Anthocyanin Regulatory/Structural Gene Expression in *Phalaenopsis*

Hongmei Ma and Margaret Pooler
U.S. Department of Agriculture/Agricultural Research Service, U.S. National Arboretum, Floral and Nursery Plants Research Unit, 3301 New York Avenue, NE, Washington, DC 20002

Robert Griesbach
U.S. Department of Agriculture/Agricultural Research Service, U.S. National Arboretum, Floral and Nursery Plants Research Unit, 10300 Baltimore Avenue, Building 003, Beltsville, MD 20705

Additional index words. Myb, Myc, Wd, Chs, Dfr, orchid

Abstract. Anthocyanin biosynthesis requires the coordinated expression of Myc, Wd, Chs, Dfr, and Myb. Chs and Dfr are structural genes, while Myc, Myb, and Wd are regulatory genes. Reverse transcription polymerase chain reaction was used to measure the expression of these genes in *Phalaenopsis amabilis* and *Phalaenopsis schilleriana*. *P. amabilis* expresses an albescent phenotype with petals and sepals that are anthocyanin free, while *P. schilleriana* has a wild-type phenotype with anthocyanin-containing petals and sepals. As expected, the petals and sepals of *P. schilleriana* expressed high levels of Chs and Dfr. The petals and sepals of *P. amabilis* expressed high levels of Chs and very low levels of Dfr. In *P. amabilis* and *P. schilleriana*, anthocyanin-specific Myc and Wd were expressed; however, Myb specific for anthocyanin biosynthesis were undetectable in *P. amabilis*. This suggests that the absence of Myb expression was responsible for the lack of dihydroflavonol 4-reductase and results in the absence of anthocyanin pigmentation in *P. amabilis* petals and sepals. This was confirmed by particle bombardment of *P. amabilis* petals with functional Mybs isolated from *P. schilleriana*. Comparisons of anthocyanin-related Myb gene expression between *P. schilleriana* and *P. amabilis* are between genetically different species. *Phalaenopsis* ‘Everspring Fairy’ expresses a harlequin phenotype with white petals and sepals containing large anthocyanin sectors. Harlequin flowers are ideal to evaluate anthocyanin-related Myb gene expression within genetically identical but differently pigmented tissue. High levels of anthocyanin-specific Myb and Dfr transcripts were present in the purple, but not in the white, sectors of *Phalaenopsis* ‘Everspring Fairy’ petals and sepals. There was no differential expression of Chs, Wd, and Myc between the purple and white sectors. These results are in agreement with the results from *P. amabilis* and *P. schilleriana*.

Anthocyanin biosynthesis was first characterized in *Zea mays* and was demonstrated to be predominantly regulated at the transcription level by two families of regulatory factors, R-like MYC (R, B, Lc, Sn) and R2R3-MYB (C1/Pl) proteins (Brevitz et al., 2000; Chandler et al., 1989; Cone et al., 1986; Grotewold et al., 1989; Hernandez et al., 2004; Irani et al., 2003; Ludwig et al., 1989; Paz-Ares et al., 1987; Radicella et al., 1991; Sainz et al., 1997; Zhang et al., 2000). The defining characteristic of the MYC family is the presence of a conserved basic helix-loop-helix (bHLH) bipartite domain consisting of a presumed DNA-binding basic region and a protein-protein dimerization motif. MYB proteins are defined by the presence of a conserved DNA-binding domain that is generally formed by one to three imperfect amino acid motif(s) (R1, R2, and R3). Each repeat encodes three α-helices, with the second and third helices forming a helix-turn-helix (HTH) structure when bound to DNA. The *Z. mays* anthocyanin regulatory genes *Le* and *Cl* were the first plant MYC and MYB genes, respectively, to be isolated.

Recently, WD40 (Wd) repeat proteins were also recognized as part of the anthocyanin regulatory network (deVetten et al., 1997; Smith, 1999). They comprise a 40-residue WD motif core region that contains the glycine-histidine (GH) dipeptide at the N-terminus and the tryptophan-aspartate (WD) dipeptide at the C-terminus. The WD motif forms part of the WDR protein, a conserved β-propeller structure that acts as a stable surface to facilitate protein-protein interactions. Several plant WDR proteins involved in the anthocyanin pathway have been isolated (e.g., ZmPAC1, PhAN11, and AtTTG1).

