The Plant Cell, Vol. 13, 2099–2114, September 2001, www.plantcell.org © 2001 American Society of Plant Biologists

The Arabidopsis *TT2* Gene Encodes an R2R3 MYB Domain Protein That Acts as a Key Determinant for Proanthocyanidin Accumulation in Developing Seed

Nathalie Nesi, Clarisse Jond, Isabelle Debeaujon, Michel Caboche, and Loïc Lepiniec[1]

Laboratoire de Biologie des Semences, Unité Mixte de Recherche, Institut National de la Recherche Agronomique, Centre de Versailles, Route de Saint-Cyr, 78026 Versailles Cedex, France

In Arabidopsis, proanthocyanidins specifically accumulate in the endothelium during early seed development. At least three *TRANSPARENT TESTA* (*TT*) genes, *TT2*, *TT8*, and *TTG1*, are necessary for the normal expression of several flavonoid structural genes in immature seed, such as *DIHYDROFLAVONOL-4-REDUCTASE* and *BANYULS* (*BAN*). *TT8* and *TTG1* were characterized recently and found to code for a basic helix-loop-helix domain transcription factor and a WD-repeat–containing protein, respectively. Here the molecular cloning of the *TT2* gene was achieved by T-DNA tagging. *TT2* encoded an R2R3 MYB domain protein with high similarity to the rice OsMYB3 protein and the maize COLORLESS1 factor. A TT2–green fluorescent protein fusion protein was located mostly in the nucleus, in agreement with the regulatory function of the native TT2 protein. *TT2* expression was restricted to the seed during early embryogenesis, consistent with *BAN* expression and the proanthocyanidin deposition profile. Finally, in gain-of-function experiments, TT2 was able to induce ectopic expression of *BAN* in young seedlings and roots in the presence of a functional TT8 protein. Therefore, our results strongly suggest that stringent spatial and temporal *BAN* expression, and thus proanthocyanidin accumulation, are determined at least partially by TT2.

INTRODUCTION

Flavonoids are secondary metabolites that are unique to higher plants. They are well known for the red, purple, and brown pigmentation they give to flowers, fruit, and seed. Flavonoids fulfill numerous physiological functions during plant life and also serve as beneficial micronutrients in human and animal diets (reviewed by Koes et al., 1994; Shirley, 1996; Mol et al., 1998; Harborne and Williams, 2000). Arabidopsis contains three major classes of flavonoids: the anthocyanins (red to purple pigments), the flavonols (colorless to pale yellow pigments), and the proanthocyanidins (colorless pigments that turn to brown), which also are known as condensed tannins (Figure 1). Anthocyanins and flavonols are synthesized in vegetative parts, whereas flavonols and proanthocyanidins accumulate in seed (Chapple et al., 1994).

As shown in Figure 1, the different flavonoid subpathways share common initial biosynthetic steps, including synthesis of naringenin chalcone by chalcone synthase (CHS), conversion of naringenin chalcone to naringenin by chalcone isomerase (CHI), and subsequent hydroxylations of naringenin by flavanone 3-hydroxylase (F3H) and flavonoid 3′-hydroxylase (F3′H). Genetic and molecular study of flavonoid biosynthesis in Arabidopsis has revealed that the CHS, CHI, F3H, and F3′H enzymes are encoded by the *TRANSPARENT TESTA4* (*TT4*), *TT5*, *TT6*, and *TT7* genes, respectively (Shirley et al., 1992, 1995; Wisman et al., 1998; Schoenbohm et al., 2000). Then, an NADPH-dependent dihydroflavonol reductase (DFR), encoded by the *TT3* gene in Arabidopsis (Shirley et al., 1992), leads to the production of flavan-3,4-diols (leucoanthocyanidins), which are the last common intermediates in anthocyanin and proanthocyanidin biosynthesis. Leucoanthocyanidins are then converted to catechins by an NADPH-dependent leucoanthocyanidin reductase (LAR) (Tanner and Kristiansen, 1993), which is probably encoded by the *BANYULS* (*BAN*) gene in Arabidopsis (Devic et al., 1999). Recently, Debeaujon and coworkers (2001) showed that the Arabidopsis *TT12* gene codes for a putative transporter that is likely to participate in vacuolar sequestration of proanthocyanidin precursors. Finally, hypothetical condensing enzymes perform dimerization between leucoanthocyanidin and catechin monomers followed by sequential addition of leucoanthocyanidin-derived single units to form proanthocyanidin polymers (reviewed by Stafford, 1989; Jende-Strid, 1993). With the exception of the *FLAVONOL SYNTHASE* (*FLS*) gene, all flavonoid biosynthetic enzymes known to date are encoded by single-copy genes in Arabidopsis.

[1] To whom correspondence should be addressed. E-mail lepiniec@versailles.inra.fr; fax 33-01-30-83-30-99.

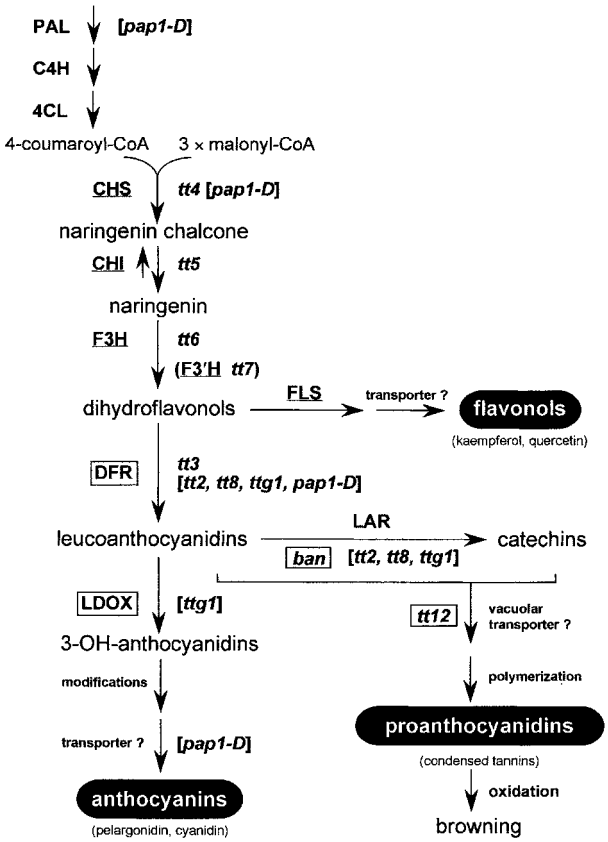

Figure 1. Scheme of the Flavonoid Biosynthetic Pathway of Arabidopsis Leading to the Synthesis of Anthocyanins, Flavonols, and Proanthocyanidins.

Enzymes are indicated in uppercase letters, with the corresponding genetic loci given in lowercase italic letters. Mutants for regulatory genes are indicated in brackets. Enzymes encoded by flavonoid EBGs are underlined, and those encoded by flavonoid LBGs are boxed. Abbreviations are as follows: *ban*, *banyuls*; CHI, chalcone isomerase; CHS, chalcone synthase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; DFR, dihydroflavonol 4-reductase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; FLS, flavonol synthase; LAR, leucoanthocyanidin reductase; LDOX, leucoanthocyanidin dioxygenase; PAL, phenylalanine ammonia-lyase; *pap1-D*, *production of anthocyanin pigment1-dominant*; *tt*, *transparent testa*; *ttg1*, *transparent testa glabrous1*.

Proanthocyanidins have been shown to be major determinants of seed agronomic and nutritional properties (Winkel-Shirley, 1998; Debeaujon et al., 2000). In immature seed of Arabidopsis, proanthocyanidins are located specifically in the innermost cell layer of the seed coat, called the endothelium (Devic et al., 1999). They begin to accumulate as colorless compounds from the very early stages of embryo development and turn brown upon oxidative reactions, which occur after the completion of embryogenesis, with the onset of seed desiccation (Debeaujon et al., 2001). Expression of the *BAN* gene also is restricted to the endothelium during early seed development (Devic et al., 1999), thus demonstrating a correlation between the accumulation of structural gene transcripts and proanthocyanidin deposition. The stringent spatial and temporal specificity of proanthocyanidin accumulation in Arabidopsis seed coat raises the question of the molecular basis of the regulatory mechanisms involved in pigment deposition during seed development.

Numerous studies have suggested that the specific accumulation of the different flavonoid end products in plants is tightly controlled by independent groups of regulators (reviewed by Mol et al., 1998). Extensive genetic analyses with maize revealed that anthocyanin production requires combinatorial interactions between a member of the RED1 (R1)/BOOSTER1 (B1) family, encoding basic helix-loop-helix (bHLH) DNA binding domain proteins, and a member of the COLORLESS1 (C1)/PURPLE LEAF (Pl) family, encoding MYB transcription factors (Cone et al., 1986; Paz-Ares et al., 1987; Ludwig et al., 1989). Each family exhibits allelic diversity and multiple paralogs that control pigmentation in a specific organ or tissue (Ludwig and Wessler, 1990). For instance, alleles at the standard *R1* locus mainly condition the color of the kernel in addition to the pigmentation of embryo and plant tissues, whereas members of the duplicated *B1* locus mostly affect anthocyanin accumulation in vegetative parts. In addition, the interaction between an R1- and a C1-related protein is necessary and sufficient to activate the entire anthocyanin pathway in maize (Goff et al., 1990). In contrast, the MYB factor P regulates a subset of flavonoid structural genes by itself, leading to phlobaphene biosynthesis in maize floral organs (Grotewold et al., 1994).

Flavonoid biosynthesis in dicot plants is controlled in at least two separate subsets, with the flavonoid early biosynthetic genes (EBGs) on the one hand and the flavonoid late biosynthetic genes (LBGs) on the other hand (Martin et al., 1991; Kubasek et al., 1992; Quattrocchio et al., 1993; Nesi et al., 2000) (Figure 1). In petunia, the *ANTHOCYANIN1* (*AN1*), *AN2*, and *AN11* genes regulate, at least partially, the expression of flavonoid LBGs in flower and encode a bHLH protein, an R2R3 MYB factor, and a WD-repeat regulator, respectively (Quattrocchio et al., 1993, 1999; de Vetten et al., 1997; Spelt et al., 2000). AN1 and AN11 also are required for anthocyanin accumulation in almost all pigmented tissues, whereas *AN2* expression apparently is restricted to the corolla limb. Paralogs of AN2 have been identified and apparently control pigmentation in different organs (Spelt et al., 2000).

In Arabidopsis, we demonstrated previously that functional TTG1 (TRANSPARENT TESTA GLABROUS1), TT8, and TT2 proteins are required for the normal expression of at least two flavonoid LBGs, *DFR* and *BAN*, during seed formation (Nesi et al., 2000) (Figure 1). *TTG1* was shown to encode a WD-repeat–containing protein and probably represents an ortholog of the petunia AN11 product (Walker et

al., 1999). The molecular cloning of *TT8* revealed that it encodes a bHLH domain protein (Nesi et al., 2000). At least two lines of evidence support the idea that TT8 belongs to a novel class of flavonoid bHLH-containing regulators: (1) TT8 is more closely related to the petunia *AN1* gene (Spelt et al., 2000) than to the other bHLH proteins involved in flavonoid production, and (2) overexpression of *TT8* is not sufficient to functionally complement the Arabidopsis *ttg1* mutation (N. Nesi, unpublished results), unlike the maize *R* gene (Lloyd et al., 1992). Because *TT8* and *TTG1* also are expressed in plant vegetative parts (Walker et al., 1999; Nesi et al., 2000), where neither *BAN* transcripts nor proanthocyanidins have been detected, these two regulatory genes are not the determining factors for the specific accumulation of *BAN* mRNA and tannins within the seed.

