylogenetic tree (Fig. 3).

Group 2 includes DkMYB4 (Akagi et al., 2009), VvMYBPA1 (Bogs et al., 2007), and both VvMYB5a and -5b (Deluc et al., 2006, 2008). Subdivision of group 2 MYB TFs identifies two distinct clusters: DkMYB4 with VvMYBPA1, and VvMYB5a and -5b, respectively. None of the group 2 sequences contain the nine-residue consensus sequence V[I/V]R[T/P][K/R]A[I/L/V][R/K]C, whereas VvMYB5a and -5b contain the motif DDxF[S/P]SFL[N/D]SLIN[E/D] (Deluc et al., 2008). VvMYBPA1 and DkMYB4, on the other hand, appear to contain three motifs in the C-terminal region, which have not been described previously: V[H/Y][L/N]PKPxR, [S/T]L[E/D]KLY[E/D]EY[L/F]QLL, and LD[S/Y]FAxSLL[I/V] (Supplemental Fig. S2); these motifs were also found in uncharacterized R2R3-MYB TFs from *Theobroma cacao* (ADD51352), *Populus trichocarpa* (XP_002320876), *Glycine max* (ABH02841), *V. vinifera* (CI40578), and *Lupinus albus* (BAJ61453).

PA Accumulation and Expression Analysis of *MYB14*, *ANR*, and *LAR* in *Trifolium* Tissues

Staining of *T. arvense* and *T. repens* tissues from wild-type plants with 4-dimethylaminocinnamaldehyde

(DMACA; Li et al., 1996) showed that PAs accumulate in the epidermal layers of the entire leaf lamina of *T. arvense*, whereas PAs accumulate only in trichomes on the abaxial epidermal layers of *T. repens* leaves (Fig. 4, A and D).

As a preliminary test for the putative relationship of MYB14 with PA biosynthesis and accumulation, we analyzed various tissues of wild-type *T. arvense* and *T. repens* for the presence of PAs by DMACA staining and the co-occurrence of *ANR*, *LAR*, and *MYB14* transcripts in cDNA libraries isolated from the same tissues (Table I; Supplemental Fig. S3). *TaMYB14* transcripts were present in immature and mature *T. arvense* leaves but not in callus tissues, which do not accumulate PAs (Table I). Transcripts of *MYB14* orthologs were also detected in cDNA libraries from meristematic leaf tissues, trichomes, and floral tissues from *T. repens* and *T. occidentale* but not in young and mature leaf tissues, stolons, internodes, roots, and petioles (Table I). No *MYB14* transcripts were detected in a red-leafed *T. repens* line (cv Isabelle), which accumulates high levels of anthocyanins in foliar tissues but does not accumulate foliar PAs except in trichomes. Expression of *MYB14* was paralleled by the expression of the structural genes *ANR* and *LAR* (Table I; Supplemental Fig. S3), which code for enzymes catalyzing the last steps of PA monomer synthesis (Fig. 1). This indicates that expression of *MYB14* coincides with PA biosynthesis and that PA biosynthesis is activated in *T. arvense* in both immature and mature tissues, whereas it is only activated in meristematic leaf and floral tissues of *T. repens*.

Functional Analysis of TaMYB14 in *T. arvense* by Gene Silencing

To test the function of TaMYB14, the gene was effectively silenced by transformation of *T. arvense* with an RNA interference (RNAi) construct (Helliwell and Waterhouse, 2003); silencing of *TaMYB14* was confirmed by RT-PCR (data not shown). Initially, transformed callus tissues and regenerated plants were analyzed for PAs using DMACA staining. Neither wild-type nor silenced *T. arvense* callus tissues contained PAs; leaves from regenerated wild-type plants stained positive (Fig. 4A), whereas leaves from regenerated *TaMYB14*-silenced plants showed only a light blue staining in the leaves (Fig. 4E), indicating that PA biosynthesis was severely reduced in *TaMYB14*-silenced plants. RT-qPCR analysis of *MYB14* and *ANR* expression in silenced plants from nine independent transformation events revealed that transcript levels in most lines were under the detection limit (Table II); *MYB14* expression was detectable in

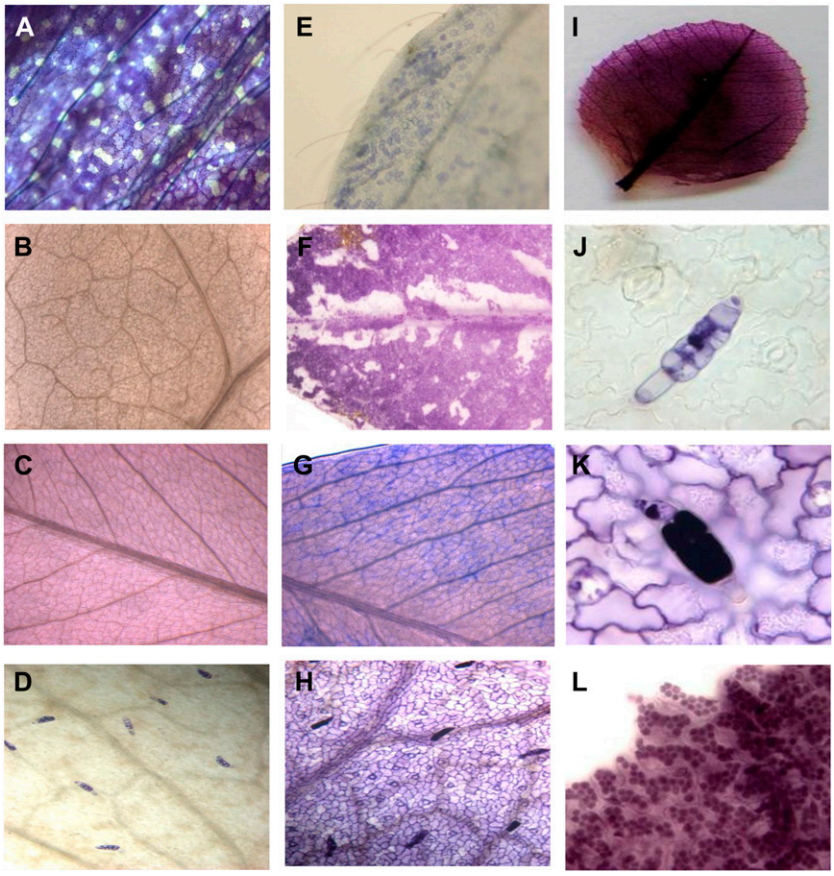

Figure 4. DMACA staining of leaves from wild-type and transgenic plants. A to D, DMACA-stained leaves (10×) of wild-type *T. arvense* (A), *N. tabacum* (B), *M. sativa* (C), and *T. repens* (D). E to H, DMACA-stained leaves (10×) of *T. arvense* with RNAi-silenced *TaMYB14* (E) and *N. tabacum* (F), *M. sativa* (G), and *T. repens* (H) plants constitutively expressing *TaMYB14*. I to L, DMACA-stained leaf from a *TaMYB14*-expressing *T. repens* line with very high levels of PA accumulation (I) and closeups (100×) of leaf epidermis with a PA-containing trichome of wild-type (J) and *TaMYB14*-expressing (K) *T. repens* and spongy mesophyll cells of *TaMYB14*-expressing *T. repens* (L).

Table I. *PA accumulation (indicated by DMACA staining) and the presence of MYB14, ANR, and LAR transcripts in tissue-specific cDNA libraries from T. arvense, T. repens, and T. occidentale*

Transcripts were detected using RT-PCR (Supplemental Fig. S3).

Species and Cultivar	cDNA Libraries	DMACA	MYB14 cDNA	ANR cDNA	LAR cDNA
T. arvense AZ2951	Immature and mature leaves	Entire leaf	+	+	+
T. arvense AZ2951	Callus	−	−	−	−
T. repens Huia	Meristematic leaf and floral tissues	Trichomes only	+	+	+
T. repens (Huia, Isabelle) and *T. occidentale* (AZ4270)	Young and mature leaves, petioles, stolons, roots	Trichomes only	−	−	−
T. repens Huia	Mature flowers	Entire petal	−	−	−

one silenced line only, whereas *ANR* transcripts were detectable in four of the nine tested silenced lines, albeit at a more than 1,000-fold lower level than in wild-type plants (Table II).

Extracts of regenerated wild-type *T. arvense* plantlets were analyzed by LC-MS^n to confirm the presence of PAs. For comparison purposes, a commercial grape seed extract was analyzed under the same conditions, as the PA composition of such extracts has been extensively examined by LC-mass spectrometry (MS; Hayasaka et al., 2003; Monagas et al., 2006; Weber et al., 2007). A targeted ion-trap MS^n approach (Fraser et al., 2012) was used to detect and provide characterizing data on PA monomers, catechin and epicatechin (procyanidins [PC]), and gallocatechin and epigallocatechin (prodelphinidins [PD]), as well as B-type PA dimers and trimers. A comprehensive study of the fragmentation of PAs reported by Li and Deinzer (2007) detailed many of the fragmentation mechanisms observed in an ion-trap mass spectrometer for PAs up to trimers, aiding in reliable detection and classification of the PA structures. For example a PC:PC dimer MS^2 (daughter ion) spectrum (Fig. 5D) gives the retro-diels alder fragments 409 and 427 mass-to-charge ratio (m/z) and a monomer fragment 291 m/z due to interflavanic bond cleavage. This approach of comparing MS^2 spectra from samples of known PA composition with unknown samples provides the qualitative structural evidence for confirming that coeluting peaks in different samples are in fact the target compound of interest (Fraser et al., 2012). The MS^2 spectra of known PAs from the grape seed extract were compared with coeluting MS^2 spectra in *T. arvense* extracts for qualitative confirmation of the PAs. Extracts of wild-type *T. arvense* leaves contained high levels of PA monomers (Supplemental Fig. S4), mainly catechin and, to a lesser extent, gallocatechin, as well as PC:PC (Fig. 5B), PC:PD and PD:PD dimers (data not shown), and trimers of the various PC:PD combinations (Supplemental Fig. S5). Figure 5, D and E, shows MS^2 spectral evidence that the highlighted PC:PC dimer peak in the grape seed control and the *T. arvense* wild-type extracts are identical. Our method also detects multiple dimers and trimers in the grape seed and experimental samples by LC-MS^n; however, without authentic standards, we are unable to assign their identities, so all polymers are referred to by the PA subunits they are constructed of (e.g. PC:PC dimers being those containing catechin and/or epicatechin). No higher degree of polymerization (DP) polymers were seen in MS^1 (parent ion) scans, but further work on concentrated acetone fractions revealed the presence of up to DP6 PAs in wild-type *T. arvense* leaves (data not shown). In contrast, only traces of PA monomers, but no dimers (Fig. 5C), trimers (Supplemental Fig. S5), or larger PA polymers were detected in leaf extracts from silenced plants, and no PAs were detected using butanol-HCl extraction. These results clearly indicate that PA biosynthesis had been severely reduced by *TaMYB14* silencing and that the expression of *TaMYB14* is required for PA synthesis in *T. arvense*.