Although R-like MYC proteins contain a DNA-binding region, thus far, there is no evidence that they directly bind DNA. R-like MYC proteins were shown to physically interact with C1 MYB proteins to coactivate anthocyanin synthesis by cooperatively binding to the promoters of the biosynthetic genes. Interactions between R-like MYC and WD40 proteins were also found. Based on this evidence (Baudry et al., 2006; Hartmann et al., 2005; Hernandez et al., 2004; Ramsay and Glover, 2005; Zimmermann et al., 2004), a working hypothesis was proposed for activation of anthocyanin structural genes by the triad of MYB, MYC, and WDR proteins (Ramsay and Glover, 2005). In a functional MYB-MYC-WD complex, the MYB factor binds directly to the cis-element of the structural gene, while R-like MYC might bind indirectly via a hypothetical R interaction protein (RIP). R-like MYC is centered in the complex that interacts with a MYB factor with WD proteins on its sides. Together, they activate the entire set of anthocyanin biosynthesis genes.

Six enzymes are generally involved in the anthocyanin biosynthesis pathway: chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), dihydroflavonol
4-reductase (DFR), anthocyanidin synthase (ANS), and UDP-glucose-flavonoid 3-O-glucosyltransferase (UGFT). CHS is the first and key regulatory enzyme of flavonoid biosynthesis, and DFR is the first committed enzyme of anthocyanin biosynthesis in the flavonoid pathway (Griesbach, 2005).

We previously confirmed that the albescient Phalaenopsis amabilis phenotype (white petals and sepals with an anthocyanin-pigmented labelum) could be complemented by transient expression of Z. mays Lc and CI regulatory genes (Griesbach and Klein, 1993; Ma et al., 2008). To further our understanding of the regulation of the anthocyanin biosynthesis pathway in Phalaenopsis, we compared the expression of the triad of anthocyanin regulatory proteins (WD40, MYC, and MYB) and two structural genes (Chs and Dfr) between species with anthocyanin-free petals (P. amabilis) and anthocyanin-containing petals (Phalaenopsis schilleriana).

Materials and Methods

Plant Material. Several commercial white-flowered P. amabilis hybrids (Kerry’s Bromeliad Nursery, Homestead, FL), purple-spotted Phalaenopsis ‘Everspring Fairy’, and purple-flowered P. schilleriana were used in this study. P. amabilis expresses an albescent phenotype with anthocyanin-free petals and sepals and an anthocyanin-pigmented labelum (Griesbach and Klein, 1993). Phalaenopsis ‘Everspring Fairy’ expresses a harlequin phenotype with white sepals, one or two large anthocyanin spots on white petals, and an anthocyanin-pigmented labelum (Tang and Chen, 2007). P. schilleriana expresses the wild-type phenotype with solid purple petals, sepals, and labellum. All plants were grown in commercial orchid greenhouses until flowering. Flowering plants were then held in the laboratory for the duration of the study. The flower buds of P. amabilis and P. schilleriana harvested for RNA isolation were 1.0 to 1.5 cm in length. Flower petals of Phalaenopsis ‘Everspring Fairy’ were harvested when flowers were just fully open.

RNA Extraction, Primer Design, Reverse Transcription (RT) Polymerase Chain Reaction (PCR), and Subcloning. Total RNAs of P. amabilis and P. schilleriana were extracted from 100 mg of flower bud petals using the RNAeasy Plant Mini Kit (Qiagen, Valencia, CA). White and purple sections from P. amabilis and P. schilleriana total RNAs using the SMART RACE Technology Kit (Clontech, Mountain View, CA) according to the manufacturer’s protocol.