Finally, information on the protein encoded by the *TT2* gene remains unknown. Recently, Borevitz and co-workers reported the characterization of the *PRODUCTION OF ANTHOCYANIN PIGMENT1* (*PAP1*) locus (Borevitz et al., 2000), the overexpression of which deregulates anthocyanin accumulation in Arabidopsis vegetative parts concomitant with a broad transcriptional activation of the overall phenylpropanoid pathway (Figure 1). The *PAP1* locus was shown to code for the Arabidopsis MYB75 factor that displays high sequence similarity to the petunia AN2 protein (Quattrocchio et al., 1999). However, plants carrying the *PAP1* antisense construct produce wild-type brown seed, suggesting that *PAP1* is different from *TT2*. Together, these previous results prompted us to search for the molecular nature and function of the *TT2* gene product.

Here we report on our molecular and functional analyses of the *TT2* gene from Arabidopsis. Cloning of *TT2* revealed that it specifies an R2R3 MYB domain protein. The expression pattern of the *TT2* gene was shown to be fully consistent with the involvement of the protein in proanthocyanidin biosynthesis, being restricted to the seed during early stages of embryogenesis. In addition, TT2 was necessary for the expression of flavonoid LBGs, and its ectopic expression was sufficient to activate the transcription of *BAN* in young seedlings and roots. Therefore, we conclude that TT2 acts as a major factor for the determination of the *BAN* mRNA expression pattern and thus for tannin accumulation in seed.

RESULTS

Isolation of *tt2* Mutants

The dro55 mutant was isolated from the Versailles T-DNA mutagenized population. The mutation gave the seed a golden yellow color (Figure 2A), which segregated as a recessive, monogenic, nuclear, and maternal trait. Indeed, in progeny from a cross between dro55 and wild-type plants, we observed that (1) all F1 seed had the phenotype conferred

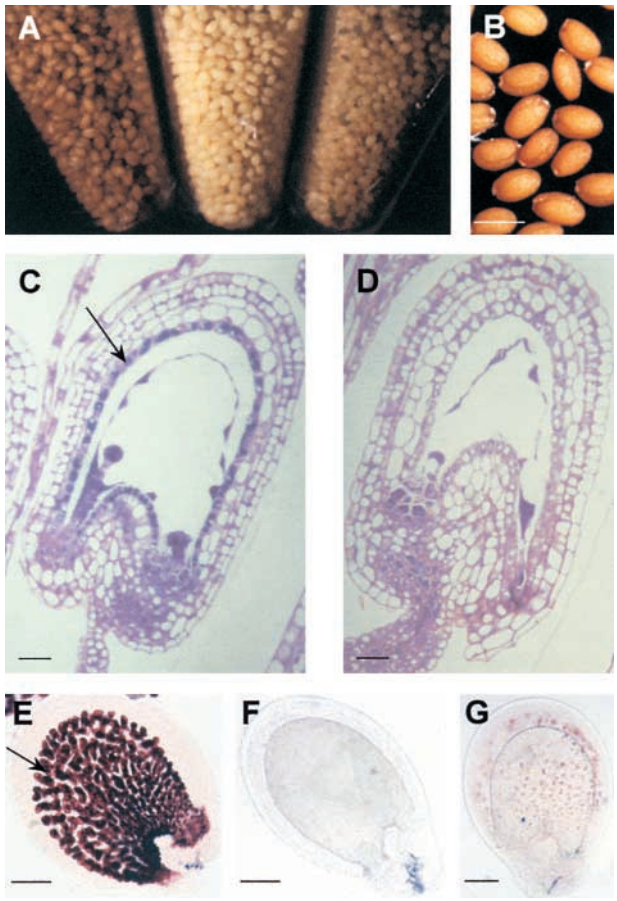

Figure 2. Seed Phenotypes.

(A) Seed from wild-type plants and from *tt2-3* and *tt2-2* mutants (left to right). Seed of the *tt2-3* mutant are golden yellow, whereas the *tt2-2* mutation gives the seed a buff color.

(B) T2 seed from a *tt2-3* homozygous mutant transformed with the *TT2* cDNA under the control of a double enhanced 35S promoter. The brown pigmentation of transgenic seed demonstrates the phenotypic complementation of the *tt2-3* mutation by overexpression of the *TT2* cDNA. Bar = 450 μm.

(C) and **(D)** Localization of phenolic compounds in the testa of immature seed from a wild-type plant **(C)** and the *tt2-3* mutant **(D)** after staining with toluidine blue. The flavonoids, stained blue, form granules that are localized in vacuoles of the endothelium cells in the wild-type seed coat (arrow in **[C]**). **(D)** shows the absence of blue granules in the *tt2-3* endothelium. Bars = 20 μm.

(E) to **(G)** Detection of proanthocyanidins and their precursors in immature seed treated with vanillin HCl. The vanillin test stains the proanthocyanidins and their precursors (leucoanthocyanidins and catechins) red in the endothelium of the wild type (arrow in **[E]**). The striking difference is the complete absence of these compounds in *tt2-3* **(F)**, whereas a slight red color is observed in the *tt2-2* seed coat **(G)**. Bars = 50 μm.

by the maternal genotype, and (2) 31 of 119 F2 plants produced seed with the mutant phenotype. Vegetative parts of dro55 seemed to synthesize normal anthocyanins, as revealed by the purple color of aging rosette leaves (data not shown), and the plants showed no obvious additional defects under our growth conditions. These observations suggested that the effect of the mutation was restricted to seed coat pigmentation. The dro55 line was shown to be allelic to the *tt2-1* mutant identified by Koornneef (1981, 1990) and was named *tt2-3*, as part of an allelic series isolated during this project (see below). Cytological analysis of immature seed showed that the *tt2-3* seed coat consisted of the five characteristic cell layers found in the wild type (Figures 2C and 2D). However, the granules, which accumulate in the wild-type endothelium layer and are stained blue-green after toluidine blue treatment (Figure 2C), were completely absent in *tt2-3* (Figure 2D). These granules accumulate in the endothelium of wild-type seed starting with the earliest stages of embryogenesis and contain flavonoid compounds (Devic et al., 1999).

Two other *tt2* alleles were identified during the course of this work (Table 1). The 3C line was obtained during a screening for reduced seed dormancy mutants on a γ-ray–mutagenized population (Léon-Kloosterziel et al., 1996). The JSM9-3 line came from the Sendai ethyl methanesulfonate–mutagenized collection. The 3C and JSM9-3 mutants were found to be allelic to *tt2-1* and were named *tt2-2* and *tt2-4*, respectively. Surprisingly, the seed of *tt2-2* harbored a buff-colored seed coat (Figure 2A), which clearly differed from the golden yellow seed of the three other *tt2* mutant alleles. Using a vanillin test, which stains proanthocyanidins as well as the catechin and leucoanthocyanidin precursors dark red in wild-type seed (Aastrup et al., 1984) (Figure 2E), we confirmed the presence of a very small amount of these products in immature seed of *tt2-2* (Figure 2G). Conversely, no red stain appeared in *tt2-3* seed (Figure 2F). Thus, the *tt2-2* mutation leads to an effective but severely reduced accumulation of condensed tannins, which may account for the buff color of the mutant mature seed.

Cloning of the Arabidopsis *TT2* Gene

By outcrossing to wild-type Arabidopsis plants, we detected tight genetic linkage between the *tt2-3* phenotype and one active kanamycin resistance locus. Indeed, F2 progeny from the backcross segregated in a 3:1 ratio for kanamycin-resistant to kanamycin-sensitive plants, as expected for a monogenic dominant trait. In addition, homozygous *tt2-3/tt2-3* plants were all kanamycin resistant (n = 1468 F3 plantlets generated from 31 independent homozygous F2 plants), and all of the homozygous *TT2/TT2* wild-type plants were kanamycin sensitive. Furthermore, DNA gel blot analyses of the *tt2-3* mutant with different T-DNA–specific probes revealed the presence of a single full-length T-DNA copy (data not shown). Together, the genetic and molecular data strongly supported the idea that the putative *TT2* gene was tagged by one complete T-DNA unit, thus allowing its molecular cloning.

The *tt2-3* genomic regions surrounding both T-DNA borders were isolated by means of polymerase chain reaction (PCR) walking (Devic et al., 1997). Two genomic fragments of ~1150 and 1500 bp were isolated from T-DNA right and left borders, respectively, sequenced, and used to design primers TT2-5′ and TT2-3′ from plant genomic DNA (Figure 3). The amplification of the corresponding wild-type genomic sequence resulted in a 2700-bp contig, which spanned the entire putative *TT2* open reading frame (ORF) (Figure 3). Submission to databases revealed that the putative *TT2* gene was located on the MOK9 genomic clone (Figure 3), which was found in the Arabidopsis genomic P1 library (Liu et al., 1995). This P1 clone was sequenced, annotated, and mapped to chromosome 5 (Nakamura et al., 1998), which was consistent with the genetic positioning of *tt2* (Koornneef et al., 1983).

Using the 2700-bp genomic fragment as a probe, we screened several Arabidopsis silique cDNA libraries but failed to isolate any cDNA clones. In addition, no expressed sequence tag corresponding to the putative *TT2* ORF was

Table 1. Identification and Characterization of Four *tt2* Mutant Alleles

Allele	Seed Stock[a]	Reference	Mutagen	Phenotype[b]	Mutation in *TT2*	Predicted Effect
tt2-1	N83 (L*er*)	Koornneef (1981, 1990)	x-ray	Yellow seed, wild-type anthocyanin levels in vegetative tissues	Deletion of genomic DNA and chromosomal rearrangement	Lack of transcript
tt2-2	3C (L*er*)	Léon-Kloosterziel et al. (1996)	γ-ray	Buff-colored seed	A-525[c]→T	K-127[d]→stop codon
tt2-3	dro55 (Ws-2)	This work	T-DNA	Similar to *tt2-1*	T-DNA insertion	Lack of transcript
tt2-4	JSM9-3 (L*er*)	Sendai collection, N. Goto	EMS[e]	Similar to *tt2-1*	G-269[c]→A	G-66[d]→R

[a] The parental ecotype is indicated within parentheses. L*er*, Landsberg *erecta*; Ws-2, Wassilewskija-2.
[b] Phenotype refers only to flavonoid contents in different parts of the plants. *tt2-1* has clavata siliques. The three other *tt2* mutant alleles show no additional phenotype under our growth conditions.
[c] Nucleotides and [d] amino acid residues are numbered from the predicted translation start site.
[e] EMS, ethyl methanesulfonate.