Comparison of LC-UV chromatograms (280 and 350 nm) revealed substantial differences in flavonoid and isoflavonoid levels between different genotypes of wild-type as well as silenced plants (data not shown), but there was no consistent effect observed on the levels of flavonoids between the silenced and wild-type plants, suggesting that only PA biosynthesis was reduced by *TaMYB14* silencing.

Functional Analysis of TaMYB14 in Transgenic *N. tabacum* Plants

N. tabacum is a plant species that, like Arabidopsis, accumulates PAs in seed coats only and has been used previously to test the function of PA-regulating R2R3-MYB TFs (Deluc et al., 2006, 2008). It has also been shown that constitutive expression of PA-regulating TFs in Arabidopsis can result in dwarfed phenotypes and death at early stages of development (Bogs et al., 2007). Therefore, we tested the function of TaMYB14 in the heterologous plant system tobacco (rather than Arabidopsis) by transforming plants with *TaMYB14*

Table II. *Analysis of MYB14 and ANR transcript levels in leaf tissues from T. arvense wild type and TaMYB14-silenced lines (A–I)*

Transcript levels were normalized using the housekeeping gene *Actin*.

T. arvense Line	*MYB14*	*ANR*
Wild type	1.29E-03	4.30E-01
A, B, H, I	0	0
C	0	2.95E-04
D	0	1.54E-03
E	3.68E-04	0
F	0	2.50E-04
G	0	2.81E-04

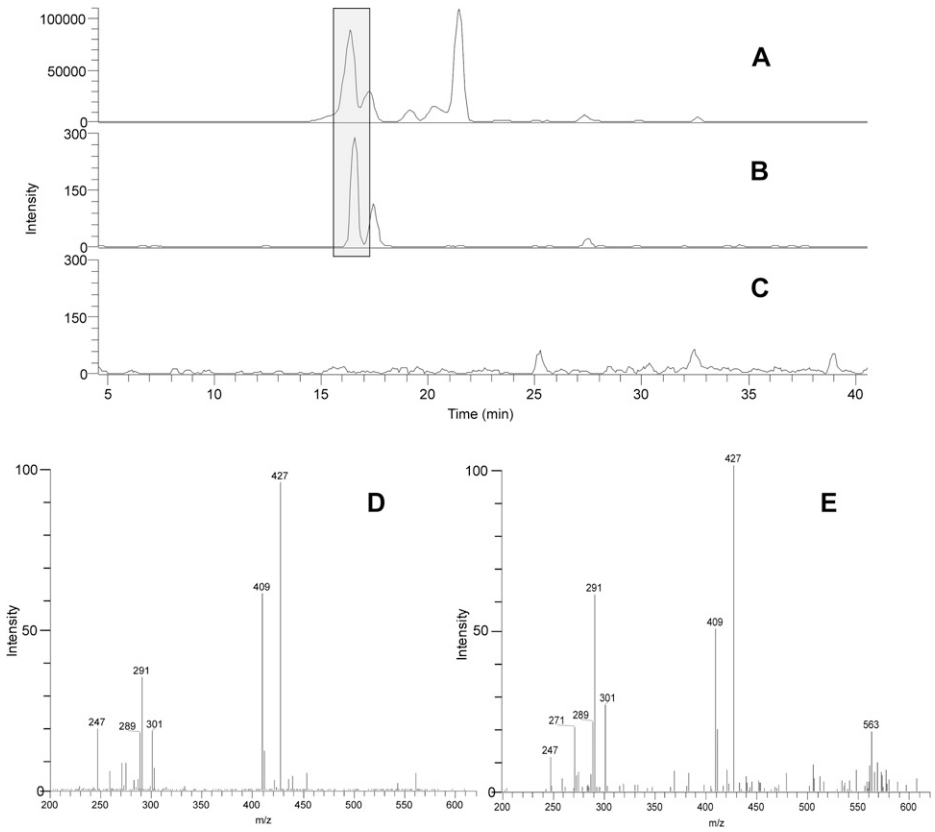

Figure 5. LC-MS^n composite extracted ion chromatograms of ions for PA dimers in *T. arvense* leaf extracts. A to D, Chromatograms are the sum of ions 291 + 409 + 427 *m/z* from MS^2 product ion scans of 579 *m/z* (PC:PC) dimers in extracts of grape seeds (A), *T. arvense* wild type (B), *T. arvense* *TaMYB14*-silenced leaves (C), and MS^2 spectrum of the highlighted PC:PC dimer in grape seeds (D). E, MS^2 spectrum of the highlighted PC:PC dimer in *T. arvense* wild type.

under the control of the CaMV *35S* promoter. Integration into the genome and expression of *TaMYB14* were confirmed by RT-PCR (data not shown). The effects of introducing *TaMYB14* into wild-type ("green" tobacco) plants as well as into tobacco plants expressing the *Solanum lycopersicum* (tomato) *LeANT1* TF under the control of the constitutive cassava vein mosaic promoter (Verdaguer et al., 1996; Mathews et al., 2003) on PA accumulation were tested by DMACA and LC-MS^n analysis. *LeANT1*-expressing tobacco plants accumulate high levels of anthocyanins ("red" tobacco), and we used these plants for transformation with the CaMV *35S*::*TaMYB14* construct to ensure the presence of sufficient precursor flavonoids for the production of PAs. Leaves from both the green and red tobacco plants transformed with the CaMV *35S*::*TaMYB14* construct stained positive for PAs with DMACA, whereas untransformed tobacco plants did not (Fig. 4, B and F), indicating that PA biosynthesis occurred due to the expression of *TaMYB14*. PA monomers (mainly epicatechin with small amounts of epigallocatechin and gallocatechin; Supplemental Fig. S6), PC:PC dimers (Fig. 6), and trace levels of trimers (data not shown) were detected by LC-MS^n analysis of flavonoid extracts from *TaMYB14*-expressing tobacco plants, whereas PAs were undetectable in control plants. The red *TaMYB14*-expressing tobacco lines did not accumulate more PA monomers and oligomers compared with the green lines; however, a loss of red color and corresponding anthocyanins in the former was noticeable (data not shown). Comparison of LC-UV chromatograms (280 and 350 nm) showed that there was no consistent effect of TaMYB14 expression on tobacco leaf flavonol levels (data not shown). These results provide strong evidence that TaMYB14 is able to up-regulate the biosynthesis of PAs, even in plants phylogenetically distant from *T. arvense*, and indicate that the up-regulation of PA biosynthesis competes with substrates for anthocyanin biosynthesis or accumulation.

Functional Analysis of TaMYB14 in Transgenic *M. sativa* Plants

M. sativa plants were transformed with *TaMYB14* under the control of the CaMV *35S* promoter to test the function of TaMYB14 in a legume; the presence and expression of *TaMYB14* were confirmed by RT-PCR (data not shown). Leaves from regenerated plantlets were screened for PA accumulation using DMACA staining, and a number of plants transformed with *TaMYB14* tested positive. Leaves from nontransformed wild-type plants stained positive with DMACA in the trichomes on the abaxial leaf layers only, whereas plants transformed with *TaMYB14* stained positive in epidermal leaf cells as well (Fig. 4, C and G). The presence of PA monomers (epicatechin and catechin; Supplemental Fig. S7), PC:PC dimers (Fig. 7), PC:PC:PC and PC:PC:PD trimers (Supplemental Fig. S8), and trace levels of tetramers in leaf extracts of *M. sativa* plants transformed with the *TaMYB14* construct was