RT-PCR was carried out using the Titan One Tube RT-PCR Kit (Roche Applied Science, Indianapolis). The RT-PCR reaction control was set up as described in the kit, except that the reactions were incubated at 95°C for 10 min to inactivate reverse transcriptase before proceeding to the PCR step. Different cycle numbers were tested to determine the linear phase of amplification in conventional PCR. Under linear conditions, mRNA concentration could be quantified by agarose gel band intensity. The following temperatures and times resulted in linear amplification of all the genes: 30 cycles of 94°C for 30 s, 57°C for 30 s, 72°C for 30 s. Amplification of a Phalaenopsis Actin gene (forward 5′-GGCACCATACCCTTCTACAACG-3′; reverse 5′-GATCAGTTGATGTA GTCTCATGG-3′) was used as a quantitative control between P. amabilis and P. schilleriana.

RT-PCR was carried out multiple times with similar results.

The resulting RT-PCR fragments of expected size were cloned in the TOPO vector (Invitrogen, Carlsbad, CA) and were sequenced using the BigDye Terminator v3.1 cycle sequencing Kit (Applied Biosystems, Foster City, CA). All the sequence data were edited manually based on the chromatogram output files before being compared against the GenBank nr database using blastx. DNA sequence translation and alignment were carried out using Clone Manager Suite 7 (Sci Ed Central, Cary, NC). DNA and protein sequences were compared using the standard linear and blosum62 scoring matrix, respectively, and dendrograms were constructed using the neighbor-joining algorithm in Clone Manager.

5′- and 3′-RACE primers and Race of Anthocyanin-Related Myb. Two 25-mer oligos, 5′- and 3′-RACE-Ready primers, were designed based on nine Myb clones derived from P. schilleriana: 5RMyb: 5′-CTCTCCACGATATGAAGACCACCTA-3′; 3RMyb: 5′-TACGGTTGCTTTTAAATTGCTGGAGG-3′. The 5′- and 3′-RACE-Ready cDNAs were prepared from P. schilleriana total RNAs using the SMART RACE CDNA Amplification Kit (Clontech, Mountain View, CA) according to the manufacturer’s protocol.

PCR products, obtained after 30 cycles of 94°C for 30 s, 57°C for 30 s, 72°C for 30 s, were purified from the gel. The fragments of 5′ RACE (5RMyb) and 3′ RACE (3RMyb) were then cloned into the TOPO vector (Invitrogen) and sequenced to identify the correct clones.

Full-Length PsMyb and PhRedMyb cDNA isolation. Based on the sequences of six 5RMyb and five 3RMyb clones, a primer set flanking the start codon and terminator was designed with BamHI and KpnI sites added to the respective 5′ ends of the P. schilleriana Myb (PsMyb) primers: PsMyb5′ (BamHI): 5′-GGATCCAGATGAAAGCGAAGGCTTAA-3′; PsMyb3′ (KpnI): 5′-GGTCCTAAAGATTTGAGTCAA TAAGCTAC-3′.
The full-length \textit{PsMybs} were amplified using this \textit{PsMyb} primer set and 5'-RACE-Ready cDNAs as templates. Fragments in the range of 700 to 800 bp were obtained after 30 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 60 s.

In addition, another \textit{PsMyb} 5' primer flanking the 5'-untranslated region (UTR) was also devised (\textit{PsUMybs}3, 5'-AAGTTTATCTATCTAAGCTGGCTGA-3'). \textit{PsUMybs} (full-length \textit{PsMyb} plus the 5'-UTR) were amplified as described for the full-length \textit{PsMyb}, except the 5' primer was \textit{PsUMybs} instead of \textit{PsMyb}5', and the annealing temperature was lowered to 55 °C. Fragments in the range of 800 to 900 bp were observed.

RT-PCR was carried out to amplify the full-length cDNAs from \textit{Phalaenopsis} ‘

\textit{Everspring Fairy}’ (\textit{PhredMyb}, harlequin mutant) using \textit{PsMyb} primers. Total RNA extracted from the purple sections of the \textit{Phalaenopsis} ‘Everspring Fairy’ was used as template for reverse transcription at 55 °C for 30 min, followed by 33 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 75 s.