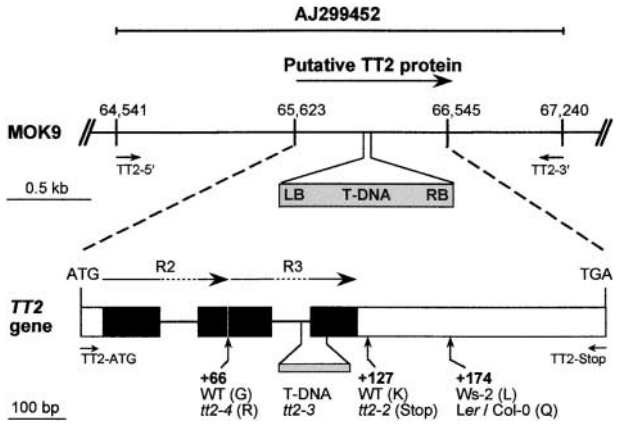

Figure 3. Isolation of the *TT2* Gene.

Top, physical map of the TT2 genomic region on the P1 clone MOK9 (chromosome 5). The arrow indicates the orientation of the putative TT2 protein. The T-DNA insertion is shown in the gray box. Numbers above the scheme indicate nucleotide positions on the MOK9 sequence (in base pairs). Bottom, exon-intron structure of the Arabidopsis *TT2* gene. Boxes represent exons and lines represent introns. The region encoding the R2R3 MYB DNA binding domain is shown by black boxes. Numbering of amino acid residues is given from the predicted translation start codon and is shown in boldface below the scheme. The T-DNA copy (7 kb) in *tt2-3* is inserted between intron 2 and exon 3 and causes a 45-bp deletion. Mutations in *tt2-2* and *tt2-4* are localized, with amino acid substitutions indicated in parentheses. Position 174 shows an amino acid polymorphism between Arabidopsis ecotypes. The different primers and probes used for molecular analyses are noted. Col-0, Columbia-0; LB, left border; L*er*, Landsberg *erecta*; RB, right border; R2, repeat 2; R3, repeat 3; Ws-2, Wassilewskija-2; WT, wild type.

available in the databases. Therefore, a cDNA clone was generated using reverse transcription (RT)-PCR with RNA from immature siliques. For this purpose, we used primers TT2-ATG and TT2-Stop (Figure 3), which were designed from the predicted translation start codon and the predicted translation stop codon, respectively, according to the annotation of MOK9. Determination and comparison of the 777-bp amplified cDNA with genomic DNA sequences showed that the *TT2* gene contains three exons interrupted by two introns (Figure 3). The T-DNA insertion is located at the junction between intron 2 and exon 3 in *tt2-3* and caused a 45-bp deletion of genomic DNA.

Functional Complementation of the *tt2-3* Mutation

To correlate the *tt2* phenotype with the disruption of the putative *TT2* gene, we conducted a functional complementation of homozygous *tt2-3* plants. The cDNA was fused to the double enhanced constitutive 35S promoter of *Cauliflower mosaic virus* (CaMV) (hereafter referred to as the 70S-*TT2* construct), cloned into a binary vector containing hygromycin resistance, and used to transform homozygous *tt2-3* plants. Among 33 independent primary transformants (T1) selected for hygromycin resistance, 31 exhibited a complete reversion of the mutant phenotype in their T2 progeny (Figure 2B). Further genetic analyses revealed that T2 plants were all kanamycin resistant, which confirmed that they harbored the original T-DNA, and that they segregated for the hygromycin resistance marker as well as for seed color phenotype. Nevertheless, all hygromycin-resistant plants produced brown seed in their progeny. These results strongly suggested that the *tt2-3* mutation was due to the disruption of the *TT2* gene by T-DNA insertion and provided compelling evidence that the cloned ORF was functional.

TT2 Sequence Features

The *TT2* gene encodes a protein of 258 amino acid residues (Figure 4A) with a predicted molecular mass of 29.7 kD and a calculated pI of 8.91. Analyses of the primary structure revealed that TT2 contains a region with high similarity to the DNA binding domain found in the vertebrate or insect proto-oncogene MYB proteins (Klempnauer et al., 1982). MYB DNA binding domains consist of ~50 amino acid residues, form a helix-helix-turn-helix motif, and are involved in transcriptional regulation mechanisms (Frampton et al., 1991; Ogata et al., 1992). The TT2 protein contains two imperfect MYB repeats within its N-terminal region (Figure 3) that correspond approximately to the R2 and R3 MYB repeats of the human c-MYB prototype (Figure 4B). The comparison of amino acid sequences of the R2R3 domain from different MYB-related proteins showed that TT2 shares general structural features that are evolutionarily conserved among this widespread class of DNA binding proteins (Figure 4B). The highly conserved W residues, which are thought to be involved in the folding of the DNA binding domain, are present in TT2, except that the first W in the R3 repeat is substituted by an I residue. In addition, the amino acid residues in the R3 MYB repeat of maize C1, which are supposed to drive specific interaction with R-like proteins (Grotewold et al., 2000), were found in TT2. Finally, the short linker sequence between the two MYB repeats also displays amino acid conservation.

Distance analysis supports the idea that TT2 is more closely related to plant MYB proteins than to the animal c-MYB (Figure 4C). Interestingly, within the large family of plant R2R3 MYB domain proteins, the rice OsMYB3 gene product shares the highest similarity with TT2. The similarity between TT2 and OsMYB3 is particularly obvious within the R2R3 DNA binding domain, where amino acid identity rates are 82% between TT2 and OsMYB3, 74% between TT2 and the maize C1 protein, 67% between TT2 and maize P, 66% between TT2 and the petunia AN2 factor, and 65% between

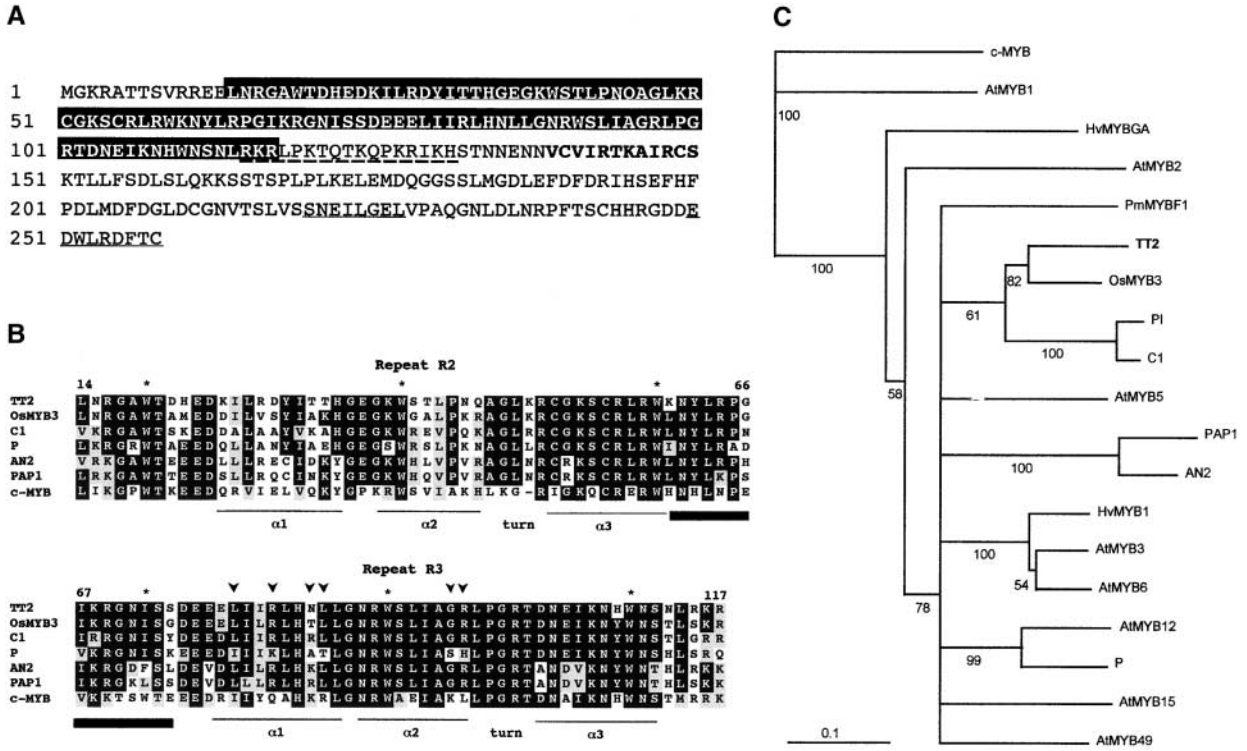

Figure 4. TT2 Shows Features of an R2R3 MYB DNA Binding Domain Protein.

(A) Deduced amino acid sequence of TT2. The R2R3 MYB DNA binding domain is boxed in black. The amino acid residues conserved between the TT2 and OsMYB3 C-terminal halves are shown in boldface. The dotted line denotes the putative NLS found by the PSORT prediction program. The two putative α-helices found in the C-terminal sequence of TT2 are underlined. Numbers at left indicate amino acid positions (from the translation start codon).

(B) Sequence comparison of the conserved MYB DNA binding domain of TT2 with other MYB-related proteins from rice, maize C1, maize P, petunia AN2, Arabidopsis PAP1, and human c-MYB (for GenBank accession numbers see Methods). Identical amino acids are boxed in black, and similar amino acids are boxed in gray. The dash in the c-MYB sequence indicates a gap introduced to perform the alignment. Asterisks denote the conserved W residues. Arrowheads indicate the amino acid residues in the C1 MYB domain that determine interaction with bHLH-related factors, according to Grotewold and co-workers (Grotewold et al., 2000). The three putative α-helices are noted below the diagram, and the closed box corresponds to the linker sequence between the R2 and R3 repeats. Amino acid residues in TT2 are numbered from the translation start codon.

(C) Dendrogram of relationships among the R2R3 domains from several MYB-related proteins. For construction of the tree, we used only the R2R3 MYB domain sequence (104 amino acid residues; see **[B]**) of each selected MYB-related protein. The matrix of sequence similarities was calculated with the CLUSTAL program from the CLUSTAL X package (Thompson et al., 1997) and submitted to a neighbor-joining analysis to generate a branching pattern. The numbers below the branches indicate the percentage of bootstrap support after 1000 replicates. Nodes with bootstrap support of <50% were discarded. The human c-MYB sequence was included as an outgroup. The consensus tree was drawn using the TreeView program (version 1.5.3, Roderic D.M. Page, University of Glasgow, UK). Sequences used are human c-MYB, maize C1, maize PI, maize P, rice OsMYB3, spruce PmMYBF1, barley HvMYBGA, barley HvMYB1, petunia AN2, Arabidopsis AtMYB1, Arabidopsis AtMYB2, Arabidopsis AtMYB3, Arabidopsis AtMYB5, Arabidopsis AtMYB6, Arabidopsis AtMYB12, Arabidopsis AtMYB15, Arabidopsis AtMYB49, Arabidopsis PAP1, and Arabidopsis TT2.