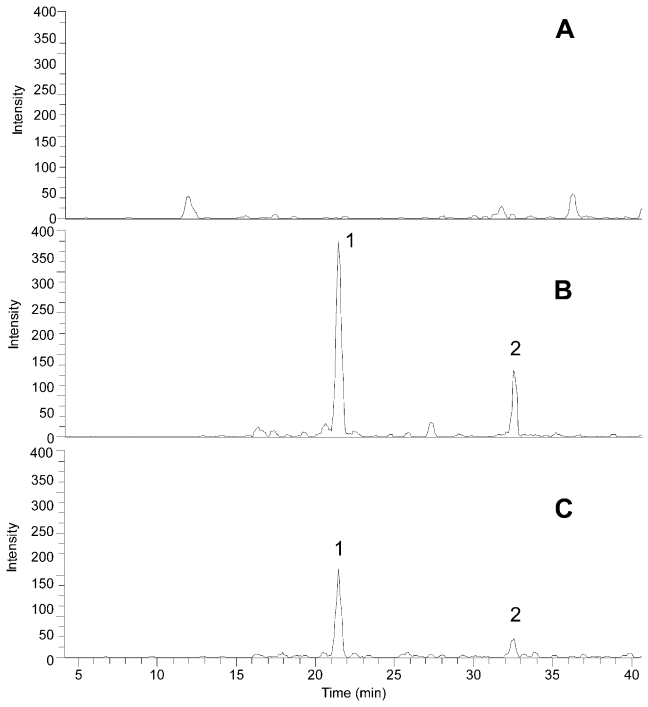

Figure 6. LC-MS^n composite extracted ion chromatograms of ions for PA dimers in *N. tabacum* leaf extracts. Chromatograms are the sum of ions 291 + 409 + 427 *m/z* from MS^2 product ion scans of 579 *m/z* (PC:PC dimers). Peaks 1 and 2 have MS^2 spectra consistent with PC:PC dimers. The chromatograms are from leaf extracts of *N. tabacum* wild type (A), green *N. tabacum* transformed with CaMV *35S*::*TaMYB14* (B), and red *N. tabacum* transformed with CaMV *35S*::*TaMYB14* (C).

confirmed by LC-MS^n analysis, whereas PAs were undetectable in control plants. A glycosylated monomer previously identified in *M. truncatula* seed coats, epicatechin glycoside (Pang et al., 2008), was also detected by LC-MS^n (MS^1 *m/z* 453, MS^2 *m/z* 291, MS^3 *m/z* 123, 139, 151, 165) in *TaMYB14*-transformed plants only, with levels 10-fold lower relative to free epicatechin (data not shown). There was no consistent effect of *TaMYB14* expression on leaf flavonoid levels (data not shown).

Quantification of soluble PAs using the butanol-HCl method (Terrill et al., 1992) in leaves of eight independent *M. sativa* lines constitutively expressing *TaMYB14* showed an accumulation of PAs in six of the tested lines, ranging from 0.1% to 1.0% dry mass (Supplemental Table S2).

Functional Analysis of TaMYB14 in Transgenic *T. repens* Plants

To test the function of TaMYB14 in *T. repens*, a close relative of *T. arvense* and widely used in pastoral mixed grazing systems, plants were transformed with *TaMYB14* under the control of the CaMV *35S* promoter; the presence and expression of the transgene were confirmed by RT-PCR (data not shown). Leaves from regenerated plantlets were screened for PA accumulation using DMACA staining, and a number of plants transformed with *TaMYB14* tested positive. Leaves from nontransformed wild-type plants showed, as expected, positive DMACA staining in the trichomes on the abaxial leaf layers only (Fig. 4D), whereas plants transformed with *TaMYB14* accumulated PAs in epidermal leaf cells as well as in trichomes (Fig. 4H). Microscopic examination under higher magnification revealed that stomatal guard cells as well as mesophyll cells stained positive with DMACA (Fig. 4, H and L). Staining in trichome tier and apical cells was much stronger compared with wild-type trichomes (Fig. 4, J and K), and staining of spongy mesophyll cells indicated PA accumulation in multiple vacuole-like organelles (Fig. 4L). In comparison, wild-type *T. arvense* plants accumulate PAs in epidermal leaf cells only as shown by DMACA staining of a young leaf cross-section (Supplemental Fig. S9). In some transgenic *T. repens* plants, PAs were also detected in root and petiole cells (data not shown). A number of plants appeared to accumulate very high levels of foliar PAs, indicated by almost black staining with DMACA (Fig. 4I); however, all of these plants died before sufficient material for LC-MS^2 analyses could be harvested. These results indicate that the constitutive expression of *TaMYB14* results in the ectopic production and accumulation of PAs in foliar (and other) tissues and cell layers of *T. repens* that normally do not accumulate PAs.

To confirm the presence of PA units in *T. repens* leaves expressing *TaMYB14*, leaf extracts were analyzed by LC-MS^n. No PA units were detected in wild-type *T. repens* leaves, whereas PA monomers (Supplemental Fig. S10), dimers (Fig. 8), as well as trace levels of trimers (Supplemental Fig. S11) were detected in *T. repens* plants transformed with the CaMV *35S*::

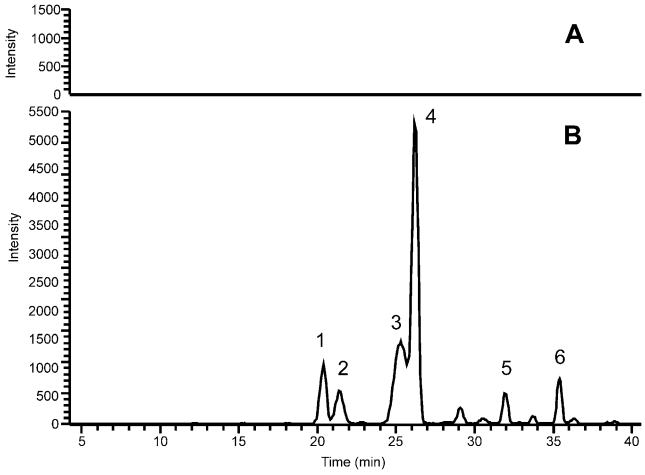

Figure 7. LC-MS^n composite extracted ion chromatograms of ions for PA dimers in *M. sativa* leaf extracts. Chromatograms are the sum of ions 291 + 409 + 427 *m/z* from MS^2 product ion scans of 579 *m/z* (PC:PC dimers). The chromatograms are from leaf extracts of *M. sativa* wild type (A) and *M. sativa* transformed with CaMV *35S*::*TaMYB14* (B).

TaMYB14 construct. The major monomer detected was epigallocatechin, with 10-fold lower levels of epicatechin and gallocatechin and only traces of catechin detected (Supplemental Fig. S10). The dimers and trimers detected were PC and PD or all PD based, consistent with the observation of both PC and PD monomers in transformed *T. repens* leaves. As seen for *T. arvense*, there were considerable differences in flavonol and isoflavonoid levels in different *T. repens* genotypes; however, there was no consistent effect of *TaMYB14* expression on leaf flavonol and isoflavonoid levels (Supplemental Fig. S12). These results provide evidence for *TaMYB14* being able to specifically activate the biosynthesis and accumulation of PAs in white clover foliar tissues.

Quantification of soluble PAs in leaves of 16 independent *T. repens* lines constitutively expressing *TaMYB14* showed an accumulation of PAs ranging from 0.5% to 1.8% dry mass. We selected five of these lines for further analysis of transcript levels of phenylpropanoid biosynthetic and transporter genes (Fig. 9).

Analysis of the Expression of Phenylpropanoid Pathway Genes in *T. repens* Plants Expressing *TaMYB14*

To analyze the effects of TaMYB14 on the expression of genes coding for enzymes of the phenylpropanoid pathway and a putative MATE-like PA transporter, RNA was isolated from leaves of *T. repens* plants constitutively expressing *TaMYB14*. Nucleotide sequences coding for putative *TrCHS*, *TrF3′H*, *TrF3′5′H*, *TrFLS*, *TrDFR*, *TrANS*, *TrANR*, *TrLAR*, *TrMATE1*, and *TrMATE2* were identified from our in-house *Trifolium* gene databases, and mRNA levels were quantified in five independent lines transformed with *TaMYB14* using RT-qPCR with gene-specific primers (Fig. 9; for primer sequences, see Supplemental Table S1).

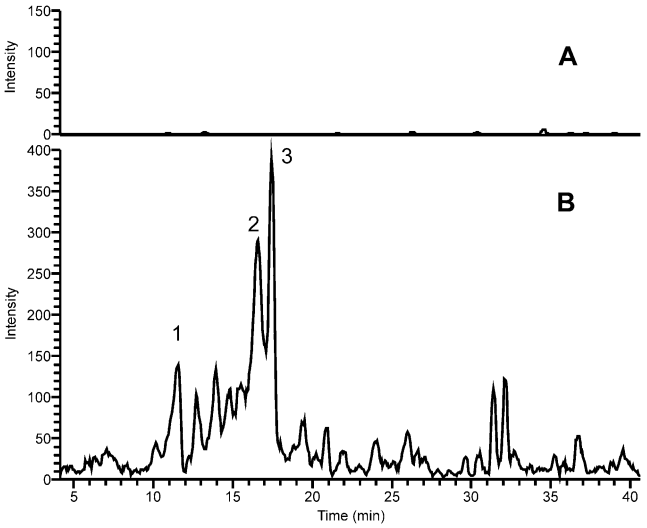

Figure 8. LC-MS^n composite extracted ion chromatograms of ions for PA dimers in *T. repens* leaf extracts. Chromatograms are the sum of ions 291 + 307 + 409 + 427 + 443 *m/z* from MS^2 product ion scans of 595 *m/z* (PC:PD) dimers. Peaks 1 to 3 have spectra consistent with PC:PD dimers. The chromatograms are from leaf extracts of *T. repens* wild type (A) and *T. repens* transformed with CaMV *35S::TaMYB14* (B).