The fragments of \textit{PsMybs}, \textit{PsUMybs}, and \textit{PhredMybs} were subsequently cloned into the TOPO vector and sequenced. Full-length \textit{Myb} cDNA fragments from \textit{P. schilleriana} and \textit{Phalaenopsis} ‘Everspring Fairy’ were identified; their GenBank accession numbers are: \textit{PsMyb}1, FJ039853; \textit{PsMyb}3, FJ039854; \textit{PsMyb}4, FJ039855; \textit{PsMyb}5, FJ039856; \textit{PsMyb}8, FJ039857; \textit{PsUMybs}1, FJ039858; \textit{PsUMybs}4, FJ039859; \textit{PsUMybs}6, FJ039860; \textit{PsUMybs}7, FJ039861; \textit{PsUMybs}8, FJ039862; \textit{PhredMyb}9, FJ039863; \textit{PhredMyb}12, FJ039864; and \textit{PhredMyb}13, FJ039865.

\textbf{Construction of \textit{PsMybs}, \textit{PsUMybs}, and \textit{PhredMybs expression constructs and their transient expression.} All constructs were placed under the control of the 35S promoter, with the backbone region of the plasmid vector identical to the S-R construct, as described in Ma et al. (2008). Biolistic transformation was carried out using a gene gun, and transient expression was monitored using fluorescence microscopy (Ma et al., 2008).

\textbf{Results and Discussion}

\textbf{Comparison of genes in \textit{P. amabilis} and \textit{P. schilleriana}.} Anthocyanin biosynthesis requires the coordinated expression of \textit{Myc}, \textit{Wd}, \textit{Chs}, \textit{Dfr}, and \textit{Myb} (Griesbach, 2005). \textit{Chs} and \textit{Dfr} are structural genes, while \textit{Myc}, \textit{Myb}, and \textit{Wd} are regulatory genes. Primers for \textit{Chs} were largely based on the \textit{Ps} gene of \textit{Phalaenopsis} species (Han et al., 2006). The expected amplified fragment size is 398 bp. RT-PCR showed no detectable difference in size or expression level between the white-flowered \textit{P. amabilis} and purple-flowered \textit{P. schilleriana} (Fig. 1). Sequencing three randomly selected clones of each revealed that they all shared highest homology to the \textit{Phalaenopsis} chalcone synthase, although nine single-nucleotide polymorphisms (SNPs) were detected, one of which caused a nonsynonymous change in one of the clones from \textit{P. schilleriana} (data not shown).

\textit{Dfr} primers were designed based on the consensus sequences derived from \textit{Dfr} genes of four orchid genera, \textit{Bromheadia} (Liew et al., 1998), \textit{Dendrobium} (Mudalige-Jayawickrama et al., 2005), \textit{Cymbidium} (Johnson et al., 1999), and \textit{Oncidium} (Hieber et al., 2005). The primers flanked a 275-bp region in the N-terminal domain that spans the section identified as responsible for the specificity of the substrate (Petit et al., 2007). From RT-PCR, there was a significant difference in \textit{Dfr} expression between \textit{P. amabilis} and \textit{P. schilleriana} (Fig. 1). \textit{Dfr} expression was significantly reduced in \textit{P. amabilis}, resulting in an accumulation of quercetin derivatives (data not shown). Dihydroquercetin is a substrate for DFR. Therefore, the white flower color of \textit{P. amabilis} is the result of a block in the biosynthetic pathway at DFR. We sequenced three clones from \textit{P. schilleriana} and one from \textit{P. amabilis}. The results indicated that all are very similar to the \textit{Dfr} from \textit{Oncidium}.

The R-like MYCs that function in transcriptional regulation associated with anthocyanin biosynthesis have two additional conserved regions other than the bHLH domain: a highly conserved N-terminal transactivation domain and a weakly conserved C-terminal domain (Purugganan and Wessler, 1994). Known anthocyanin-related \textit{Myc} from \textit{P. ×hybrida}, \textit{A. majus}, and \textit{Z. mays} were used to design the RT-PCR primers. The \textit{Myc} primers were based on the N-terminal region instead of the bHLH domain, which is shared by the entire MYC family because proteins sharing these additional motifs will more likely share similar functions.