TT2 and Arabidopsis PAP1 (Figure 4B). In addition, downstream of the MYB region, a short sequence (139-VXXIRTKAI/LRCS/N-150) is conserved between TT2 and OsMYB3 (Figure 4A). To date, more than 100 Arabidopsis R2R3 MYB genes have been sequenced and divided into 22 different subgroups on the basis of limited sequence conservation within their C-terminal regions (Kranz et al., 1998; Romero et al., 1998). Surprisingly, no significant similarity was found within the C terminus between TT2 and any other Arabidopsis MYB proteins. On the basis of these data, we assume

that the *TT2* gene unequivocally represents a novel member within the large Arabidopsis R2R3 MYB gene family.

Finally, the TT2 amino acid sequence was submitted to the Multiple Protein Sequence Analysis program (http://pbil.ibcp.fr/mpsa/) to predict putative protein secondary structures. The algorithm revealed that the C-terminal end of TT2 is capable of forming amphipathic α-helices between amino acids 220 to 227 and 250 to 258 (Figure 4A). Such secondary structures have been described in regulatory proteins with transcriptional activation activity (Ptashne, 1988).

TT2 Is a Nuclear Protein

The in vivo subcellular localization of TT2 was investigated with the aid of green fluorescent protein (GFP). A derivative of GFP, mGFP5 (Siemering et al., 1996), was fused to the *TT2* cDNA and placed under the control of the CaMV dual 35S promoter for ubiquitous and high expression. The *TT2*-GFP transgene was stably introduced into Arabidopsis, and plantlets from primary transformants were observed for GFP activity. The TT2-GFP fusion protein appeared to be localized mostly in plant cell nuclei (Figure 5), demonstrating the presence of a functional nuclear localization signal (NLS) within the TT2 protein. Interestingly, a putative NLS was identified using the PSORT prediction program (http://psort.nibb.ac.jp/) (amino acids 115 to 131; marked on Figure 4A). Whether this NLS is functional remains to be determined.

Mutations Are Present in All *tt2* Mutant Alleles

To identify the lesions in the mutants and to confirm the cloning of *TT2*, sequences of the *TT2* gene were examined in the *tt2* alleles and compared with the corresponding parental ecotype. Sequence comparison of three Arabidopsis ecotypes (Wassilewskija [Ws-2], Landsberg *erecta* [L*er*], and Columbia-0 [Col-0]) showed that L-174 in Ws-2 was replaced by Q in L*er* and Col-0 (Figure 3). No other amino acid polymorphisms were detected within the coding sequences.

Compared with the parental wild type, all *tt2* mutant alleles were found to contain DNA sequence alterations (Table 1, Figure 3), demonstrating conclusively that the *TT2* gene had been cloned. The mutations in *tt2-2* and *tt2-4* were single base pair transitions in the coding sequence. In *tt2-2*, an A-to-T change resulted in a premature translation termination so that the mutant protein lacked the 131 C-terminal amino acids. In *tt2-4*, a G-to-A nucleotide substitution induced the amino acid change from G to R at position 66. This mutation affected the last amino acid of the R2 repeat. In *tt2-1*, the mutation caused a complex lesion in genomic DNA. Indeed, amplification with primers TT2-ATG and TT2-Stop (Figure 3) led to a 450-bp fragment instead of the 920-bp wild-type amplicon. In addition, the sequence of the mutated PCR product matched a portion of Arabidopsis chromosome 4 (data not shown). Thus, deletion and chromosomal rearrangement of the *TT2* region could account for the phenotype of the *tt2-1* mutant. In *tt2-3*, the insertion of a T-DNA copy disrupted the R3 MYB repeat (Figure 3). Finally,

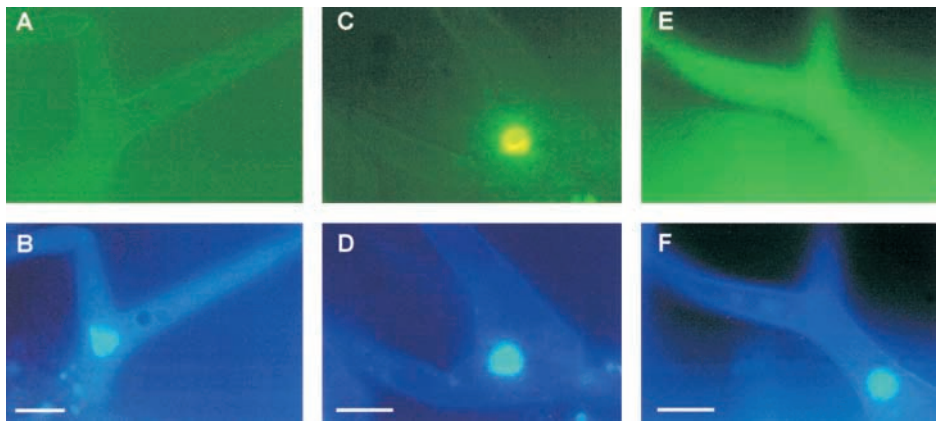

Figure 5. Targeting of the TT2-GFP Fusion Protein to Arabidopsis Cell Nuclei.

Shown are trichome cells from 2-week-old Arabidopsis plants. The positions of the nuclei in **(A)**, **(C)**, and **(E)** are deduced from the comparison with UV light images after 4′,6-diamidino-2-phenylindole staining shown below in **(B)**, **(D)**, and **(F)**, respectively.
(A) and **(B)** Trichome cells of a wild-type nontransformed Arabidopsis plant used as a negative control.
(C) and **(D)** Trichome cells of a TT2-GFP–expressing Arabidopsis transgenic plant.
(E) and **(F)** Trichome cells of a GFP-expressing Arabidopsis transgenic plant.
Bars = 50 μm for **(A)** to **(F)**.

analysis of *TT2* expression in these mutants revealed that *TT2* mRNA was detected in *tt2-2* and *tt2-4*, whereas *tt2-1* and *tt2-3* mutants did not contain a complete *TT2* transcript of the correct size (data not shown).

TT2 Expression

As described previously, *TT2* is required for the normal expression of the *DFR* gene in Arabidopsis immature siliques but not in seedlings (Shirley et al., 1995; Nesi et al., 2000). In addition, cytological analysis showed that *TT2* is necessary for the accumulation of flavonoid compounds in the endothelium of the seed coat (Figures 2D and 2F). Thus, we predicted that *TT2* should be expressed specifically in seed. Because there was not enough *TT2* transcript to be detected by RNA gel blot analysis (data not shown), the spatiotemporal gene expression pattern was determined by quantitative RT-PCR in various vegetative and reproductive organs of Arabidopsis. The transcript was detected at a high level in immature siliques and at a lower level in flowers, but it was undetectable in young seedlings, roots, rosette leaves, and inflorescence stems (Figure 6). Very faint expression was observed in flower buds when the number of PCR cycles was increased to 40 (data not shown). In immature siliques, the *TT2* transcript content was high from the very early stages of embryogenesis to the globular stage of embryo development (Figure 6, stage 3), which corresponds approximately to the third day after pollination under our conditions. The amount of *TT2* mRNA decreased rapidly from the late heart-torpedo stage ($\sim$5 to 6 days after pollination) through subsequent stages of seed development. Finally, the *TT2* transcript did not persist after the completion of embryogenesis (data not shown).

TT2-Regulated Flavonoid Gene Expression

Previous experiments have demonstrated that immature *tt2* siliques lack normal levels of transcripts for at least two flavonoid LBGs, *DFR* and *BAN* (Nesi et al., 2000). To expand on these preliminary results, the expression of several flavonoid genes was analyzed in detail in siliques from *tt2-3* plants. The results are shown in Figure 7.

Quantitative RT-PCR analysis revealed that *tt2-3* siliques contain wild-type amounts of *CHS*, *CHI*, *F3H*, *F3′H*, and *FLS1* transcripts. These genes have been classified as flavonoid EBGs in siliques according to their temporal expression patterns (Nesi et al., 2000) (Figure 1). Because *tt2-3* is a null allele, we inferred that *TT2* is not essential for the transcriptional activation of flavonoid EBGs in siliques. Conversely, transcripts of four flavonoid LBGs, *DFR*, *LEUCOANTHOCYANIDIN DIOXYGENASE* (*LDOX*), *BAN*, and *TT12*, were downregulated in siliques of *tt2-3* (Figure 7). The accumulation of *LDOX* mRNA was reduced severely in *tt2-3* siliques, as observed earlier for *DFR*, whereas *TT12* transcripts were undetectable in *tt2-3*, as shown for *BAN*. Therefore, our results indicated that *TT2* is involved specifically in the genetic control of flavonoid late metabolism in developing siliques.

Finally, we assessed the effect of the *tt2-3* mutation on *TT8* and *TTG1* mRNA accumulation. *tt2-3* siliques did not show any significant differences in the steady state levels of *TT8* and *TTG1* mRNA compared with the wild type (Figure 7), suggesting that *TT2* is not necessary for the expression of these two regulatory genes. Reciprocally, additional experiments revealed that both *TT8* and *TTG1* are not required for *TT2* expression (data not shown).

The Simultaneous Expression of *TT2* and *TT8* Genes Induces Ectopic Activation of Flavonoid LBG Expression

To gain further insight into the function of *TT2* and *TT8* as regulatory genes of flavonoid metabolism in Arabidopsis, we examined how the ectopic expression of *TT2* or *TT8* could affect the expression pattern of several flavonoid structural and regulatory genes. For this purpose, we used plants harboring the 70S-*TT2* transgene (see above) and plants expressing the 70S-*TT8* construct (Nesi et al., 2000).

First, the accumulation of transcripts for several flavonoid genes was examined in young seedlings from transgenic plants relative to nontransformed wild-type plants. Young wild-type seedlings displayed a transient accumulation of anthocyanin pigments just after seed germination (Kubasek et al., 1992) and were shown to accumulate transcripts for *CHS*, *DFR*, *TT8*, and *TTG1* genes but not for *TT2* and *BAN* genes (Figure 8, lane 1). In 70S-*TT2*–expressing seedlings, *BAN* transcripts were induced (Figure 8, lane 2). This finding demonstrates that the overexpression of *TT2* was sufficient to induce *BAN* transcription in seedlings, indicating that TT2 controls the activation of the *BAN* gene.

Further experiments were conducted on roots, a tissue that normally is not pigmented but does accumulate flavonols. In Arabidopsis wild-type roots, transcripts for flavonoid LBGs, as well as for the *TT2* and *TT8* regulatory genes, were undetectable, whereas *CHS* and *TTG1* transcripts accumulated (Figure 8, lane 3). In roots of 70S-*TT2* transgenic plants, high levels of *DFR* and *BAN* transcripts were observed (Figure 8, lane 4), providing additional evidence that TT2 controls the activation of these two flavonoid structural genes. Surprisingly, it was shown that TT2 also induced the expression of the *TT8* gene in roots (Figure 8), indicating that *TT8* mRNA accumulation is controlled by TT2 in this tissue. It should be noted that overexpression of the *TT2* cDNA did not seem to affect the level of *CHS* and *TTG1* mRNA in roots, thus corroborating the finding that TT2 is not required for the expression of these genes in siliques (Figure 7).