No consistent differences in the expression of *TrCHS*, *TrF3H*, *TrFLS*, and *TrMATE2* in transgenic *TaMYB14*-expressing compared with wild-type *T. repens* leaves were detected, although CHS transcript levels were higher (approximately 4-fold) in three of the five tested lines. Expression of *TrF3′5′H*, *TrDFR*, and *TrANS* was consistently up-regulated in transformed plants compared with wild-type plants, with a very strong induction of *TrF3′5′H* (more than 600-fold). Expression of the putative PA-specific genes *TrANR* and *TrLAR* could only be detected in leaves from plants constitutively expressing *TaMYB14* but not in wild-type plants. We also analyzed the expression of two genes putatively coding for MATE transporters and show here that *TrMATE1* was not expressed in wild-type plants but was highly expressed in *T. repens* plants transformed with CaMV *35S::TaMYB14*, whereas *TrMATE2* was expressed in wild-type plants and not up-regulated in the transgenic lines. *TaMYB14* expression varied considerably in the transgenic lines and was highest in the line with the highest PA levels (Fig. 9). Our results demonstrate that expression of the R2R3-MYB TF *TaMYB14* is necessary and sufficient to up-regulate genes of the PA pathway and to activate the synthesis and accumulation of PAs in leaves of *T. repens*. All transgenic plants were grown in a greenhouse under the same conditions. However, one previously green-leafed line that accumulated PAs to levels comparable with the other transgenic lines (indicated by DMACA staining; data not shown) changed its phenotype after 6 months at the onset of winter and subsequently exhibited a red-leafed phenotype. At this stage, high levels of anthocyanins but only trace amounts of PAs were detectable. The three anthocyanins detected were shown by LC-MS^n to be pelargonidin, cyanidin, and delphinidin glycosides (Supplemental Fig. S13). Expression analysis (Supplemental Table S3) revealed that all of the putative phenylpropanoid pathway genes related to the production of anthocyanins (i.e. *CHS, F3H, DFR, ANS,* and *F3′5′H*) as well as both *MATE* transporters were up-regulated, with expression levels for *DFR, ANS*, and *F3′5′H* in this line being much higher (more than 200-fold) compared with the PA-accumulating green-leafed *T. repens* lines. No expression of *TrANR* and *TaMYB14* could be detected in this line, whereas *TrLAR* expression was much lower compared with the green-leafed *TaMYB14* transformants.

DISCUSSION

TaMYB14 Is an R2R3-MYB TF Regulating PA Biosynthesis in Legumes

Only a few out of the 255 known *Trifolium* species accumulate PAs in leaves, one of which is *T. arvense* (Jones et al., 1976). Therefore, we used leaf tissues of *T.*

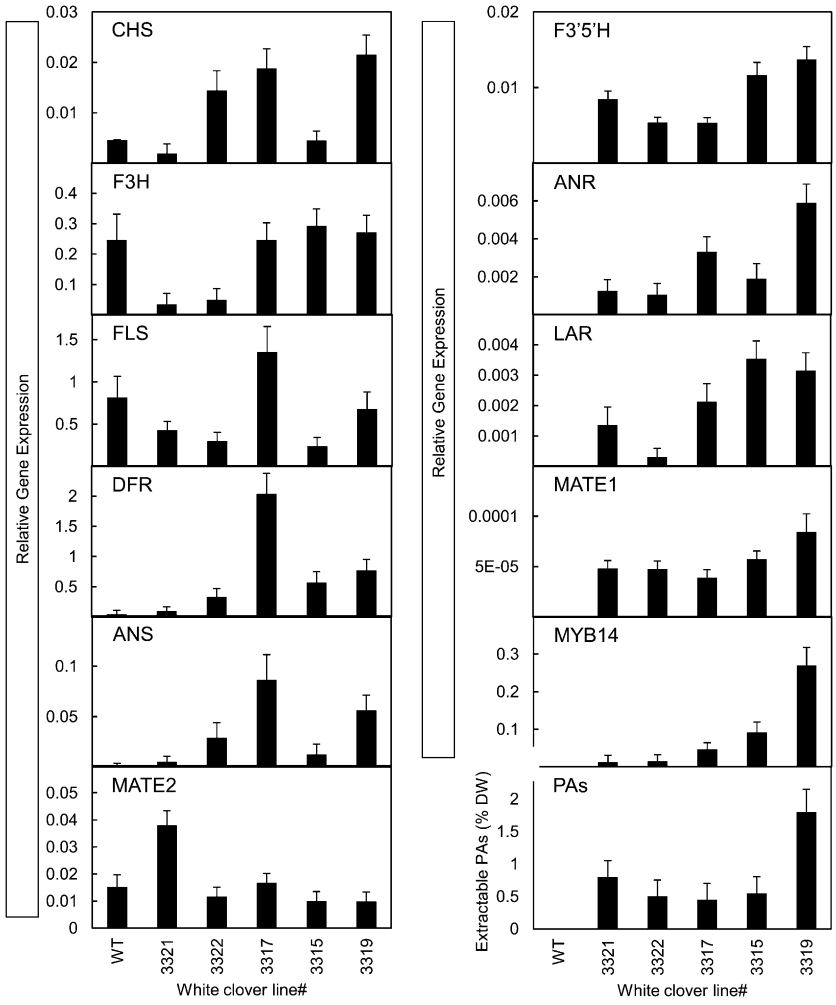

Figure 9. Analysis of phenylpropanoid and transporter gene expression and PA accumulation in *T. repens* leaf tissue. Transcript and extractable PA levels in wild-type (WT) and *TaMYB14*-expressing *T. repens* lines are shown. Transcript levels were normalized relative to the transcript levels of *T. repens* actin. PAs were extracted and quantified using the butanol-HCl method. Three technical replicates were analyzed for each line. Error bars represent SD. DW, Dry weight.

arvense to isolate the R2R3-MYB14 TF regulating foliar PA biosynthesis. Several lines of evidence are presented here that indicate that TaMYB14 is in fact regulating PA biosynthesis: (1) expression of *TaMYB14* coincides with *ANR* and *LAR* expression and PA accumulation in *T. arvense* tissues; (2) silencing of *TaMYB14* in *T. arvense* almost completely blocked foliar PA biosynthesis; and (3) constitutive expression of *TaMYB14* in *N. tabacum*, *M. sativa*, and *T. repens* resulted in ectopic accumulation of PAs in leaf tissues of these plants as well as up-regulation of PA biosynthetic pathway genes in *T. repens* transformants.

Silenced *TaMYB14 T. arvense* plants still showed very weak DMACA staining and traces of PA monomers, whereas PA dimers and trimers were completely absent. The synthesis of traces of PA monomers might have been the result of very low residual expression of *TaMYB14*, although we were not able to detect any *TaMYB14* transcripts in these plants. It is also possible that *T. arvense* possesses one or more additional PA-regulating R2R3-MYB TFs, as has been described for grapevine and poplar (Bogs et al., 2007; Deluc et al., 2008; Terrier et al., 2009; Wilkins et al., 2009). The *TaMYB14* ortholog *TaMYB14_2* identified in this study was not expressed in any of the tissues tested, and it is not clear if this gene is in fact coding for a functional variant of TaMYB14 in *T. arvense*.

R2R3-MYB TFs have been described to regulate PA biosynthesis, often in a complex with bHLH and WD40 proteins (Baudry et al., 2004; Yoshida et al., 2008, 2010; Hichri et al., 2011), in a wide range of plants. In our study, ectopic expression of *TaMYB14* under the control of the CaMV *35S* promoter resulted in PA accumulation in leaves of *N. tabacum*, *M. sativa*, and *T. repens* without the need for coexpression of additional factors. Transcript analysis of *T. repens* transformants showed expression of the putative PA-specific genes *TrANR*, *TrLAR*, and *TrMATE1*, which were completely silent in wild-type *T. repens* leaves, and a more than 600-fold up-regulation of *TrF3′5′H*. The expression profile of *TrMATE1* would be consistent with it being the ortholog of the PA-specific MATE transporter *TT12* (Debeaujon et al., 2001; Marinova et al., 2007), whereas *TrMATE2*, which was expressed in both wild-type and transformed plants, seems to be an ortholog of the anthocyanin/flavonol *MtMATE2* transporter from *M. truncatula* (Zhao et al., 2011). Genes required for precursors of both the anthocyanin and PA pathways (i.e. *DFR* and *ANS*) were also up-regulated,

but to a lesser extent, whereas early flavonoid pathway genes (i.e. *CHS, F3H,* and *FLS*) were not consistently upregulated. In comparison, overexpression of the recently identified TF *MtPAR* in *M. truncatula* hairy roots did not induce the expression of *CHS*, *FLS*, *DFR*, and *LAR* but did result in higher transcript levels of *F3H, ANS,* and *ANR* (Verdier et al., 2012). In contrast, overexpression of *VvMYBPA2* or *VvMYBPA1* in grapevine hairy roots and *PtrMYB134* in poplar induced the expression of both early and late flavonoid pathway genes as well as the PA-specific genes *ANR* and *LAR1* (Mellway et al., 2009; Terrier et al., 2009). Constitutive expression of these R2R3-MYB TFs was sufficient to induce PA biosynthesis without the need for additional factors such as bHLH or WD40 repeat proteins, as was seen in our study for TaMYB14 as well. This is in contrast to the situation in grapevine cell suspensions, where cotransformation with a bHLH factor was required for VvMYBPA1 and VvMYB5b to activate the promoters of PA-specific genes in a transient luciferase assay (Bogs et al., 2007; Deluc et al., 2008). This cotransformation was also required for activating *AtDFR*, *AtANS*, and *AtANR* promoters with LjTT2s in transient Arabidopsis leaf cell assays, whereas constitutive overexpression of *LjTT2a* was sufficient to result in the expression of *AtCHS*, *AtDFR*, and *AtANR* and the accumulation of PAs in stably transformed Arabidopsis seedlings (Yoshida et al., 2010). Constitutive expression of R2R3-MYB TFs generally did not result in PA accumulation in all cells of transformed plants in the above studies but rather in particular cell types, as seen in our study as well. These results could indicate that the required cofactors, such as the AtTT8 and AtTTG1 orthologs, are constitutively expressed in at least some plant tissues and hairy root cells, but not in cell cultures. Constitutive expression of a WD40 repeat protein (MtWD40-1) has been demonstrated in *M. truncatula* (Pang et al., 2009).