RT-PCR (Fig. 1) yielded a ≈400-bp fragment for \textit{Myc} in \textit{P. amabilis} and \textit{P. schilleriana}. Sequencing revealed that the fragments derived from both species were similar to a \textit{Cornus} Myc-like anthocyanin regulatory protein (Fan et al., 2004). However, when compared using neighbor-joining algorithms, the randomly selected 18 clones (nine from each species) that we sequenced formed two distinct groups, with clones from each species in each group (Fig. 2). Because there was no difference in the expression of \textit{Mycs} from \textit{P. amabilis} and \textit{P. schilleriana}, we concluded that \textit{Myc} expression was not responsible for the lack of \textit{Dfr} expression in \textit{P. amabilis}.

\textit{Wd} primers were designed within the \textit{Wd} motif. The expected amplified fragment of 404 bp was found in \textit{P. amabilis} and \textit{P. schilleriana} RT-PCR (Fig. 1). Two clones of each were randomly selected for sequencing. Resulting sequences showed that both shared high homology to \textit{ZmPAC1} (Carey et al., 2004). Alignment of the four sequences displayed nine SNPs, but only one amino acid difference at the deduced protein level (data not shown). Because there was no difference in the expression of \textit{Wds} from \textit{P. amabilis} and \textit{P. schilleriana}, we concluded that \textit{Wd} expression was not responsible for the lack of \textit{Dfr} expression in \textit{P. amabilis}.

Most plant MYB proteins are members of the R2R3 family, to which anthocyanin MYB regulatory factors belong (Stracke et al., 2007). To amplify the \textit{Myb(s)}, we designed a set of primers based on the highly conserved region from the third helix of R2 to the third helix of R3 (residues 50–108, with
residue numbering based on the *Z. mays* *Myb* gene, C1). This region was selected because very little sequence similarity occurs outside of the highly conserved MYB domain, especially the third \( \alpha \)-helix within each repeat (Jiang et al., 2004).

Our RT-PCR showed no differences in expression of the expected 178-bp *Myb* fragment between *P. amabilis* and *P. schilleriana* (Fig. 1). However, sequencing of 33 randomly selected clones with an insert size of 178 bp (18 from *P. amabilis* and 15 from *P. schilleriana*) revealed differences in the fragments from the two species. Overall, three clusters were exhibited when a phylogenetic tree of the 35 deduced protein sequences was constructed (Fig. 3). The *Z. mays* C1, Pl, and P proteins were used as landmarks. Two of the groups, which possessed highest homology to DwMyb8 or ZmP, contain clones from *P. schilleriana* and *P. amabilis*. DwMyb8 and ZmP are flavonol-specific regulators (Mehrtens et al., 2005). DwMYB8 was reported to be constitutively expressed in *Dendrobium* orchid leaves, stems, and flowers at different stages in development (Wu et al., 2003). The third subgroup shared homology to ZmPl/C1, the MYB specific for anthocyanin biosynthesis. All nine sequences in this group were from *P. schilleriana*. OgMYB1 from *Oncidium* orchid flowers also shared homology with ZmPl/C1 (Chiou and Yeh, 2008).

The *Myb* gene family is the one of largest regulatory gene families in plants. Only a few members of this family regulate flavonoid biosynthesis (Stracke et al., 2007). Because we used primers derived from the most conserved region, amplification of different classes of MYBs was expected. It appears that *P. amabilis* lacks anthocyanin-specific MYB expression because only the MYBs from *P. schilleriana* shared homology with the anthocyanin-specific ZmPl/C1 cluster. Even though a band was amplified with the Myb primers from *P. amabilis*, it did not correspond to an anthocyanin-specific *Myb*. Because there was a difference in the expression of the anthocyanin-specific *Myb* from *P. amabilis* and *P. schilleriana*, we hypothesized that *Myb* expression was related to the much reduced *Dfr* expression in *P. amabilis*, and cloned the complete *Myb* gene for transient expression.