Because it was crucial to determine whether TT2 was able

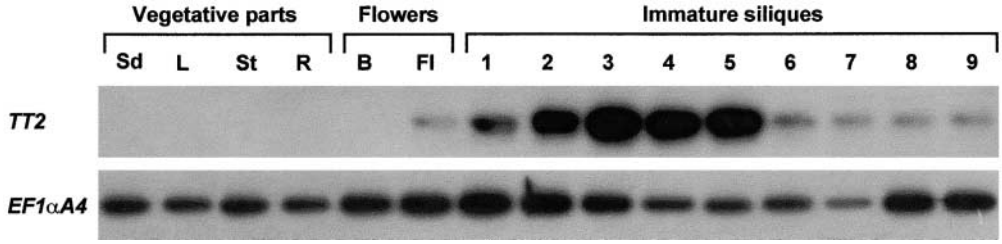

Figure 6. Expression Pattern of the *TT2* Gene.

The accumulation of *TT2* transcript was measured by quantitative RT-PCR with RNA from 4-day-old seedlings (Sd), rosette leaves (L), stems (St), 10-day-old roots (R), flower buds (B), flowers (Fl), and immature siliques at different developmental stages (1 to 9) as indicated. Silique samples were numbered according to the prevailing embryo stage they contained: 1, one cell; 2, one to four cells; 3, early globular to globular; 4, heart; 5, late heart to torpedo; 6, late torpedo to curled cotyledons; 7, late curled cotyledons; 8, green cotyledons; 9, mature embryo. The expression profile of the Arabidopsis *EF1αA4* gene was determined as an mRNA loading control.

to activate flavonoid LBGs in the absence of TT8, the 70S-*TT2* construct was introduced into *tt8-3*, carrying a null mutation at the *TT8* locus (Nesi et al., 2000). Roots of *tt8-3* plants carrying the 70S-*TT2* transgene produced high levels of *TT2* mRNA but failed to accumulate transcripts for either *TT8* or the other flavonoid LBGs tested (Figure 8, lane 6). This result demonstrated conclusively the key role of TT8 in the activation of the flavonoid LBGs examined. Likewise, transgenic plants harboring the 70S-*TT8* construct accumulated *TT8* mRNA in roots, but no effect was observed for any of the transcripts tested (Figure 8, lane 5). Together, these results led to the following conclusions: (1) the ectopic activation of flavonoid LBGs requires TT2 and TT8; (2) there are no functional homologs of *TT2* and *TT8* genes in Arabidopsis roots; and (3) TT2 is capable of inducing *TT8* mRNA accumulation in roots, but the reciprocal induction has not been observed. As a consequence, when expressed ectopically, TT2 is able to trigger by itself the activation of all flavonoid LBGs in roots. Nevertheless, although the roots of 70S-*TT2* plants displayed high levels of *DFR* and *BAN* transcripts (Figure 8, lane 4), they failed to accumulate either proanthocyanidins or their precursors (leucoanthocyanidins and catechins), as revealed using a vanillin test (data not shown). This observation indicated that other tissue-specific factors are probably required.

The fact that *BAN* was expressed in seedlings and roots of 70S-*TT2* plants, whereas the accumulation of these transcripts was restricted to the endothelium of the seed coat in wild-type plants, demonstrated that the ectopic expression of *TT2* was sufficient to alter the spatial expression pattern of *BAN*. This finding indicates that TT2 plays a key role in determining the tissue-specific expression pattern of *BAN*. Finally, *DFR* transcripts were present in seedlings of wild-type plants (Figure 8, lane 1), in which *TT2* was not expressed, suggesting that TT2 function is assumed by other gene(s) in this organ.

DISCUSSION

In Arabidopsis seed, proanthocyanidins begin to accumulate at early stages of embryo development and are located specifically in the endothelium (Devic et al., 1999). Several lines of evidence reported in this study suggest that the *TT2* gene is a major limiting factor in the proanthocyanidin regulatory network. First, *TT2* knockout specifically affected seed pigmentation. Furthermore, it dramatically reduced the expression of several structural genes involved in tannin metabolism. Finally, gain-of-function experiments demonstrated that TT2 induced the ectopic expression of *BAN*, which is supposed to define the first enzyme committed to proanthocyanidin biosynthesis.

TT2 Encodes an R2R3 MYB Domain Protein

Here we report the isolation of a new *tt2* mutant allele, *tt2-3*, which is T-DNA tagged, thereby allowing the cloning of the *TT2* gene. *TT2* was shown to encode an MYB-related protein. MYB proteins are characterized by one to three N-terminal copies of a conserved sequence (the MYB repeat) and by a C-terminal region with little sequence conservation. The MYB domain is assumed to participate in nuclear translocation, DNA binding activity, and protein–protein interaction (reviewed by Lipsick, 1996), whereas the C terminus is presumed to contain transcriptional activation domains (Goff et al., 1991). With regard to the structure of TT2, residues known to be important for the DNA binding activity of MYB-related proteins were well conserved. The three regularly spaced W residues, which are involved in the folding of the MYB domain, were found in each repeat of TT2, except that the first W of the last repeat was substituted by an I residue. Substitution of this W residue occurs in many plant MYB proteins, and it was

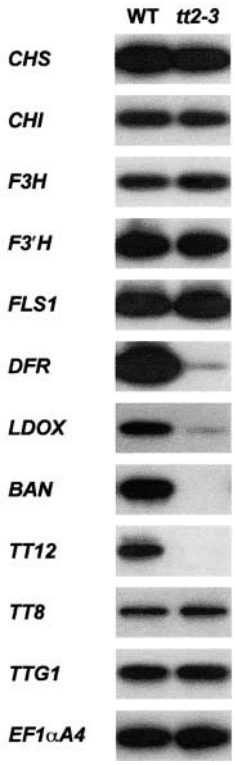

Figure 7. *TT2* Is Essential for the Activation of Flavonoid LBGs.

Transcripts for five flavonoid EBGs (*CHS*, *CHI*, *F3H*, *F3′H*, and *FLS1*), four flavonoid LBGs (*DFR*, *LDOX*, *BAN*, and *TT12*), two regulatory genes (*TT8* and *TTG1*), and, as a control, *EF1αA4* were detected by quantitative RT-PCR in immature siliques of wild-type (WT) and *tt2-3* plants. In these experiments, siliques from approximately the unicellular to the torpedo stage of embryo development (stages 1 to 5 in Figure 6) were pooled and examined.

shown that its change for an aromatic/hydrophobic residue in an animal c-MYB protein did not alter the DNA binding ability (Saikumar et al., 1990; Martin and Paz-Ares, 1997).

Each MYB repeat in TT2 contains three predicted α-helices, as described for the animal c-MYB protein, with the second and third helices forming an HTH motif (Ogata et al., 1992). However, as in most other plant MYB–related proteins, it is unlikely that the HTH motif can form within repeat 2 of TT2 because there is a P residue in the middle of the second helix (Figure 4B) (Jackson et al., 1991; Lipsick, 1996). Nevertheless, in the TT2 protein, residues of the third α-helix of R2 and R3, the predicted DNA recognition helix, were the most conserved, which would support the functionality of TT2 in DNA binding processes. In addition, a computer-assisted search for protein secondary structures revealed that two α-helices could be predicted within the C terminus of TT2. Such structures have been implicated in transcriptional activation (Ptashne, 1988). These findings, together with the nuclear location of TT2, suggest that TT2 acts as a transcriptional activator.

In plants, the first MYB-related protein, C1, was reported in maize (Cone et al., 1986; Paz-Ares et al., 1987). The TT2 protein has two MYB repeats, R2 and R3, in common with most identified plant MYBs (Jin and Martin, 1999). To date, more than 100 R2R3 MYB domain proteins have been identified in Arabidopsis on the basis of sequence conservation (Kranz et al., 1998; Romero et al., 1998; Riechmann et al., 2000). Information on the function of the majority of these MYB-related proteins is scarce, and diverse research programs have been undertaken to identify the corresponding genetic loci (Meissner et al., 1999; Borevitz et al., 2000).

In the *tt2-3* allele, the T-DNA copy is inserted within the R3 MYB repeat. The *tt2-3* mutation was definitely null (i.e., no transcript was produced) and gave the seed a golden-yellow color attributable to the absence of proanthocyanidin accumulation. Because the structure of the R2R3 MYB domain from mutated TT2 in *tt2-3* was disrupted, the T-DNA–tagged allele represents an invaluable tool for subsequent gene expression analyses. Three additional *tt2* alleles were characterized, providing the opportunity to better understand the relationships between the structure of TT2 domains and their respective functions in planta. In *tt2-1*, a complex genomic rearrangement was induced after x-ray mutagenesis that led to a null mutation. In *tt2-2*, a single mismatch introduced a precocious stop codon, so that the encoded protein was truncated just after the R2R3 MYB domain. In this mutant, proanthocyanidin accumulation was detected but severely reduced. In addition, the subcellular location of pigments seemed to be altered, as judged by the diffuse red coloration after vanillin staining (Figure 2G). The leaky phenotype of *tt2-2* seed was correlated with very faint expression of the *BAN* gene (data not shown). This finding suggests that the truncated protein retained DNA binding activity or, alternately, the ability to interact with another transcription factor. Finally, the *tt2-4* mutation was caused by a single amino acid change within the linker sequence between the two MYB repeats (Figure 4B). Transcripts were produced normally, but because the *tt2-4* phenotype is similar to the *tt2-1* and *tt2-3* phenotypes, the mutated protein probably was not functional at all. Functional analyses of the c-MYB protein in animals have proven that the linker sequence is of great importance for the DNA binding properties of MYB domains. In particular, substitutions within the last four amino acids in repeat R2 (LNPE; see Figure 4B) led to the reduced stability of protein-DNA complexes and even the loss of DNA binding activity (Hegvold and Gabrielsen, 1996). Thus, we assumed that the G-to-R amino acid transition at position 66 in *tt2-4* abolished TT2 activity, probably by severely reducing DNA binding affinity or, alternately, by impairing protein interaction with another transcription factor. Among more than 100 R2R3 MYB sequences found in Arabidopsis, only AtMYB52, AtMYB54, and AtMYB56 harbor an R residue at the end of the R2 repeat (Romero et al., 1998). Whether these proteins are functional remains unknown.

