

## Research Article

# A Genomic Approach to Study Anthocyanin Synthesis and Flower Pigmentation in Passionflowers

Lilian Cristina Baldon Aizza and Marcelo Carnier Dornelas

Departamento de Biologia Vegetal, Rua Monteiro Lobato 970, Instituto de Biologia, Universidade Estadual de Campinas, Cidade Universitária Zeferino Vaz, 13083-970 Campinas, SP, Brazil

Correspondence should be addressed to Marcelo Carnier Dornelas, dornelas@unicamp.br

Received 14 January 2011; Accepted 1 March 2011

Academic Editor: Francisco Cruz-Sosa

Copyright © 2011 L. C. B. Aizza and M. C. Dornelas. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most of the plant pigments ranging from red to purple colors belong to the anthocyanin group of flavonoids. The flowers of plants belonging to the genus *Passiflora* (passionflowers) show a wide range of floral adaptations to diverse pollinating agents, including variation in the pigmentation of floral parts ranging from white to red and purple colors. Exploring a database of expressed sequence tags obtained from flower buds of two divergent *Passiflora* species, we obtained assembled sequences potentially corresponding to 15 different genes of the anthocyanin biosynthesis pathway in these species. The obtained sequences code for putative enzymes are involved in the production of flavonoid precursors, as well as those involved in the formation of particular (“decorated”) anthocyanin molecules. We also obtained sequences encoding regulatory factors that control the expression of structural genes and regulate the spatial and temporal accumulation of pigments. The identification of some of the putative *Passiflora* anthocyanin biosynthesis pathway genes provides novel resources for research on secondary metabolism in passionflowers, especially on the elucidation of the processes involved in floral pigmentation, which will allow future studies on the role of pigmentation in pollinator preferences in a molecular level.

## 1. Introduction

Anthocyanins belong to a diverse group of secondary metabolites of the phenylpropanoid class, the flavonoids, which are found in different plant species. They represent some of the most important natural pigments, which are responsible for the wide range of red to purple colors present in many flowers, fruits, seeds, leaves, and stems. Besides having great economical relevance, flower and fruit pigments play an important ecological role in the animal attraction for pollination and seed dispersal, which is a spectacular example of coevolution between plants and animals [1–3].

The biosynthetic pathway of anthocyanins has been well characterized biochemically and genetically in species with different floral morphology, pigmentation pattern, and pollination syndromes such as *Petunia hybrida* [4, 5], *Matthiola* [6], *Dianthus* [7], *Eustoma* [8], *Gerbera* [9], *Zea mays* [10, 11], *Antirrhinum majus* [12], and *Ipomoea* [13, 14]. A representa-

tion of a general anthocyanin biosynthetic pathway is shown in Figure 1.

Briefly, the pathway is initiated with chalcone synthase (CHS) catalyzing the stepwise condensation of three molecules of acetate residues from malonyl-CoA with one molecule of 4-coumaroyl-CoA to form the basic structure of flavonoids (tetrahydroxychalcone), which is rapidly isomerized to the colorless naringenin by chalcone isomerase (CHI). Naringenin is then converted to dihydroflavonol by flavanone 3-hydroxylase (F3H). Dihydroflavonol 4-reductase (DFR), which is a specific enzyme for the anthocyanin synthesis, catalyses the production of leucoanthocyanidins from dihydroflavonols, which can be hydroxylated on the 3' or 5' position of the B-ring by flavonoid 3'-hydroxylase (F3'H) to produce dihydroquercetin or by flavonoid 3'5'-hydroxylase (F3'5'H) to form dihydromyricetin. Subsequently, leucoanthocyanidin oxidase/anthocyanidin synthase (LDOX/ANS) is responsible for the formation of the anthocyanidins

from the colorless leucoanthocyanidins. GT enzymes (O-glucosyltransferases) represent the final step in anthocyanin biosynthesis: anthocyanidins are converted in differentially “decorated” anthocyanin molecules [15, 16]. Biochemical approaches have demonstrated that all anthocyanin pigments are derived from one of three aglycones: pelargonidin, cyanidin, and delphinidin. The main determinants of the apparent color of these pigments are the hydroxylation and methylation patterns, as well as the number and type of sugars on the beta ring of the flavonoid molecule [1, 3, 17–19].

Figure 1 depicts a generalized anthocyanin biosynthesis pathway. At least, two groups of genes are required for anthocyanin biosynthesis: the first group is represented by the structural genes encoding enzymes for the production of the flavonoid precursors, as well as those involved in the formation of particular (“decorated”) anthocyanin molecules. The second group includes the genes encoding regulatory factors that control the expression of structural genes which are mainly overexpressed by complexes formed by MYB and basic helix-loop-helix (bHLH) transcription factors that include WDR (WD40 repeats) proteins [2, 4, 15, 16, 20–23].