**ISOLATION OF ANTHOCYANIN-RELATED MYB RACE CLONES AND 5′-UTR CHARACTERIZATION.** To identify a region specific to *P. schilleriana*, we aligned anthocyanin-specific *Myb* sequences from *P. schilleriana* and flavonol-related (ZmP class) *Myb* from *P. amabilis* (Fig. 4). The region unique to *P. schilleriana* (25 bp, 261–285 bp from the start codon) was used to design primers for the 5′- and 3′- rapid amplification of cDNA ends (RACE). Using RACE-ready first-strand cDNA and appropriate 5′- or 3′-RACE primers, we successfully amplified the corresponding RACE fragments.

Sequences from six 5′-RACE clones (5RMyb) and five 3′-RACE clones (3RMyb) from *P. schilleriana* were obtained. A Blast and an open reading frame search against GenBank revealed the start and stop codon of the *Myb* genes (Fig. 5).

Sequences of the 5′-UTR regions of the 5RMyb clones have very distinctive characteristics: 1) the suboptimum start codon context (CAGATGA); the 5′-UTR region contains another upstream start codon that is not functional based upon the transcript and protein sequences of the functional *Myb* and MYB; 2) a perfect repeat in the start codon region; and 3) various lengths of trinucleotide (CAG) repeats in the 5′ UTR (Fig. 6).

Among the six 5RMyb clones, all contained a CAG triplet SSR with repeat numbers varying from 3 to 26. It has been...
digression of our Myb genes from the consensus sequence might provide an alternative regulatory mechanism to control the cellular availability of anthocyanin-related Mybs to ensure the precise control of its cellular concentration.

**ISOLATION OF FULL-LENGTH ANTHOCYANIN-RELATED MYBS AND THEIR TRANSIENT EXPRESSION.** Alignment of sequences in 5′- and 3′-RACE groups revealed consensus sequences that allowed us to design primers to amplify the full-length Myb cDNAs with the 5′-UTR (PsU-Myb) and without the 5′-UTR (PsMyb). We isolated five full-length PsMybs and seven full-length Mybs with 5′-UTR (PsUMybs). All encoded predicted proteins of 239 amino acids.

The 5′-UTR regions of the PsU-Myb clones revealed the same pattern as in the 5RMyb clones (Fig. 6). Interestingly, we found two different types of Myb DNA sequences that were associated with the shorter (PsU-Myb-3, -5, and -7) or longer repeats (PsU-Myb-1, -4, -6, -8) respectively. Among the three PsU-Mybs with shorter repeats, two had a premature stop codon due to a point mutation (CGA → TGA, PsU-Myb-5) or frame shift that resulted from an insertion (PsU-Myb-3, data not shown).

Alignment of both DNA and predicted protein sequence (not including the two with premature stop codons) revealed two different types of Mybs (Fig. 7). One group consisted of PsMybs-3, -8 and PsU-Myb-1, 4, 6, and -8; the other, PsMybs-1, -4, -5, and PsU-Myb-7. It seemed there was a correlation between difference in the length of trinucleotide (CAG) repeats in 5′-UTR and difference in the gene coding sequences, with the longer length of the repeat (>20) associated with the shorter DNA sequences and seven full-length Mybs with 5′-UTR (PsUMybs). All encoded predicted proteins of 239 amino acids.

Two CAG units were also present, one immediately upstream and one three nucleotides downstream of the start codon. SSRs in the 5′-UTRs have been reported to serve as protein-binding sites, hence regulating gene translation (Paul et al., 1987). We speculate that the CAG repeat binding site in our Myb 5′-UTR might not only influence its translation activity, but also dictate its proper start codon.

The location of the start codon of the 5RMyb clones was predicted based on GenBank searches using blastx. A stretch of nine nucleotides from the –9 position to the –1 position in the Mybs, ATGAACGCAG, is identical to that from the +1 position to the +9 position, which means there is an untranslated start codon (uAUG) six nucleotides upstream of the real start codon. The presence of uAUG has been shown to decrease the mRNA translational efficiency (Kochetov et al., 2002). Joshi et al. (1997) extensively surveyed over 5000 unbiased plant gene transcripts and discovered that more than 80% of them have a purine (A or G) at the −3 position and G at the +4 position. It was concluded that deviation from this norm might result in lower protein translation efficiency. In our case, both criteria were violated, with a pyrimidine at the −3 and a non-G base at the +4 positions.