Within the widespread R2R3 MYB family in plants, TT2 shows the highest similarity with the rice OsMYB3 (Suzuki et

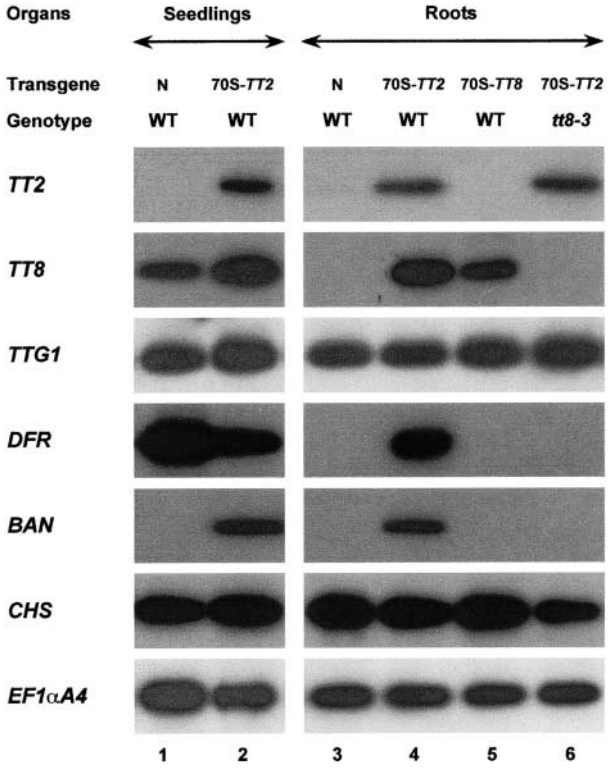

Figure 8. Gene Expression Analysis in Transgenic Plants That Overexpress *TT2* or *TT8*.

Transcripts for the *TT2*, *TT8*, *TTG1*, *DFR*, *BAN*, *CHS*, and *EF1αA4* genes were detected by RT-PCR with RNA from 4-day-old seedlings (lanes 1 and 2) and 10-day-old roots (lanes 3 to 6) of wild-type (WT) nontransformed plants (lanes 1 and 3), wild-type plants harboring the 70S-*TT2* transgene (lanes 2 and 4), wild-type plants harboring the 70S-*TT8* transgene (lane 5), and *tt8-3* plants harboring the 70S-*TT2* transgene (lane 6). N, none.

al., 1997). The two MYB proteins are related structurally, especially in their DNA binding domains, which share 84% amino acid identity, but also in their C-terminal ends, where a short sequence is conserved (Figure 4A). This amino acid stretch was not found in any other plant MYB sequence and may be the signature for a novel subgroup of MYB proteins in the classification established previously (Kranz et al., 1998). The presence of a short signature uniting TT2 and OsMYB3 coupled with the results of a neighbor-joining analysis (Figure 4C) establish that these proteins are likely to be orthologs, although it is not clear at this time that OsMYB3 is involved in rice flavonoid metabolism. The TT2 protein also is closely related to the maize C1 factor. In addition, transient assays on immature maize kernels demonstrate that TT2 can substitute efficiently for C1 when used in combination with B-Peru (P. Perez, personal communication), thus indicating that C1 and TT2 activate the same target genes in the maize anthocyanin pathway. Strikingly, TT2 displays a lower amino acid conservation with the petunia AN2 MYB factor, although these two proteins appear to regulate similar sets of flavonoid structural genes in their respective hosts (Quattrocchio et al., 1993, 1998, 1999).

TT2 Is Expressed Specifically in Immature Seed, in Which It Controls Flavonoid LBGs

The *TT2* gene was shown to be expressed specifically in seed and temporally restricted to a short developmental window that spans only the early stages of embryogenesis. Consistent with the phenotype of *tt2* mutants, these results support the conclusion that *TT2* expression is seed specific. The stringent temporal expression pattern of *TT2* may explain why the corresponding cDNA was not found in Arabidopsis silique cDNA libraries, in which young immature siliques are underrepresented (J. Giraudat, personal communication). Similar reasons may account for the fact that no *TT2* expressed sequence tag was found in the databases.

Because we demonstrated that both the *tt2-3* mutation and ectopic expression of the *TT2* cDNA altered the expression pattern of flavonoid LBGs, whereas no effect was observed on the expression of flavonoid EBGs, we assumed that TT2 is involved specifically in the regulation of flavonoid late metabolism. In addition, the time course of *TT2* gene expression correlated with those of flavonoid LBGs, such as *BAN* and *TT12*, which were restricted to the seed during early embryogenesis (Devic et al., 1999; Debeaujon et al., 2001). It also overlapped the precocious deposition of proanthocyanidins in the endothelium. Indeed, proanthocyanidins start to accumulate at the two-cell stage of embryo development as colorless compounds (Debeaujon et al., 2001) and turn brown during the seed maturation–desiccation process, when *TT2*, as well as other flavonoid LBG transcripts, has disappeared completely. Inspection of the sequences from flavonoid LBG promoters revealed the presence of putative recognition sites for MYB-related DNA binding proteins, including matches to the animal consensus (C/TAACG/TG; Luscher and Eisenman, 1990) and to the DNA binding site of the maize C1 transcription factor (AC/ACT/AAC/AC; Sainz et al., 1997). Together, these results raise the possibility that TT2 interacts directly with flavonoid LBGs and activates their transcription.

Control of Flavonoid LBG Expression Involves the Combinatorial Action of TT2 and TT8

Our results strongly suggest that TT2 activity is tightly linked to the presence of TT8. Previously, we reported that TT8, like TT2, is required for the normal expression of flavonoid LBGs (Nesi et al., 2000). Here we demonstrate that the ectopic activation of *BAN* in *TT2*-overexpressing plants depends strictly on the presence of a functional TT8 protein in the organs examined. These findings highlight the importance of

TT8 in the flavonoid regulatory network in Arabidopsis siliques. Nevertheless, a comparison of temporal gene expression patterns revealed that a steady state level of *TT8* mRNA was maintained throughout seed development (Nesi et al., 2000), whereas transcripts for *TT2* and LBGs decreased rapidly from the torpedo embryo stage onward. In addition, *TT8* also was expressed in young seedlings, in which *BAN* mRNA was not detected. Together, these results suggest that although *TT8* undoubtedly is required for the full transcriptional activation of flavonoid LBGs, the developmental competence of tissues to accumulate *BAN* transcripts is rate limited both spatially and temporally by *TT2* expression. Similar data were found in petunia flower, in which the tissue-specific accumulation of anthocyanins depends on the tissue-specific expression of MYB genes, in particular *AN2* and *AN4*, whereas *AN1* and *AN11* genes are expressed in all pigmented tissues (Quattrocchio et al., 1993, 1999; Spelt et al., 2000). Conversely, in maize, the tissue-specific pigmentation is caused in many cases by the tissue-specific expression of particular alleles of the bHLH genes *R1* and *B1* rather than, or in addition to, control by the MYB gene (Ludwig and Wessler, 1990).

The nature of the relationship between TT2 and TT8 has not been established. The occurrence of a protein–protein interaction has emerged from extensive studies with maize. Indeed, the activation of anthocyanin biosynthesis in maize requires an MYB domain protein (C1 or Pl) in combination with a bHLH-related factor (R or B), which were shown to interact physically in yeast two-hybrid assays (Goff et al., 1992). A similar situation was described for the MYB and bHLH transcription factors involved in the trichome development process in Arabidopsis (Payne et al., 2000). However, the interaction between MYB-related factors and other regulatory proteins is not a compulsory mechanism. For instance, the maize P protein is able to trigger the activation of flavonoid structural genes by itself (Grotewold et al., 1994). Recent work identified the residues in the MYB domain of C1 that determine the specific interaction with R-like proteins (Grotewold et al., 2000). The conservation of these residues in TT2 suggests that they may be functionally relevant, thus supporting the hypothesis that TT2 action on gene regulation is mediated through combinatorial interaction with another factor. In this way, TT8 appears to be a good candidate, but a direct association between TT2 and TT8 remains to be demonstrated.

It is interesting that the ectopic expression of *TT2* in roots induces the transcriptional activation of *TT8*. The fact that a null allele displayed wild-type amounts of *TT8* mRNA in siliques suggests that the loss of TT2 function may be compensated for by the presence of other protein(s) that have at least partially redundant functions during seed development. Similar results were obtained with petunia, in which the MYB protein AN2 did not appear to regulate *AN1* mRNA levels in petals; but when ectopically expressed, AN2 apparently activated *AN1* expression in other tissues such as the leaf (Spelt et al., 2000). In maize, there is no evidence that MYB and bHLH proteins control each others expression (Goff et al., 1990; Sainz et al., 1997).

Finally, this study proved that both *TT2* and *TT8* genes are involved in the regulation of flavonoid late metabolism in Arabidopsis seed. It also demonstrated that TT2 overexpression is sufficient to induce ectopic expression of the *BAN* gene, thus suggesting that TT2 is a limiting factor for the accumulation pattern of *BAN* transcripts. A crucial issue to be investigated now is how the spatial and temporal regulation of the *TT2* gene is determined. The occurrence of an upstream regulator of TT2 may be suggested by studies with maize, in which the transcriptional regulator VIVIPAROUS1 was shown to control the transcription of the *C1* gene (Hattori et al., 1992).

Other Factors Involved in the Regulation of Flavonoid Metabolism Remain to Be Found

Several lines of evidence indicate that additional regulatory elements are still to be found to complete our current view of the flavonoid regulatory network in Arabidopsis seed. A previous study reported that *BAN* is epistatic to *TT2* (Albert et al., 1997), which seems inconsistent with the fact that *BAN* transcripts were undetectable in immature siliques of the *tt2-3* null mutant. To explain this apparent discrepancy, we assumed that very low levels of *BAN* are present in *tt2-3* but are undetectable using classic molecular methods. The basal level of *BAN* may be independent of TT2 activity. Therefore, the remaining BAN in *tt2-3* may allow entry into the proanthocyanidin biosynthetic subpathway, although the *tt2* mutation subsequently avoids the production and/or accumulation of proanthocyanidins. Conversely, the complete loss of BAN in *ban* and *tt2 ban* mutants probably leads to the accumulation of anthocyanins as a result of metabolic rechanneling, as hypothesized by Devic and co-workers (Devic et al., 1999).

Our results demonstrated that TT2 overexpression was not sufficient to induce the ectopic accumulation of proanthocyanidins, suggesting that additional tissue-specific factors are required for a complete functional biosynthetic pathway. Interestingly, overexpression of the Arabidopsis *PAP1* gene in *pap1-D* mutants is sufficient to induce the overall anthocyanin pathway, leading to the overaccumulation and ectopic accumulation of these compounds (Borevitz et al., 2000). Therefore, it will be instructive to learn whether the *pap1-D* mutation, combined with the overexpression of *TT2*, might circumvent a requirement for specific regulators involved in proanthocyanidin deposition.

METHODS

Plant Materials

The dro55 transformant was identified by visual examination of the progeny of 15,000 independent T-DNA–mutagenized lines of *Arabi-*

dopsis thaliana (Wassilewskija-2 [Ws-2] ecotype) generated at the Station de Génétique et d'Amélioration des Plantes at the Institut National de la Recherche Agronomique, Versailles, France (Bechtold et al., 1993; Bouchez et al., 1993). The *tt2-1* mutant was described by Koornneef (1981, 1990) and is available through the Nottingham Arabidopsis Stock Centre seed catalog (http://nasc.nott.ac.uk/; seed stock number N83). The 3C mutant was isolated after γ-ray mutagenesis on the basis of seed color and reduced seed dormancy (Léon-Kloosterziel et al., 1996). Line JSM9-3 was obtained from the Sendai ethyl methanesulfonate mutant collection generated by Dr. N. Goto (available at http://www.shigen.nig.ac.jp). The three mutants *tt2-1*, 3C, and JSM9-3 are in the Landsberg *erecta* (L*er*) ecotype. The *tt8-3* mutant allele used for the expression analyses has been described previously (Nesi et al., 2000). Details of plant growth conditions and in vitro selection of transgenic plantlets were described by Nesi et al. (2000). Plant materials used in this study are available at the Nottingham Arabidopsis Stock Centre with the seed stock numbers given in parentheses: *tt2-3* (N890), *tt8-1* (N891), 70S-TT2 construct in wild-type background (N892), 70S-TT2 construct in *tt2-3* (N893), TT2-GFP construct in wild-type background (N894), 70S-TT8 in wild-type background (N895), and 70S-TT8 in *tt8-3* background (N896). No restrictions or conditions will be placed on the use of any novel materials described in this publication that would limit their use in noncommercial research. Details on the constructs are given below.