One white clover line transformed with *TaMYB14* exhibited a deeply red-colored phenotype due to anthocyanin production and accumulation. This line initially expressed *TaMYB14*, contained PAs, and showed no anthocyanin accumulation for the first 6 months until the onset of winter, after which it started to accumulate large amounts of anthocyanins. At that stage, only trace levels of PA and no *TaMYB14* and *ANR* transcripts were detectable in leaves of this line, whereas *DFR, ANS,* and *F3′5′H* transcripts were strongly induced. Interestingly, both *MATE1* and *MATE2* genes were also more highly expressed in the red line compared with the other *TaMYB14* transformants. Recent studies have shown that the Arabidopsis MATE transporter TT12 takes up epicatechin-3′-glucoside (E3′G), a putative PA precursor in *M. sativa* (Pang et al., 2008), more efficiently than the anthocyanin cyanidin-3-glucoside (Cy3G), whereas MtMATE2 preferentially takes up Cy3G (Zhao et al., 2011). It was also shown that at high Cy3G concentrations, E3′G uptake was inhibited (Zhao and Dixon, 2009). Therefore, the high concentrations of anthocyanins found in the red *TaMYB14*-expressing clover line might have inhibited vacuolar PA monomer uptake and accumulation. This could have prevented PA polymerization in the vacuole, as indicated by the absence of oligomers in this line. However, it is unclear what caused the repression of *TaMYB14* expression, and more studies are needed to identify factors leading to the switch from PA to anthocyanin accumulation in this line.

Chemical Composition of PAs in *TaMYB14*-Expressing Plants

It has been shown previously that PA detection by DMACA staining or HPLC analysis alone can be misleading (Peel et al., 2009). Therefore, in this study, we used targeted LC-MS^n analysis and comparison with known grapevine PA standards, which allows the unambiguous and sensitive identification of PA monomers and oligomers based on their specific ion fragmentation patterns (Fraser et al., 2012). The majority of PA monomers accumulating in *TaMYB14*-transformed white clover leaves were epigallocatechin, with 10-fold lower concentrations of epicatechin and gallocatechin, and only traces of catechin. Analysis of the PAs in wild-type *T. repens* flowers has shown that they consist mostly of gallocatechin- and epigallocatechin-based monomeric units, with equal proportions of gallocatechin and epigallocatechin in the terminator units and a 1:2 ratio of gallocatechin to epigallocatechin in the extender units (Foo et al., 2000). Our findings of preferential accumulation of epigallocatechin in *TaMYB14*-expressing clover leaves is consistent with this; however, concentrations of epigallocatechin were at least 10-fold higher (rather than 2-fold, as seen in wild-type clover flowers) compared with gallocatechin, which might indicate that TaMYB14 induced *TrANR* more than *TrLAR*, although this is not apparent from the RT-qPCR results. Nothing is known about the composition of PAs detected in DMACA-positive trichomes of *T. repens* leaves, and we were not able to detect any PA-specific ions by LC-MS/MS even in trichome-enriched extracts of immature wild-type *T. repens* leaves, which could simply be due to PA concentrations being below the detection limit. However, it is also possible that the compounds responsible for positive DMACA staining of wild-type trichomes are not PAs, as Peel et al. (2009) reported that in *LAP1*-overexpressing *Medicago* plants, DMACA-positive compounds accumulated, which, upon more detailed analysis, turned out not to be PAs.

Native PAs in *M. sativa* seeds are mainly based on epicatechin and epigallocatechin (91%), with much lower concentrations of catechin- and gallocatechin-based PAs (Koupai-Abyazani et al., 1993). In contrast, *M. sativa* foliar monomers induced by ectopic expression of *TaMYB14* in this study consisted almost equally of epicatechin and catechin, whereas (epi)gallocatechin monomers were undetectable. This might indicate that TaMYB14 induced both *ANR* (resulting in epicatechin production) and *LAR* (resulting in catechin production) almost equally, as was seen in our *T. repens*

transformants, whereas the native *M. sativa* PA regulator(s) in seeds might be more active with *MsANR* rather than *MsLAR*. This would be consistent with the induction of *ANR*, but not *LAR*, in hairy roots of *M. truncatula* overexpressing *MtPAR* (Verdier et al., 2012). We could not detect any 5′-hydroxylated epigallocatechin or gallocatechin monomers in *TaMYB14*-transformed *M. sativa*, and the majority of the foliar PA dimers and trimers were PC based, with only traces of a PC:PC:PD trimer identified. This is consistent with a 14:1 ratio of PC- to PD-based PAs found in seeds (Koupai-Abyazani et al., 1993) and could mean that *M. sativa* has only minor F3′5′H activity, which is also not induced by TaMYB14. Transcriptomic analysis of *MtPAR*-transformed hairy roots also did not show any up-regulation of a *F3′5′H* (Verdier et al., 2012).

Pang et al. (2008) showed that more than 75% of epicatechin accumulates as E3′G in developing *M. sativa* seed coats, whereas its content declined during seed development and was absent in mature seeds. We also detected E3′G in *TaMYB14*-transformed *M. sativa* leaves, but only at one-tenth of the concentration of the free PA monomers, which could reflect insufficient up-regulation of the corresponding UDP-glucosyltransferase by TaMYB14. We did not detect any glycosylated PA monomers in *T. repens* or tobacco plants transformed with *TaMYB14* or in wild-type *T. arvense*, and it is not clear if the production of E3′G is specific for *Medicago* species or if levels of E3′G were below the detection limit in the other three species analyzed here.

CONCLUSION

Our study conclusively demonstrates that expression of *TaMYB14* specifically induces the biosynthesis and accumulation of PAs in foliar tissue of the forage legume *T. repens* through the up-regulation of late flavonoid biosynthetic pathway genes and the PA-specific transporter *TrMATE1*. Constitutive expression of *TaMYB14* in *M. sativa* and *T. repens* resulted in foliar PA accumulation of up to 1.8% dry mass, with PA complexes as high as DP6 and a monomeric unit composition comparable to that of wild-type flower PAs. It is conceivable, therefore, that the production of genetically modified *T. repens* cultivars expressing *TaMYB14* is a viable option for a bloat-safe, highly productive pastoral agriculture with reduced greenhouse gas emissions.

MATERIALS AND METHODS

Plant Materials and Histochemical Analysis

Seeds of *Trifolium arvense* (AZ2925, AZ4755, AZ1731), *Trifolium affine* (AZ925), *Trifolium repens* (DC111, 161/21), *Trifolium occidentale* (AZ4270), and *Medicago sativa* (AF2971, AF30061, AF3009) were obtained from the Margot Forde Forage Germplasm Centre. The *T. repens* line with high levels of foliar anthocyanins (cv Isabelle) was obtained from Dr. W. Williams (AgResearch). Seeds of wild-type (cv W38) and red-leafed *LeANT1*-expressing *Nicotiana tabacum* (Mathews et al., 2003) were obtained from Dr. B. Franzmayer (AgResearch).

Seeds were germinated on seed trays and plants grown in a glasshouse. Plant tissues were harvested at various developmental stages and either immediately processed for histochemical staining or frozen in liquid nitrogen and stored at −80°C for subsequent DNA, RNA, and PA isolation.

Fresh tissue samples (mature leaves, flowers, roots, immature/meristematic leaves, and trichomes) were collected from plants, and PAs were histochemically analyzed using the acidified DMACA (Sigma NZ) method essentially as described by Li et al. (1996). Briefly, tissue samples were decolorized in ethanol:acetic acid (3:1) overnight, stained with DMACA (3 mg mL^{-1}; methanol:hydrochloric acid, 1:1) for 20 min at room temperature, and destained with several washes of 70% ethanol. Meristematic leaves and trichomes were dissected from end tips of stolons using a microscope. Histochemical staining for GUS activity was performed as described (Jefferson et al., 1987).

LC-MS^n Analysis and Quantitation of PAs in Plant Tissues

To extract flavonoids for LC-MS^n analysis, leaf tissue (500 mg fresh weight) was frozen in liquid N_2, ground to a fine powder, and extracted with acetic acid:methanol (80:20, v/v) for 30 min at 4°C. Plant debris was pelleted in a microcentrifuge at 13,000 rpm for 10 min. The supernatant was removed and placed at −20°C for 30 min. Extracts were analyzed by LC-MS^n using a linear ion-trap mass spectrometer (Thermo LTQ) coupled to a Thermo Finnigan Surveyor HPLC system equipped with a Thermo photo diode array detector. Thermo Finnigan Xcalibur software (version 2.0) was used for data acquisition and processing. A 5-μL aliquot of sample was injected onto a 150- × 2.1-mm Luna C18(2) column (Phenomenex) held at a constant 25°C. The HPLC solvents used were as follows: solvent A = 0.1% formic acid in water; solvent B = 0.1% formic acid in acetonitrile. The flow rate was 200 μL min^{-1}, and the solvent gradient used is shown in Supplemental Table S4. Photo diode array data were collected across the range of 220 to 600 nm for the entire chromatogram.

The mass spectrometer was set for electrospray ionization in positive mode. The spray voltage was 4.5 kV, the capillary temperature was 275°C, and the flow rates of sheath gas, auxiliary gas, and sweep gas were set to 20, 10, and 5 arbitrary units min^{-1}, respectively. The first 4 min and last 11 min of flow from the HPLC device were diverted to waste. The MS apparatus was programmed to scan from 150 to 2,000 m/z (MS^1 scan) and then to perform data-dependent MS^3 on the most intense MS^1 ion. The isolation window for the data-dependent MS^3 method was 2 μ (nominal mass units), and fragmentation (35% relative collision energy) of the most intense ion from the MS^1 spectrum was followed by the isolation (2 μ) and fragmentation (35% relative collision energy) of the most intense ion from the MS^2 spectrum. The mass spectrometer then sequentially performed product ion scans for selected masses on the different combinations of PC and PD masses shown in Supplemental Table S5, with an isolation window for each selected m/z value of 2.0 μ and a relative collision energy of 35%. Each sample was run twice with two data collection methods to ensure that enough points were collected in the mass spectrometer across the chromatographic peaks. The first method included an MS^1 scan, the data-dependent MS^3 scans, followed by product ion scans based on selected ions (Supplemental Table S5) for the PA and PD monomers, dimers, and trimers. The second method contained an MS^1 scan, the data-dependent MS^3 scans, followed by product ion scans based on selected ions (Supplemental Table S5) for the tetramers; this method was used for *M. sativa* extracts only. The masses listed cover the different combinations of procyanidin (catechin and/or epicatechin) and prodelphinidin (gallocatechin and/or epigallocatechin) masses up to DP4.