Joshi’s survey also found that those transcripts with suboptimal start codon contexts often encode transcription factors, regulatory proteins/signal transducers, etc. Therefore, such
and \(Wd\) were expressed; however, \(Myb\) specific for anthocyanin biosynthesis was undetectable in \(P. amabilis\) petals and sepals. This suggests that the absence of \(Myb\) expression is responsible for the lack of DFR and results in the absence of anthocyanin pigmentation in \(P. amabilis\) petals and sepals.

Even though \(P. amabilis\) petals and sepals do not produce functional \(Myb\) and \(Df\) transcripts, functional copies of both genes must be present because \(P. amabilis\) flowers are not totally white; the labellum contains areas of deep anthocyanin pigmentation. Because the labellum produces anthocyanins, the lack of anthocyanin in the petals and sepals must be the result of differential tissue-specific gene expression.

Comparisons of anthocyanin-related \(Myb\) gene expression between \(P. schilleriana\) (solid purple) and \(P. amabilis\) (albescent phenotype) are between genetically different species. Neither an albescent \(P. schilleriana\) nor a solid purple \(P. amabilis\) exist. Even though flowers of \(P. amabilis\) (albescent phenotype) contain pigmented areas in the labellum, it is not possible to isolate quality RNA from labellum tissue. Within flowers of the \(Phalaenopsis\) ‘Everspring Fairy’ (harlequin...
phenotype), quality RNA can be isolated from both the pigmented and unpigmented areas within the petal. Therefore, harlequin flowers are ideal to evaluate anthocyanin-related *Myb* gene expression within genetically identical but differently pigmented tissue.

To further confirm that the primers we used to isolate the full-length *Mybs* from *P. schilleriana* are indeed anthocyanin-related, we carried out RT-PCR using PsMyb primers on RNAs extracted from purple and white sections of the harlequin mutant *Phalaenopsis* ‘Everspring Fairy’. With PsMyb primers, a fragment of 700 to 800 bp was only present in the purple section (Fig. 9). This fragment was isolated and several randomly selected clones were sequenced. A sequence homology search revealed that three clones (PhredMyb-9, -12, and -13) were full-length *Mybs*. PhredMybs were then compared with PsMYBs and PsUMybs (Fig. 7). One noticeable difference was a one-amino acid deletion in PhredMyb-12 and -13 with respect to amino acid 146 of PhredMyb-9 and PsMYBs/PsUMybs. The length of the MYB domain is expected to be highly conserved because it could affect Myb DNA-binding activity. However, the effects of the altered length outside of the MYB domain are not clear. PhredMyb 9 and PhredMyb 13 were made into expression constructs for transient expression. These assays showed that both were able to induce anthocyanin production (data not shown). Therefore, we concluded that the PsMyb primers were able to amplify anthocyanin-related full lengths *Mybs* and *Phalaenopsis* ‘Everspring Fairy’ petal color pattern was likely due to the differential expression of anthocyanin *Mybs*.

*Myb* expression is also responsible for anthocyanin pigmentation in *P. xhybrida* and *Oncidium* flowers. In *P. xhybrida*, a mutation in *An2* (*Myb*) reduces the expression of *An6* (*Dfr*) and results in white flowers (Quattrocchio et al., 1999). In *Oncidium* ‘Gower Ramsey’, the absence of anthocyanin pigmentation was attributed to lack of DFR and CHI (chalcone isomerase) expression (Chiu and Yeh, 2008). Several *Mybs* were expressed in *Oncidium* flowers, one of which, *OgMybB1*, was associated with anthocyanin production.