Nucleic Acid Analyses

Genomic DNA extraction, polymerase chain reaction (PCR) amplification, analyses of DNA sequences, and quantification of mRNA by reverse transcription (RT)-PCR analysis were performed as described previously (Nesi et al., 2000).

Primers TT2-5′ (5′-GGTCAACGCCGGTCAACAC-3′) and TT2-3′ (5′-AATGTATTCTAATACATTTCT-3′) were derived from plant genomic DNA and used to amplify a 2700-bp DNA fragment (DDBJ/EMBL/GenBank accession number AJ299452). BLAST searches (http://www.ncbi.nlm.nih.gov/BLAST/) with the AJ299452 sequence allowed us to recover an Arabidopsis genomic P1 clone (clone MOK9; GenBank accession number AB015477). Data on the MOK9 sequence and gene information are available on the KAOS (Kazusa Arabidopsis Data Opening Site) database at http://www.kazusa.or.jp/kaos/.

Gene-specific primers used for RT-PCR assays were as follows: for *TT2*, TT2-ATG (5′-ATGGGAAAGAGAGCAACTACTAGTGTG-3′) and TT2-Stop (5′-TCAACAAGTGAAGTCTCGGAGC-3′); for chalcone isomerase (CHI; GenBank accession number M86358), CHI-UP (5′-TGTCTTCATCCAACGCCTGCGC-3′) and CHI-RP (5′-TGAAGATCTCCAAGAACTTCTCC-3′); for flavanone 3-hydroxylase (F3H; GenBank accession number U33932), F3H-UP (5′-AAATTCGTCCGAGACGAAGACG-3′) and F3H-RP (5′-GTCTTGTAGCAGCAAGGTAATGG-3′); for flavonoid 3′-hydroxylase (F3′H; GenBank accession number AH009204), F3′H-UP (5′-CATGGCAACTCTATTTCTCAC-3′) and F3′H-RP (5′-CGTCACCGTCAAGATCAGTTCC-3′); for flavonol synthase (FLS1; GenBank accession number U72631), FLS1-UP (5′-ATGGAGGTCGAAAGAGTCCAAG-3′) and FLS1-RP (5′-TCCACTGAGATCTGTATGAGCC-3′); for leucoanthocyanidin dioxygenase (LDOX; GenBank accession number U70478), LDOX-UP (5′-ATGGTTGCGGTTGAAAGAGTTG-3′) and LDOX-RP (5′-ATCGGTGTGAGCTTCCACACCG-3′); and for *TTG1* (GenBank accession number AJ133743), TTG1-UP (5′-CTCCAGATTCGTTATCCAGATCG-3′) and TTG1-RP (5′-CGTATCAGGCTGAGGACTCTCG-3′). For *TT12* (GenBank accession number AJ294464), we used primers TT12-fw and TT12-rev designed by Debeaujon et al. (2001), and for *BAN* (GenBank accession number AF092912), we used primers BAN-RT5′ and BAN-RT3′ designed by Devic et al. (1999). Primers for chalcone synthase (*CHS*; CHS-UP and CHS-RP), dihydroflavonol reductase (*DFR*; DFR-UP and DFR-RP), *TT8* (deb122RB1 and deb122LB3), and *EF1αA4* (EF1αA4-UP and EF1αA4-RP) were those designed by Nesi et al. (2000).

Construction and Analyses of Transgenic Plants

All PCR amplifications to generate constructs used for plant transformation were conducted with high-fidelity DNA polymerase (PfuTurbo DNA polymerase; Stratagene), and PCR products were verified by DNA sequencing. Final constructs were inserted into the binary T-DNA vector pBIB-Hyg, which carries a hygromycin resistance marker for plant selection (Becker, 1990), and then introduced into *Agrobacterium tumefaciens* C58C1Rif(pmp90). Plant transformation was performed as described by Clough and Bent (1998). Arabidopsis adult plants were dipped into a solution containing Agrobacterium, 5% sucrose, and 50 μL/L of the surfactant Silwet (Witco, Geneva, Switzerland).

The 70S-*TT2* expression cassette was obtained by ligation of the *TT2* cDNA between the double enhanced Cauliflower mosaic virus (CaMV) 35S promoter and the CaMV polyadenylation signals of plasmid pLBR19 (Guerineau et al., 1992). Details of the 70S-*TT8* transgene can be found elsewhere (Nesi et al., 2000). For gene expression analyses of plants transformed with the 70S-*TT2* (or 70S-*TT8*) construct, care was taken that the analyzed plants harbored only one functional hygromycin-resistant locus (segregation ratio of 3:1 for hygromycin-resistant to hygromycin-sensitive plants). All experiments were conducted with individuals selected on hygromycin medium, ensuring that each plant carried at least one copy of the transgene.

For construction of the TT2−green fluorescent protein (GFP) fusion protein, the complete *TT2* cDNA (without stop codon) was amplified using a 5′ upstream primer (5′-GTGACCATGGGAAAGAGAGCAACTACTAGTGTG-3′) and a 3′ downstream primer (5′-ATATCCATGGAACAAGTGAAGTCTCGGAGCC-3′), both designed with an engineered NcoI site. The PCR fragment was cloned as an in-frame N-terminal fusion to the GFP of vector pAVA393. Plasmid pAVA393 is the same as pAVA319 (von Arnim et al., 1998) except that the GFP cDNA is version mGFP5 (Siemering et al., 1996). Gene expression was driven by the CaMV dual 35S promoter and terminator signals and enhanced by the translational leader sequence of *Tobacco etch virus* present on plasmid pAVA393. Live seedlings of GFP-expressing transgenic plants were subjected to direct microscopic inspection (Axioplan 2; Zeiss, Jena, Germany) with filter sets for fluorescence (BP450/490, BSFT510, and EMLP515) to visualize the TT2-GFP fusion and GFP. To determine the localization of nuclei, leaf fragments were stained with 1% 4′,6-diamidino-2-phenylindole (Sigma) in distilled water and 0.1% Triton X-100 and observed under UV light illumination (filter sets BP365/12, BSFT395, and EMLP397).

Histological Analysis

Histological sections were prepared as follows. Immature siliques of wild-type and *tt2-3* plants were harvested, fixed immediately in 4% *p*-formaldehyde in PBS (130 mM NaCl, 7 mM Na_2HPO_4, and 3 mM NaH_2PO_4, pH 7.2) under vacuum for 1 hr, and transferred to fresh fixative for 16 hr at 4°C. After fixation, the material was rinsed twice for

10 min in PBS and then dehydrated through a series of acetone solutions (50, 70, 90, and 100%) for 2 hr each. The material was embedded in Technovit 7100 resin (Heraeus Kulzer Histo-Technik, Wehrheim, Germany) according to the manufacturer's instructions. The Technovit-embedded samples were sectioned to a thickness of 2 μm using a rotary microtome. Sections were stained in toluidine blue (1% in distilled water) for 1 min, rinsed abundantly with sterile water, and observed with a light microscope.

The vanillin test (Aastrup et al., 1984) was performed by direct incubation of immature siliques or root samples in a freshly prepared solution of 1% (w/v) vanillin (4-hydroxy-3-methoxybenzaldehyde; Sigma) in 6 N HCl for 30 min at room temperature. Under acidic conditions, vanillin turns red upon binding to flavan-3,4-diols (leucoanthocyanidins) and flavan-3-ol (catechins), which are present as monomers or as terminal subunits of proanthocyanidins (Deshpande et al., 1986).

GenBank Accession Numbers

The GenBank accession numbers are as follows (see Figure 4): rice OsMYB3 (D88619), maize C1 (M37153), maize P (U57002), petunia AN2 (AF146702), Arabidopsis PAP1 (AtMYB75, AF062908), human c-MYB (M15024), maize Pl (L19494), spruce PmMYBF1 (U39448), barley HvMYBGA (X87690), barley HvMYB1 (X70877), Arabidopsis AtMYB1 (D10936), Arabidopsis AtMYB2 (D14712), Arabidopsis AtMYB3 (AF062859), Arabidopsis AtMYB5 (U26935), Arabidopsis AtMYB6 (U26936), Arabidopsis AtMYB12 (AF062864), Arabidopsis AtMYB15 (X90384), Arabidopsis AtMYB49 (AF175991), and Arabidopsis TT2, this work (AJ299452).

ACKNOWLEDGMENTS

We are especially grateful to Brenda Winkel-Shirley and Annie Marion-Poll for critical review of the manuscript and to Martine Devic for helpful discussion during the progress of this work. We thank Karen Léon-Kloosterziel for providing the 3C mutant, Nobuharu Goto for providing the JSM9-3 mutant, the Nottingham Arabidopsis Stock Centre for providing the *tt2-1* allele, and Albrecht von Arnim for the generous gift of plasmid pAVA393. We also thank Pascual Perez for sharing unpublished data. Nicole Bechtold and Georges Pelletier are acknowledged for allowing us to screen the Versailles T-DNA collection. Many thanks to Joël Talbotec and Hervé Ferry for technical assistance with plant culture. This work was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (Grant 97-5-11772) to N.N.

Received March 8, 2001; accepted June 9, 2001.

REFERENCES

Aastrup, S., Outtrup, H., and Erdal, K. (1984). Location of the proanthocyanidins in the barley grain. Carlsberg Res. Commun. **49,** 105–109.

Albert, S., Delseny, M., and Devic, M. (1997). *BANYULS*, a novel negative regulator of flavonoid biosynthesis in the Arabidopsis seed coat. Plant J. **11,** 289–299.

Bechtold, N., Ellis, J., and Pelletier, G. (1993). In planta Agrobacterium-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris **316,** 1194–1199.

Becker, D. (1990). Binary vectors which allow the exchange of plant selectable markers and reporter genes. Nucleic Acids Res. **18,** 203.

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

Bouchez, D., Camilleri, C., and Caboche, M. (1993). A binary vector based on Basta resistance for in planta transformation of *Arabidopsis thaliana*. C. R. Acad. Sci. Paris **316,** 1188–1193.

Chapple, C.C.S., Shirley, B.W., Zook, M., Hammerschmidt, R., and Somerville, S.C. (1994). Secondary metabolism in Arabidopsis. In Arabidopsis, E.M. Meyerowitz and C.R. Somerville, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 989–1030.

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.