Soluble PAs were quantified using the butanol hydrochloric acid methodology (Terrill et al., 1992). A calibration curve was set up with purified PAs isolated from *Lotus pedunculatus* (Sivakumaran et al., 2006).

Standard Methods for Isolation, PCR, Cloning, and Sequencing of DNA and RNA

Genomic DNA was isolated from fresh or frozen plant tissues using the DNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. DNA preparations were treated with RNaseH (Sigma NZ) to remove RNA from the samples. Total RNA was isolated from fresh or frozen tissues using the RNeasy Plant Mini Kit (Qiagen). Isolated total RNA was treated with RNase-free DNaseI (Roche NZ) to remove DNA from the samples following the manufacturer's instructions. The concentration and purity of DNA and RNA samples were assessed by determining the ratio of A_{260} and A_{280} using a NanoDrop ND-100 spectrophotometer (NanoDrop Technologies; Thermo Scientific).

Standard PCR was carried out in a Thermal Cycler (Applied Biosystems) using 5 ng of DNA or 1 μL of cDNA as template. The thermal cycle conditions were as follows: initial reaction at 94°C for 30 s; 35 cycles at 94°C for 30 s, 50°C to 64°C for 30 s (depending on the melting temperature of the primers), and 72°C for 1 to 2 min (1 min kb^{-1}), and a final reaction at 72°C for 10 min.

PCR products were separated by agarose gel electrophoresis and visualized by ethidium bromide staining. Bands of interest were cut out, and DNA was subsequently extracted from the gel slice using the QIAquick Gel Extraction Kit (Qiagen) following the manufacturer's instructions. Extracted PCR products were cloned into TOPO 2.1 vectors (Invitrogen NZ) and transformed into OneShot *Escherichia coli* cells by chemical transformation following the manufacturer's instructions. Bacteria were plated, selected, and grown according to the manufacturer's protocols. Bacterial plasmid DNA was extracted and purified from the bacterial cultures using the Qiagen Prep Plasmid Miniprep Kit (Qiagen). The ZAP XR vector system used the ZAP-cDNA Synthesis and Gigapack III Gold Kits (Stratagene).

Isolated plasmid DNA was sequenced using the dideoxynucleotide chain termination method (Sanger et al., 1977) using Big-Dye (version 3.1) chemistry (Applied Biosystems). Either M13 forward and reverse primers or specific gene primers were used. The products were separated on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems), and sequence data were compared with sequence information published in GenBank (National Center for Biotechnology Information) using the AlignX function of Vector NTI (Invitrogen NZ).

A variety of libraries were constructed from *T. arvense* and *T. repens* leaf, root, stolon, and floral tissues using the Uni-ZAP XR vector system with the ZAP-cDNA Synthesis and Gigapack III Gold Kits (Stratagene), amplified using the Clontech BD SMART PCR cDNA Synthesis Kit (Clontech), and cloned using TOPO 2.1 vectors (Invitrogen NZ).

Isolation of Putative R2R3-MYB Orthologs from *T. arvense*

To predict the conserved domain of legume R2R3-MYB TFs, we used the hidden Markov model method (HMMER 3.0; Eddy, 1996) based on R2R3-MYB sequences published previously (Nesi et al., 2001; Miyake et al., 2003). EST and EST contigs from available legume sequence databases (in house: *T. repens* EST CS35, *T. repens* Pastoral Genomics NR; public: *T. pratense* EST contigs, *Lotus japonicus* EST contigs, *Glycine max* contigs, *Medicago truncatula* predicted proteins, *M. sativa* glandular trichome ESTs, the Dana Farber Gene Indices databases at http://compbio.dfci.harvard.edu/tgi/plant.html, http://www.kazusa.or.jp/e/resources/index.html, and GenBank accession nos. CO511688–CO517334) were downloaded onto an in-house database system. Sequences were translated in all six reading frames and searched with the predicted R2R3-MYB domain. The total number of legume sequences predicted to have an R2R3-MYB domain identified with the HMM method was 166. A total of 130 full-length domain sequences were aligned using the HMMER alignment program (http://hmmer.wustl.edu/), and degenerate primers were designed to isolate putative R2R3-MYB TFs from *T. arvense* (Supplemental Table S1).

Total RNA was isolated from immature and mature *T. arvense* leaves as described above and reverse-transcribed into two separate cDNA libraries using the SMART cDNA Synthesis Kit (Clontech). Library A was synthesized using the SMART CDS primer IIA, which binds to the poly(A) tail; library B was synthesized using an anchored poly(T) primer and the SMART IIA oligonucleotide, which anchors to the 5′ untranslated region of the mRNA. PCR products were amplified using the PCR primer IIA and degenerate MYB primers designed to the 5′ region of the MYB domain, and amplicons were cloned into TOPO vectors as described above; 200 clones were randomly selected for sequencing and analyzed using BLAST analysis against the Database of Arabidopsis Transcription Factors (http://datf.cbi.pku.edu.cn/index.php).

Genomic DNA was PCR amplified using the primer pair MYB14ATG and MYB14TGA (Supplemental Table S1) to amplify full-length MYB14 genomic sequences from *T. arvense*, *T. repens*, *T. affine*, and *T. occidentale*.

Genetic Constructs, Plant Transformation, and Regeneration

For overexpression of *TaMYB14* or *GUS* (Jefferson et al., 1987), the genes were first cloned into the primary cloning vector pART7 between the CaMV *35S* promoter and the octopine synthase gene 3′-untranslated region as an *Eco*RI fragment (Gleave, 1992). The entire overexpression cartridge was then linearized and shuttled as a *Not*I fragment into the binary vectors pART27 (Gleave, 1992) for *N. tabacum* and pHZBAR for *M. sativa* and *T. repens* transformation. Transfer-DNAs of the genetic construct showing the orientation of the cloned gene are represented graphically in Supplemental Figure S14.

To silence *TaMYB14*, a 299-bp fragment from the 3′ end of *TaMYB14* cDNA was cloned into the hairpin RNAi vector pHANNIBAL (Helliwell and Waterhouse, 2003) and used to transform *T. arvense* cotyledons. The fragment for cloning in the sense orientation was amplified with the primer pair MYB14F1/MYB14R (Supplemental Table S1), which added *Xba*I sites to the 5′ and 3′ ends. For cloning in the antisense orientation, *Xho*I sites were added to both ends of the fragment with the primer pair MYB14F/MYB14R1 (Supplemental Table S1). The direction of cloning was determined by DNA sequence analysis. The *Not*I fragment from MYB14pHANNIBAL containing the hairpin RNA cassette was cloned into the binary vector pHZBAR (designated pHZBARSMYB; Supplemental Fig. S15) and transformed into *Agrobacterium tumefaciens* for transformation of *T. arvense*.

Leaf discs from *N. tabacum* and *M. sativa* and cotyledons from *Trifolium* species were transformed using *Agrobacterium*-mediated transformation and plant regeneration protocols as described (Blaydes, 1966; An, 1985; Horsch et al., 1985; Bingham, 1991; Shetty and McKersie, 1993; Voisey et al., 1994; Austin et al., 1995).

RT-qPCR Analysis of Gene Expression in CaMV *35S*::*TaMYB14*-Expressing *T. repens*

Total RNA from wild-type and CaMV *35S*::*TaMYB14*-expressing *T. repens* leaves was isolated using TRIzol Reagent (Invitrogen NZ) containing 2% polyvinylpyrrolidone and treated with RNase-free DNaseI (Roche NZ) to remove residual genomic DNA. DNase-treated RNA was subsequently cleaned up using the RNeasy Plant Mini Spin Kit (Qiagen) to remove enzyme, salts, and degraded DNA fragments. RNA integrity and quality were checked by denaturing gel electrophoresis, and the absence of genomic DNA was confirmed by PCR using a primer pair designed for *T. repens* 5.8S rRNA (Supplemental Table S1) prior to reverse transcription. One μg of total RNA was reverse transcribed into cDNA using the SuperScriptVILO cDNA Synthesis Kit (Invitrogen NZ) following the manufacturer's instructions. The synthesized cDNAs were then diluted 100-fold, of which 10 μL was used for subsequent qPCR analysis with a total PCR volume of 25 μL.

Gene expression profiles were quantified by qPCR amplifying the target cDNA using an iCycler (Bio-Rad Laboratories NZ) following the procedures described by Rasmussen et al. (2007). The qPCR was arrayed using SYBR Green PCR Master Mix (Applied Biosystems) at 60°C annealing temperature. Plasmids containing the target genes, ranging from 2×10^1 to 2×10^7 copies, were used for calibration of the PCR. The expression of the analyzed genes was normalized and expressed as a ratio of the target gene and actin (*TrACT*) as the housekeeping gene. Primers used for amplification were specific to *T. repens* putative *CHS*, *FLS*, *F3H*, *F3′5′H*, *DFR*, *ANS*, *ANR*, *LAR*, *MATE1*, and *MATE2* sequences as identified in our in-house *T. repens* databases; primer sequences are listed in Supplemental Table S1. Primer specificity was confirmed by direct cloning and sequencing of individual PCR amplification products. Reported are results of three technical replicates.