**Comparison of Phalaenopsis Mybs.** *Phalaenopsis* MYBs’ predicted protein sequences (Fig. 7) are consistent with the characterization of R2R3-*Mybs*, which are highly conserved in the MYB domain and strikingly divergent outside the MYB domain. The hallmark of plant R2R3-MYB domains is the number of Trp residues (three in R2 and two in R3) and their corresponding positions in each repeat. The three Trps in R2 of all the *Phalaenopsis* MYBs are spaced evenly 20 amino acids apart at positions 17, 37, and 57, while the two Trps are positioned 19 amino acids apart in R3. The Trp residues are part of a hydrophobic core to facilitate DNA binding. In contrast, the animal R2R3-MYB domain has three Trps in R3. In all 10 of the *Phalaenopsis* MYBs, the first Trp (at position 70) is replaced with Ile. This is different from *Dendrobium* and *A. thaliana* MYBs. Among the 21 partial *Dendrobium* MYBs analyzed (Wu et al., 2003), only two of them substitute Trp-70 with Ile, and the predominant substitution amino acid is Phe. A similar pattern is seen in *A. thaliana* Mybs (Stracke et al., 2007). Nevertheless, all use a hydrophobic amino acid to replace the Trp.

The R2 and R3 repeat is joined by a linker composed of about nine amino acids (63–71). The linker is believed to provide flexibility to position the DNA-recognition helices on their target DNA sequence, but the level of sequence conservation requirement for this region is unknown (Rabinowicz et al., 1999). At position 63, a Pro residue was found in both animal and plant Myb domains. It was highly conserved with few exceptions. This Pro can be replaced by Ala, Arg, and Ser, which resulted in so-called A-, R-, and S-type *Mybs*. The *Phalaenopsis* MYBs isolated thus far are P-type *Mybs*, which have Pro63.

**Key role of anthocyanin Mybs in conditioning pigmentation.** It is widely accepted that the anthocyanin biosynthesis pathway is regulated largely by the MYB-MYC-WD regulatory protein complex (Koes et al., 2005). WDs serve as a platform for protein-protein interactions with MYCs, which, in turn, recruit MYBs to form a transcription complex. This complex includes various members of the MYB, MYC, and WD families. The working hypothesis is that the ubiquitous WD interacts with multiple MYCs, and each MYC enlists different MYBs that eventually bind to the unique target gene to initiate gene expression.
To gain some insight as how this triad regulates anthocyanin expression in *Phalaenopsis*, we carried out similar RT-PCR surveys on the anthocyanin gene (Myc, *Dfr*, and *Chs*) expression between white and purple sections of the *Phalaenopsis* ‘Everspring Fairy’ petals as well. We used the same primers designed for the expression survey with *P. amabilis* and *P. schilleriana*. The RT-PCRs indicated that in addition to anthocyanin-specific Myb, *Dfr* transcripts were present in the purple, but not in the white, sectors (data not shown). There was no differential expression of *Chs*, *Wd*, and *Myc* between the purple and white sectors (data not shown). These results are in agreement with our results from *P. amabilis* and *P. schilleriana*. Based on the sequence data from three different *Phalaenopsis* species, WDs had the least variation at the protein level. This observation could be partially explained by the network regulation model because MYB and MYC together appeared to be sufficient to dictate the specificity of the regulatory complex. The RT-PCR result, however, strongly indicated that it is MYB that plays the key role in determining the anthocyanin production, with *Dfr* its potential targeted structural gene.

Plant MYBs have proven to be encoded by a large gene family in *Z. mays* (>80 MYBs) and *A. thaliana* (125 MYBs). It appears that this could be true in *Phalaenopsis* as well based on partial Myb sequences and several full-length cDNAs of anthocyanin-specific Mybs isolated from *P. schilleriana*. The 5′-UTR regions of these sequences displayed distinctive molecular features that might provide clues as to how the anthocyanin-specific regulatory gene Myb is itself regulated. Another possible mechanism for the regulation of Myb expression includes micro-RNA. The occurrence of virus-induced color break in some orchids is suggestive of interference in color break in some orchids is suggestive of interference in micro-RNA regulation due to interactions with viral suppressors of RNA silencing (Teycheney and Tepfer, 2001). Further studies are under way to identify anthocyanin-specific micro-RNAs.

**Literature Cited**