Cone, K.C., Burr, F.A., and Burr, B. (1986). Molecular analysis of the maize anthocyanin regulatory locus *C1*. Proc. Natl. Acad. Sci. USA **83,** 9631–9635.

Debeaujon, I., Léon-Kloosterziel, K.M., and Koornneef, M. (2000). Influence of the testa on seed dormancy, germination and longevity in Arabidopsis. Plant Physiol. **122,** 403–413.

Debeaujon, I., Peeters, A.J.M., Léon-Kloosterziel, K.M., and 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,** 853–872.

Deshpande, S.S., Cheryan, M., and Salunkhe, D.K. (1986). Tannin analysis of food products. CRC Crit. Rev. Food Sci. Nutr. **24,** 401–449.

de Vetten, N., Quattrocchio, F., Mol, J., and Koes, R. (1997). The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev. **11,** 1422–1434.

Devic, M., Albert, S., Delseny, M., and Roscoe, T.J. (1997). Efficient PCR walking on plant genomic DNA. Plant Physiol. Biochem. **35,** 331–339.

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.

Frampton, J., Gibson, T.J., Ness, S.A., Döderlein, G., and Graf, T. (1991). Proposed structure for the DNA-binding domain of the Myb oncoprotein based on model building and mutational analysis. Protein Eng. **4,** 891–901.

Goff, S.A., Klein, T.M., Roth, B.A., Fromm, M.E., Cone, K.C., Radicella, J.P., and Chandler, V.L. (1990). Transactivation of anthocyanin biosynthetic genes following transfer of *B* regulatory genes into maize tissues. EMBO J. **9,** 2517–2522.

Goff, S.A., Cone, K.C., and Fromm, M.E. (1991). Identification of functional domains in the maize transcriptional activator C1: Com-

parison of wild-type and dominant inhibitor proteins. Genes Dev. **5,** 298–309.

Goff, S.A., Cone, K.C., and Chandler, V.L. (1992). Functional analysis of the transcriptional activator encoded by the maize *B* gene: Evidence for a direct functional interaction between two classes of regulatory proteins. Genes Dev. **6,** 864–875.

Grotewold, E., Drummond, B.J., Bowen, B., and Peterson, T. (1994). The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell **76,** 543–553.

Grotewold, E., Sainz, M.B., Tagliani, L., Hernandez, J.M., Bowen, B., and Chandler, V.L. (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.

Guerineau, F., Lucy, A., and Mullineaux, P. (1992). Effect of two consensus sequences preceding the translation initiator codon on gene expression in plant protoplasts. Plant Mol. Biol. **18,** 815–818.

Harborne, J.B., and Williams, C.A. (2000). Advances in flavonoid research since 1992. Phytochemistry **55,** 481–504.

Hattori, T., Vasil, V., Rosenkrans, L., Hannah, L.C., McCarty, D.R., and Vasil, I.K. (1992). The *Viviparous-1* gene and abscisic acid activate the *C1* regulatory gene for anthocyanin biosynthesis during seed maturation in maize. Genes Dev. **6,** 609–618.

Hegvold, A.B., and Gabrielsen, O.S. (1996). The importance of the linker connecting the repeats of the c-Myb oncoprotein may be due to a positioning function. Nucleic Acids Res. **24,** 3990–3995.

Jackson, D., Culianez-Macia, F., Prescott, A.G., Roberts, K., and Martin, C. (1991). Expression patterns of *myb* genes from Antirrhinum flowers. Plant Cell **3,** 115–125.

Jende-Strid, B. (1993). Genetic control of flavonoid biosynthesis in barley. Hereditas **119,** 187–204.

Jin, H., and Martin, C. (1999). Multifunctionality and diversity within the plant *MYB*-gene family. Plant Mol. Biol. **41,** 577–585.

Klempnauer, K.H., Gonda, T.J., and Bishop, J.M. (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: The architecture of a transduced oncogene. Cell **31,** 453–463.

Koes, R.E., Quattrocchio, F., and Mol, J.N.M. (1994). The flavonoid biosynthetic pathway in plants: Function and evolution. BioEssays **16,** 123–132.

Koornneef, M. (1981). The complex syndrome of *ttg* mutants. Arabidopsis Inf. Serv. **18,** 45–51.

Koornneef, M. (1990). Mutations affecting the testa colour in Arabidopsis. Arabidopsis Inf. Serv. **27,** 1–4.

Koornneef, M., van Eden, J., Hanhardt, C.J., Stam, P., Braaksma, F.J., and Feenstra, W.J. (1983). Linkage map of *Arabidopsis thaliana*. J. Hered. **74,** 265–272.

Kranz, H.D., et al. (1998). Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. Plant J. **16,** 263–276.

Kubasek, W.L., Shirley, B.W., McKillop, A., Goodman, H.M., Briggs, W., and Ausubel, F.M. (1992). Regulation of flavonoid biosynthetic genes in germinating Arabidopsis seedlings. Plant Cell **4,** 1229–1236.

Léon-Kloosterziel, K.M., van de Bunt, G.A., Zeevaart, J.A.D., and Koornneef, M. (1996). Arabidopsis mutants with a reduced seed dormancy. Plant Physiol. **110,** 233–240.

Lipsick, J.S. (1996). One billion years of MYB. Oncogene **13,** 223–235.

Liu, Y.-G., Mitsukawa, N., Vazquez-Tello, A., and Whittier, R.F. (1995). Generation of a high-quality P1 library of Arabidopsis suitable for chromosome walking. Plant J. **7,** 351–358.

Lloyd, A.M., Walbot, V., and Davis, R.W. (1992). Arabidopsis and Nicotiana anthocyanin production activated by maize regulators *R* and *C1*. Science **258,** 1773–1775.

Ludwig, S.R., and Wessler, S.R. (1990). Maize *R* gene family: Tissue-specific helix-loop-helix proteins. Cell **62,** 849–851.

Ludwig, S.R., Habera, L.F., Dellaporta, S.L., and Wessler, S.R. (1989). *Lc*, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the *myc*-homology region. Proc. Natl. Acad. Sci. USA **86,** 7092–7096.

Luscher, B., and Eisenman, R.N. (1990). New light on Myc and Myb. Part II. Myb. Genes Dev. **4,** 2235–2241.

Martin, C., and Paz-Ares, J. (1997). MYB transcription factors in plants. Trends Genet. **13,** 67–73.

Martin, C., Prescott, A., Mackay, S., Bartlett, J., and Vrijlandt, E. (1991). Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. Plant J. **1,** 37–49.

Meissner, R.C., et al. (1999). Function search in a large transcription factor gene family in Arabidopsis: Assessing the potential of reverse genetics to identify insertional mutations in R2R3 MYB genes. Plant Cell **11,** 1827–1840.

Mol, J., Grotewold, E., and Koes, R. (1998). How genes paint flowers and seeds. Trends Plant Sci. **3,** 212–217.

Nakamura, Y., Sato, S., Asamizu, E., Kaneko, T., Kotani, H., Miyajima, N., and Tabata, S. (1998). Structural analysis of *Arabidopsis thaliana* chromosome 5. VII. Sequence features of the regions of 1,013,767 bp covered by sixteen physically assigned P1 and TAC clones. DNA Res. **5,** 297–308.

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.

Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S., and Nishimura, Y. (1992). Solution structure of a DNA-binding unit of Myb: A helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. Proc. Natl. Acad. Sci. USA **89,** 6428–6432.

Payne, C.T., Zhang, F., and Lloyd, A.M. (2000). *GL3* encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. Genetics **156,** 1349–1362.

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A., and Saedler, H. (1987). The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. **6,** 3553–3558.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature **335,** 683–689.

Quattrocchio, F., Wing, J.F., Leppen, H.T.C., Mol, J.N.M., and

Koes, R.E. (1993). Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. Plant Cell **5,** 1497–1512.

Quattrocchio, F., Wing, J.F., van der Woude, K., Mol, J.N.M., and Koes, R. (1998). Analysis of bHLH and MYB domain proteins: Species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J. **13,** 475–488.

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

Riechmann, J.L., et al. (2000). Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. Science **290,** 2105–2110.

Romero, I., Fuertes, A., Benito, M.J., Malpica, J.M., Leyva, A., and Paz-Ares, J. (1998). More than 80 R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. Plant J. **14,** 273–284.

Saikumar, P., Murali, R., and Reddy, E.P. (1990). Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions. Proc. Natl. Acad. Sci. USA **87,** 8452–8456.

Sainz, M.B., Grotewold, E., and Chandler, V.L. (1997). Evidence for direct activation of an anthocyanin promoter by the maize C1 protein and comparison of DNA binding by related Myb domain proteins. Plant Cell **9,** 611–625.

Schoenbohm, C., Martens, S., Eder, C., Forkmann, G., and Weisshaar, B. (2000). Identification of the *Arabidopsis thaliana* flavonoid 3′-hydroxylase gene and functional expression of the encoded P450 enzyme. Biol. Chem. **381,** 749–753.

Shirley, B.W. (1996). Flavonoid biosynthesis: "New" functions for an "old" pathway. Trends Plant Sci. **1,** 377–382.

Shirley, B.W., Hanley, S., and Goodman, H.M. (1992). Effects of ionizing radiation on a plant genome: Analysis of two Arabidopsis transparent testa mutations. Plant Cell **4,** 333–347.

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.

Siemering, K.R., Golbik, R., Sever, R., and Haseloff, J. (1996). Mutations that suppress the thermosensitivity of green fluorescent protein. Curr. Biol. **6,** 1653–1663.

Spelt, C., Quattrocchio, F., Mol, J.N.M., and Koes, R. (2000). *anthocyanin1* of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. Plant Cell **12,** 1619–1631.

Stafford, H.A. (1989). The enzymology of proanthocyanidin biosynthesis. In Chemistry and Significance of Condensed Tannins, R.W. Hemingway and J.J. Karchesy, eds (New York: Plenum Press), pp. 47–70.

Suzuki, A., Suzuki, T., Tanabe, F., Toki, S., Washida, H., Wu, C.Y., and Takaiwa, F. (1997). Cloning and expression of five myb-related genes from rice seed. Gene **198,** 393–398.

Tanner, G.J., and Kristiansen, K.N. (1993). Synthesis of 3,4-cis-leucocyanidin and enzymatic reduction to catechin. Anal. Biochem. **209,** 274–277.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL X Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25,** 4876–4882.

von Arnim, A.G., Deng, X.W., and Stacey, M.G. (1998). Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. Gene **221,** 35–43.

Walker, A.R., Davison, P.A., Bolognesi-Winfield, A.C., James, C.M., Srinivasan, N., Blundell, T.L., Esch, J.J., Marks, M.D., and Gray, J.C. (1999). The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. Plant Cell **11,** 1337–1349.

Winkel-Shirley, B. (1998). Flavonoids in seeds and grains: Physiological function, agronomic importance and the genetics of biosynthesis. Seed Sci. Res. **8,** 415–422.

Wisman, E., Hartmann, U., Sagasser, M., Baumann, E., Palme, K., Hahlbrock, K., Saedler, H., and Weisshaar, B. (1998). Knock-out mutants from an En-1 mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. Proc. Natl. Acad. Sci. USA **95,** 12432–12437.