There are about 600 *Passiflora* species widely distributed in tropical and subtropical regions. Some *Passiflora* species have economical importance due to the production of fruits (passionfruit) or use as ornamentals. Nevertheless, a large number of *Passiflora* species are rare and/or endangered, as the environment of their diversity center has been increasingly degraded by human activities [24]. An enormous floral diversity is observed among *Passiflora* species, including variation in color, size, morphology, and fusion of floral organs. These and other floral characteristics, including evolutionary innovations such as the presence of coronal filaments and an androgynophore, are indicative of the wide range of pollination syndromes found in the genus [24]. Wide passionflowers may be pollinated by insects (bees and wasps), hummingbirds, and bats [24]. The most striking feature of floral variation among passionflowers is the wide range of pigmentation patterns of the corona filaments. Most of the floral pigments in *Passiflora* are different types of anthocyanin molecules [25, 26]. Among all *Passiflora* species, *P. edulis* Deg and *P. suberosa* L. are of particular interest, because they are model *Passiflora* species for which expressed sequence tags (ESTs) were produced within the frame of the “PASSIOMA” Project [27]. *P. edulis* Deg flowers are pollinated by large bees of genus *Xylocopa*. These flowers are about 8–12 cm wide, and their coronas contain multiple series of purplish filaments with white tips. The flowers of *P. suberosa* L. are small (2–3 cm wide) and show two morphologically distinct series of corona filaments: the outer series is greenish, and the inner series is formed by smaller purple filaments. The flowers of *P. suberosa* are pollinated by wasps [28].

We are particularly interested in the characterization of genes involved in the anthocyanin biosynthetic pathway of these two *Passiflora* species. With this aim, we searched for putative *Passiflora* genes responsible for flower pigmentation, using the key proteins known to be involved in the different enzymatic steps of anthocyanin biosynthesis as baits to search for expressed sequence tags (ESTs) in the PASSIOMA database.

## 2. Material and Methods

**2.1. Searching *Passiflora* ESTs Homologous to Anthocyanin Biosynthetic Genes.** The clustered expressed sequence tags (ESTs) from the PASSIOMA Project database [27] were used as a primary source of data for our analyses. These sequences were assembled from ESTs obtained from the sequencing of several *P. edulis* or *P. suberosa* cDNA libraries, made from floral buds at different developmental stages (see [27] for details on library construction, sequencing, and database structure). Nucleotide sequences and their respective deduced amino acid sequences from genes known to be involved in anthocyanin biosynthesis (see Figure 1) were obtained from the National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/>). Searches for putative homolog sequences in the PASSIOMA database were conducted using the tBLASTN module that compares the consensus amino acid sequence with a translated nucleotide sequences database [29]. We generally used *Arabidopsis thaliana* or *Petunia hybrida* as query consensus sequences as the anthocyanin biosynthesis pathways in these model species are more thoroughly studied at the molecular level [30–32]. All sequences in the PASSIOMA database that exhibited a significant alignment (*e*-value lower than  $10^{-5}$ ) with the query were retrieved from the PASSIOMA database.

The clusterization of all reads identified using a given query sequence was performed using the CAP3 algorithm [33] from the BioEdit software [34]. The novel cluster consensus sequences obtained were reinspected for the occurrence of conserved motives using InterProScan [35] and were compared to NCBI databases using BLAST [29]. Sequences that did not show the main motives present in the query sequence were discarded. Validated sequences were then included in phylogenetic analyses.

**2.2. Comparison of the Amino Acid Sequences and Phylogenetic Analysis.** All amino acid sequences were aligned by CLUSTALX software using default parameters [36]. The obtained alignments were eventually corrected by hand and imported into the molecular evolutionary genetics analysis (MEGA) software [37]. Phylogenetic trees were obtained using parsimony and/or genetic distance calculations (in the later case using pairwise deletion option and with the Poisson correction model). Neighbor-joining [38] and Bootstrap (with 10,000 replicates) trees were also constructed.

## 3. Results

The cDNA libraries of the PASSIOMA Project were obtained from mRNA extracted from floral buds at different developmental stages, and it is expected that all EST sequences correspond to genes expressed during *Passiflora* flower development [27]. This sequence search detected a total of 75 *Passiflora* EST sequences, 34 of them corresponding to *P. edulis* sequences and 41 of them corresponding to sequences derived from *P. suberosa* libraries. When submitted to the CAP3 algorithm and detailed comparison of their deduced amino acid sequences, the number of valid clusters was reduced to 15, potentially corresponding to 15 different genes.