Sequence data from this article can be found in the GenBank/EMBL databases under the following accession numbers: TaMYB14 (*T. arvense* MYB14; JN049641), TaMYB14-2 (*T. arvense* MYB14-2; JN052773), TafMYB14-1 (*T. affine* MYB14-1; JN117921), TafMYB14-2 (*T. affine* MYB14-2; JN117922), TrMYB14-1 (*T. repens* MYB14-1; JN117923), TrMYB14-2 (*T. repens* MYB14-2; JN117924), TrMYB14-3 (*T. repens* MYB14-3; JN117925), ToMYB14-1 (*T. occidentale* MYB14-1; JN117926), ToMYB14-2 (*T. occidentale* MYB14-2; JN117927), MsMYB14-1 (*M. sativa* MYB14-1; JN120176), MsMYB14-2 (*M. sativa* MYB14-2; JN120177), MsMYB14-3 (*M. sativa* MYB14-3; JN225514), MtMYB14-1 (*M. truncatula* MYB14-1; JN157820), and MtMYB14-2 (*M. truncatula* MYB14-2; JN157821).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Amino acid alignment of TaMYB14 homologs from *Trifolium* and *Medicago* species.

Supplemental Figure S2. Amino acid alignment of R2R3-MYB TFs sharing common motifs with the PA-related VvMYBPA1 from subgroup 25 (Wilkins et al., 2009).

Supplemental Figure S3. RT-PCR analysis of *ANR*, *LAR*, and *MYB14* transcript presence in cDNA libraries isolated from various tissues of wild-

type *Trifolium* species and cultivars grown under glasshouse or tissue culture conditions.

Supplemental Figure S4. LC-MS^n composite extracted ion chromatograms of ions for catechin and gallocatechin in *T. arvense* leaf extracts.

Supplemental Figure S5. LC-MS^n composite extracted ion chromatograms of ions for PA trimers in *T. arvense* leaf extracts.

Supplemental Figure S6. LC-MS^n composite extracted ion chromatograms of ions for PA monomers in *N. tabacum* leaf extracts.

Supplemental Figure S7. LC-MS^n composite extracted ion chromatograms of ions for PA monomers in *M. sativa* leaf extracts.

Supplemental Figure S8. LC-MS^n composite extracted ion chromatograms of ions for PA trimers in *M. sativa* leaf extracts.

Supplemental Figure S9. DMACA-stained cross-section of *T. arvense* wild-type young leaf.

Supplemental Figure S10. LC-MS^n composite extracted ion chromatograms of PA monomers in *T. repens* leaf extracts.

Supplemental Figure S11. LC-MS^n composite extracted ion chromatograms of PA trimers in *T. repens* leaf extracts.

Supplemental Figure S12. LC-UV (350 nm) chromatograms of quercetin and kaempferol glycosides in *T. repens* leaf extracts.

Supplemental Figure S13. LC-MS chromatograms of the anthocyanins in a red-leafed *T. repens* line.

Supplemental Figure S14. Graphical presentation of the CaMV *35S*::*TaMYB14* gene construct used for transformation of *N. tabacum*, *M. sativa*, and *T. repens*.

Supplemental Figure S15. Graphical presentation of the RNAi *TaMYB14* gene-silencing construct used for transformation of *T. arvense*.

Supplemental Table S1. List of primer sequences used in this study.

Supplemental Table S2. PA accumulation in leaves of *M. sativa* wild-type or transgenic lines transformed with CaMV 35S::TaMYB14.

Supplemental Table S3. RT-qPCR analysis of gene expression in *T. repens*.

Supplemental Table S4. Solvent gradient used for the separation of PA oligomers.

Supplemental Table S5. Selected ion masses (m/z) for product ion scans of PA monomers, dimers, trimers, and tetramers in positive electrospray ionization mode.

ACKNOWLEDGMENTS

We thank Anar Khan (AgResearch) for providing bioinformatics support and Nick Ellison (AgResearch) for assisting with the phylogenetic analysis. We acknowledge Pastoral Genomics, Ltd., a research consortium funded by the Ministry of Science and Innovation, DairyNZ, Fonterra, Beef + Lamb New Zealand, AgResearch, and DEEResearch for providing access to their sequence databases.

Received February 12, 2012; accepted May 6, 2012; published May 7, 2012.

LITERATURE CITED

Abrahams S, Lee E, Walker AR, Tanner GJ, Larkin PJ, Ashton AR (2003) The Arabidopsis TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development. Plant J **35:** 624–636

Aerts RJ, Barry TN, McNabb WC (1999) Polyphenols and agriculture: beneficial effects of proanthocyanidins in forages. Agric Ecosyst Environ **75:** 1–12

Ahmad N, Gupta S, Mukhtar H (2000) Green tea polyphenol epigallocatechin-3-gallate differentially modulates nuclear factor kappaB in cancer cells versus normal cells. Arch Biochem Biophys **376:** 338–346

Akagi T, Ikegami A, Tsujimoto T, Kobayashi S, Sato A, Kono A, Yonemori K (2009) DkMyb4 is a Myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. Plant Physiol **151:** 2028–2045

An G (1985) High efficiency transformation of cultured tobacco cells. Plant Physiol **79:** 568–570

Austin S, Bingham ET, Mathews DE, Shahan MN, Will J, Burgess RR (1995) Production and field performance of transgenic alfalfa (*Medicago sativa* L.) expressing alpha-amylase and manganese dependent lignin peroxidase. Euphytica **85:** 381–393

Aziz N, Paiva NL, May GD, Dixon RA (2005) Profiling the transcriptome of alfalfa glandular trichomes. Planta **221:** 28–38

Bagchi D, Bagchi M, Stohs SJ, Das DK, Ray SD, Kuszynski CA, Joshi SS, Pruess HG (2000) Free radicals and grape seed proanthocyanidin extract: importance in human health and disease prevention. Toxicology **148:** 187–197

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

Beauchemin KA, McGinn SM, Martinez TF, McAllister TA (2007) Use of condensed tannin extract from quebracho trees to reduce methane emissions from cattle. J Anim Sci **85:** 1990–1996

Bingham ET (1991) Registration of the alfalfa hybrid Regen-SY germplasm for tissue culture and transformation research. Crop Sci **31:** 1098

Blaydes DF (1966) Interaction of kinetin and other inhibitors in the growth of soybean tissue. Physiol Plant **19:** 748–753

Bogs J, Jaffé FW, Takos AM, Walker AR, Robinson SP (2007) The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. Plant Physiol **143:** 1347–1361

Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell **12:** 2383–2394

Cedroni ML, Cronn RC, Adams KL, Wilkins TA, Wendel JF (2003) Evolution and expression of *MYB* genes in diploid and polyploid cotton. Plant Mol Biol **51:** 313–325

Debeaujon I, Peeters AJM, Léon-Kloosterziel KM, Koornneef M (2001) The *TRANSPARENT TESTA12* 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

Deluc L, Barrieu F, Marchive C, Lauvergeat V, Decendit A, Richard T, Carde J-P, Mérillon J-M, Hamdi S (2006) Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. Plant Physiol **140:** 499–511

Deluc L, Bogs J, Walker AR, Ferrier T, Decendit A, Merillon J-M, Robinson SP, Barrieu F (2008) The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. Plant Physiol **147:** 2041–2053

Dixon RA, Xie D-Y, Sharma SB (2005) Proanthocyanidins: a final frontier in flavonoid research? New Phytol **165:** 9–28

Douglas GB, Stienezen M, Waghorn GC, Foote AG, Purchas RW (1999) Effect of condensed tannins in birdsfoot trefoil (*Lotus corniculatus*) and sulla (*Hedysarum coronarium*) on body weight, carcass fat depth, and wool growth of lambs in New Zealand. N Z J Agric **42:** 55–64

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in *Arabidopsis*. Trends Plant Sci **15:** 573–581

Dufresne CJ, Farnworth ER (2001) A review of latest research findings on the health promotion properties of tea. J Nutr Biochem **12:** 404–421

Eddy SR (1996) Hidden Markov models. Curr Opin Struct Biol **6:** 361–365

Ellison NW, Liston A, Steiner JJ, Williams WM, Taylor NL (2006) Molecular phylogenetics of the clover genus (Trifolium-Leguminosae). Mol Phylogenet Evol **39:** 688–705

Fay MF, Dale PJ (1993) Condensed tannins in *Trifolium* species and their significance for taxonomy and plant breeding. Genet Resour Crop Evol **40:** 7–13

Foo LY, Lu Y, Molan AL, Woodfield DR, McNabb WC (2000) The phenols and prodelphinidins of white clover flowers. Phytochemistry **54:** 539–548

Fraser K, Harrison SJ, Lane GA, Otter DE, Hemar Y, Quek S-Y, Rasmussen S (January 31, 2012) HPLC MS/MS profiling of proanthocyanidins in teas: a comparative study. J Food Compos Anal http://dx.doi.org/10.1016/j.jfca.2012.01.004

Gleave AP (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. Plant Mol Biol **20:** 1203–1207

Gonzalez de Colmenares N, Ramirez-Martinez JR, Aldana-Mendez B (1998) Isolation, characterization and determination of biological activity of coffee proanthocyanidins. J Sci Food Agric **77:** 368–372

Goplen BP, Howarth RE, Sarkar SK, Lesins K (1980) A search for condensed tannins in annual and perennial species of *Medicago*, *Trigonella*, and *Onobrychis*. Crop Sci **20:** 801–804

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol **27:** 221–224

Grotewold E, Sainz MB, Tagliani L, Hernandez JM, Bowen B, Chandler VL (2000) Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R. Proc Natl Acad Sci USA **97:** 13579–13584