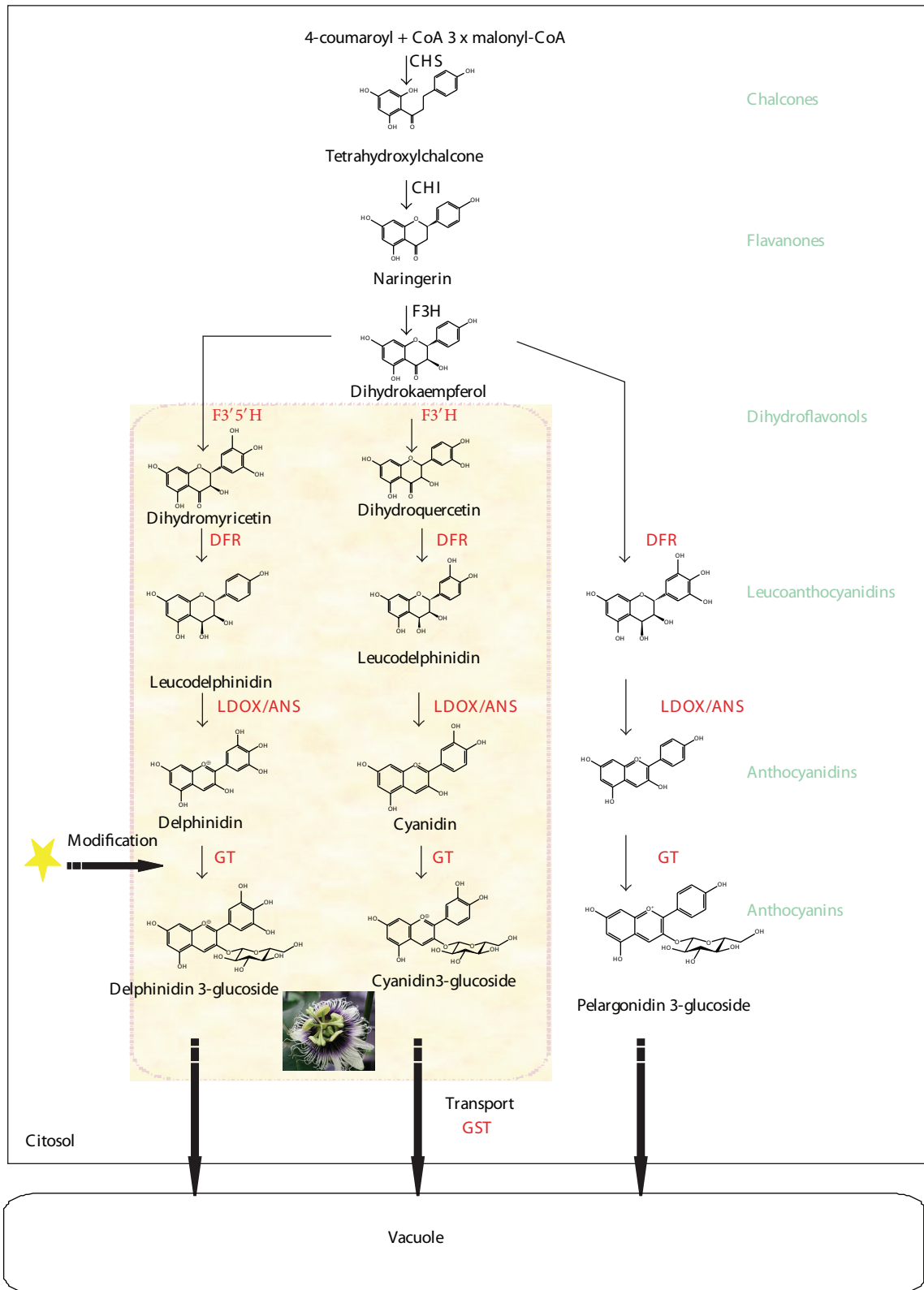


FIGURE 1: Schematic representation of the anthocyanin biosynthetic pathway (adapted from [16]). Enzymes are indicated in red, and classes of compounds are in green. Anthocyanidin is further modified with glycosyl, acyl, or methyl groups, resulting in the “decorated” anthocyanin. In this case, UF3GT is responsible for the glycosylation of anthocyanidins. The proposed anthocyanin biosynthetic pathway for *Passiflora edulis* is highlighted by the colored background. CHS: chalcone sintase; CHI: chalcone isomerase; F3H: flavanone 3-hydroxylase; F3'H: flavanone 3'-hydroxylase; F3'5'H: flavanone 3'5'-hydroxylase; DFR: dihydroflavonol 4-reductase; LDOX/ANS: leucoanthocyanidin dioxygenase/anthocyanidin synthase; GT: glucosyltransferase; GST: glutathione S-transferase.