Hayasaka Y, Waters EJ, Cheynier V, Herderich MJ, Vidal S (2003) Characterization of proanthocyanidins in grape seeds using electrospray mass spectrometry. Rapid Commun Mass Spectrom **17:** 9–16

Helliwell C, Waterhouse P (2003) Constructs and methods for high-throughput gene silencing in plants. Methods **30:** 289–295

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V (2011) Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. J Exp Bot **62:** 2465–2483

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell **7:** 1071–1083

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. Science **227:** 1229–1231

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

Jiang C, Gu X, Peterson T (2004) Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of *Arabidopsis* and *Oryza sativa* L. ssp. *indica*. Genome Biol **5:** R46.1–11

Jones WT, Broadhurst RB, Lyttleton JW (1976) The condensed tannins of pasture legume species. Phytochemistry **15:** 1407–1409

Katoh K, Asdimenos G, Toh H (2009) Multiple alignment of DNA sequences with MAFFT. *In* D Prosada, ed, Bioinformatics for DNA Sequence Analysis. Methods in Molecular Biology 537. Humana Press, New York, pp 39–64

Kingston-Smith AH, Thomas HM (2003) Strategies of plant breeding for improved rumen function. Ann Appl Biol **142:** 13–24

Koupai-Abyazani MR, McCallum J, Muir AD, Lees GL, Bohm BA, Towers GHN, Gruber MY (1993) Purification and characterization of a proanthocyanidin polymer from seed of alfalfa (*Medicago sativa* cv. Beaver). J Agric Food Chem **41:** 565–569

Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol **57:** 405–430

Li HJ, Deinzer ML (2007) Tandem mass spectrometry for sequencing proanthocyanidins. Anal Chem **79:** 1739–1748

Li Y, Tanner G, 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

Marinova K, Pourcel L, Weder B, Schwarz M, Barron D, Routaboul JM, Debeaujon I, Klein M (2007) The *Arabidopsis* MATE transporter TT12 acts as a vacuolar flavonoid/H^+-antiporter active in proanthocyanidin-accumulating cells of the seed coat. Plant Cell **19:** 2023–2038

Marshall DR, Broué P, Grace J, Munday J (1980) Tannins in pasture legumes: the annual and perennial *Medicago* species. Aust J Exp Agric **21:** 47–50

Mathews H, Clendennen SK, Caldwell CG, Liu XL, Connors K, Matheis N, Schuster DK, Menasco DJ, Wagoner W, Lightner J, et al (2003) Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport. Plant Cell **15:** 1689–1703

McMahon LR, McAllister TA, Berg BP, Majak W, Acharya SN, Popp JD, Coulman BE, Wang Y, 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

Mellway RD, Tran LT, Prouse MB, Campbell MM, Constabel CP (2009) The wound-, pathogen-, and ultraviolet B-responsive MYB134 gene encodes an R2R3 MYB transcription factor that regulates proanthocyanidin synthesis in poplar. Plant Physiol **150:** 924–941

Miyake K, Ito T, Senda M, Ishikawa R, Harada T, Niizeki M, Akada S (2003) Isolation of a subfamily of genes for R2R3-MYB transcription factors showing up-regulated expression under nitrogen nutrient-limited conditions. Plant Mol Biol **53:** 237–245

Monagas M, Hernández-Ledesma B, Gómez-Cordovés C, Bartolomé B (2006) Commercial dietary ingredients from *Vitis vinifera* L. leaves and grape skins: antioxidant and chemical characterization. J Agric Food Chem **54:** 319–327

Nesi N, Jond C, Debeaujon I, Caboche M, 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

Pang Y, Peel GJ, Sharma SB, Tang Y, Dixon RA (2008) A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. Proc Natl Acad Sci USA **105:** 14210–14215

Pang Y, Wenger JP, Saathoff K, Peel GJ, Wen J, Huhman D, Allen SN, Tang Y, Cheng X, Tadege M, et al (2009) A WD40 repeat protein from *Medicago truncatula* is necessary for tissue-specific anthocyanin and proanthocyanidin biosynthesis but not for trichome development. Plant Physiol **151:** 1114–1129

Peel GJ, Pang Y, Modolo LV, Dixon RA (2009) The LAP1 MYB transcription factor orchestrates anthocyanidin biosynthesis and glycosylation in *Medicago*. Plant J **59:** 136–149

Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R (1999) Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. Plant Cell **11:** 1433–1444

Rasmussen S (2009) Biosynthesis and manipulation of flavonoids in forage legumes. *In* K Gould, KM Davies, CS Winefield, eds, Anthocyanins: Biosynthesis, Functions and Applications. Springer-Verlag, New York, pp 257–272

Rasmussen S, Parsons AJ, Bassett S, Christensen MJ, Hume DE, Johnson LJ, Johnson RD, Simpson WR, Stacke C, Voisey CR, et al (2007) High nitrogen supply and carbohydrate content reduce fungal endophyte and alkaloid concentration in *Lolium perenne*. New Phytol **173:** 787–797

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

Reed JD (1995) Nutritional toxicology of tannins and related polyphenols in forage legumes. J Anim Sci **73:** 1516–1528

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA **74:** 5463–5467

Shetty K, McKersie BD (1993) Proline, thioproline and potassium mediated stimulation of somatic embryogenesis in alfalfa (*Medicago sativa* L.). Plant Sci **88:** 185–193

Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM (1995) Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. Plant J **8:** 659–671

Sivakumaran S, Rumball W, Lane GA, Fraser K, Foo LY, Yu M, Meagher LP (2006) Variation of proanthocyanidins in Lotus species. J Chem Ecol **32:** 1797–1816

Smith P, Martino D, Cai Z, Gwary D, Janzen H, Kumar P, McCarl B, Ogle S, O'Mara F, Rice C, et al (2008) Greenhouse gas mitigation in agriculture. Philos Trans R Soc Lond B Biol Sci **363:** 789–813

Stracke R, Werber M, Weisshaar B (2001) The *R2R3-MYB* gene family in *Arabidopsis thaliana*. Curr Opin Plant Biol **4:** 447–456

Takos AM, Jaffé FW, Jacob SR, Bogs J, Robinson SP, Walker AR (2006) Light-induced expression of a *MYB* gene regulates anthocyanin biosynthesis in red apples. Plant Physiol **142:** 1216–1232

Tanner GJ, Francki KT, Abrahams S, Watson JM, Larkin PJ, Ashton AR (2003) Proanthocyanidin biosynthesis in plants: purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. J Biol Chem **278:** 31647–31656

Terrier N, Torregrosa L, Ageorges A, Vialet S, Verriès C, Cheynier V, Romieu C (2009) Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. Plant Physiol **149:** 1028–1041

Terrill TH, Rowan AM, Douglas GB, Barry TN (1992) Determination of extractable and bound condensed tannin concentrations in forage plants, protein concentrate meals and cereal grains. J Sci Food Agric **58:** 321–329

Verdaguer B, de Kochko A, Beachy RN, Fauquet C (1996) Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. Plant Mol Biol **31:** 1129–1139

Verdier J, Zhao J, Torres-Jerez I, Ge S, Liu C, He X, Mysore KS, Dixon RA, Udvardi MK (2012) MtPAR MYB transcription factor acts as an on switch for proanthocyanidin biosynthesis in *Medicago truncatula*. Proc Natl Acad Sci USA **109:** 1766–1771

Voisey CR, White DWR, Dudas B, Appleby RD, Ealing PM, Scott AG (1994) *Agrobacterium*-mediated transformation of white clover using direct shoot organogenesis. Plant Cell Rep **13:** 309–314

Weber HA, Hodges AE, Guthrie JR, O'Brien BM, Robaugh D, Clark AP, Harris RK, Algaier JW, Smith CS (2007) Comparison of proanthocyanidins in commercial antioxidants: grape seed and pine bark extracts. J Agric Food Chem **55:** 148–156

Wei YL, Li JN, Lu J, Tang ZL, Pu DC, Chai YR (2007) Molecular cloning of *Brassica napus* TRANSPARENT TESTA 2 gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis. Mol Biol Rep **34:** 105–120

Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM (2009) Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. Plant Physiol **149:** 981–993

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

Woodward SL, Waghorn GC, Ulyatt MJ, Lassey KR (2001) Early indications that feeding Lotus will reduce methane emissions from ruminants. Proc N Z Soc Anim Prod **61:** 23–26

Xie D-Y, Sharma SB, Wright E, Wang Z-Y, Dixon RA (2006) Metabolic engineering of proanthocyanidins through co-expression of anthocyanidin reductase and the PAP1 MYB transcription factor. Plant J **45:** 895–907

Yoshida K, Iwasaka R, Kaneko T, Sato S, Tabata S, Sakuta M (2008) Functional differentiation of *Lotus japonicus* TT2s, R2R3-MYB transcription factors comprising a multigene family. Plant Cell Physiol **49:** 157–169

Yoshida K, Kume N, Nakaya Y, Yamagami A, Nakano T, Sakuta M (2010) Comparative analysis of the triplicate proathocyanidin regulators in Lotus japonicus. Plant Cell Physiol **51:** 912–922

Zhao J, Dixon RA (2009) MATE transporters facilitate vacuolar uptake of epicatechin 3′-*O*-glucoside for proanthocyanidin biosynthesis in *Medicago truncatula* and *Arabidopsis*. Plant Cell **21:** 2323–2340

Zhao J, Huhman D, Shadle G, He X-Z, Sumner LW, Tang Y, Dixon RA (2011) MATE2 mediates vacuolar sequestration of flavonoid glycosides and glycoside malonates in *Medicago truncatula*. Plant Cell **23:** 1536–1555

Zhao J, Pang Y, Dixon RA (2010) The mysteries of proanthocyanidin transport and polymerization. Plant Physiol **153:** 437–443

Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF (2004) Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. Plant J **40:** 22–34