TABLE 1: Putative *Passiflora* homologs of genes encoding elements of the anthocyanin biosynthetic pathway.

Enzyme	<i>Passiflora</i> AS*	First BLAST hit	e-value	ID/SM
CHS	PACEPE3010G11.g	ABD24222 <i>CHS Populus alba</i>	$7e^{-72}$	85/90
	PACEPE3014B06.g	ABC86919 <i>CHS Populus alba</i>	$9e^{-67}$	77/84
	PACEPE3007G06.g	XP_002305446 <i>CHS-like Populus trichocarpa</i>	$7e^{-121}$	84/92
	PACEPE3023H10.g	XP_002326830 <i>CHS-like Populus trichocarpa</i>	$3e^{-110}$	82/91
DFR	PACEPS7017D03.g	AAQ62589 <i>CHS3 Glycine Max</i>	$1e^{-98}$	82/88
	PACEPE3003G04.g	XP_002307667 <i>DFR2 Populus trichocarpa</i>	$1e^{-95}$	82/94
GT	PACEPE3030G03.g	XP_002532899 <i>UFGT Ricinus communis</i>	$6e^{-31}$	53/70
	PACEPS7021H07.g	XP_002518725 <i>UFGT Ricinus communis</i>	$8e^{-57}$	56/72
GST	PACEPE3013H01.g	AF048978 <i>GST Glycine Max</i>	$7e^{-52}$	79/90
	PACEPE3007A05.g	XM_002519342 <i>GST theta Ricinus communis</i>	$4e^{-52}$	77/89
	PACEPE3018F08.g	ADB11335 <i>GSTF7 phi Populus trichocarpa</i>	$2e^{-82}$	68/83
	PACEPS4006H06.g	ADB11332 <i>GSTF4 phi Populus trichocarpa</i>	$2e^{-61}$	63/78
MYB	PACEPS7023B03.g	AF243378 <i>GST 23 Glycine max</i>	$1e^{-51}$	80/88
	PACEPS7022E07.g	XP_002530824 <i>R2R3 MYB Ricinus communis</i>	$7e^{-80}$	88/91
WD40	PACEPE3007G07.g	XP_002512788 <i>WD-repeat protein Ricinus communis</i>	$3e^{-124}$	92/96

Abbreviations: CHS: chalcone synthase; DFR: dihydroflavonol 4-reductase; GT: glucosyltransferase and GST: glutathione S-transferase. Using the BLASTp algorithm [29].

\*AS: assembled sequence. Codes refer to the longest cDNA clone. PACEPE: *Passiflora edulis*; PACEPS: *Passiflora suberosa*. ID/SM: identity/similarity (both based on the amino acid sequence) with the first BLAST hit.

When the validated amino acid sequences obtained from the PASSIOMA database were compared to other plant protein sequences in the public databases, the first BLAST hits generally corresponded to *Populus* and *Ricinus* sequences. This was expected, as *Passiflora* and these genera belong to the same order (Malpighiales) and are considered to be closely related [39].

We obtained assembled EST sequences corresponding to genes of the following genes families: CHS, DFR, GT, GST, MYB, and WD40 (see Table 1). Therefore, we used 15 *Passiflora* assembled sequences from the PASSIOMA database and a selected set of genes from divergent plant species from the public databases to explore their evolutionary relationships. The obtained sequence comparison alignments allowed the construction of phylogenetic trees for each of these families of genes involved in the different enzymatic steps of the anthocyanin pathway.

The similarities among all genes identified in this study and those reported from other plant species were assembled in Table 1 and ranged from 70% (PACEPE3030G03.g; representing a putative member of the GST, glutathione S-transferase superfamily) to 96% (PACEPE3007G07.g; potentially encoding a WD40 protein).

Some of these gene sequences showed significant similarity to elements required for early or late steps of the pathway; others putatively encode regulatory proteins involved in the control of the spatial and temporal patterns of pigmentation,

while others are responsible for intracellular transport of the anthocyanin molecules. The role of each of these genes in the anthocyanin biosynthesis and the probable implications for the understanding of the *Passiflora* flower pigmentation are presented in the Discussion.

### 3.1. Identification and Phylogenetic Analysis of *Passiflora* Genes Potentially Involved in Anthocyanin Biosynthesis and Transport

**3.1.1. Chalcone Synthases (CHSs).** We have found 5 *Passiflora* assembled sequences (5 putative genes) encoding enzymes of the CHS family: PACEPE3010G11.g, PACEPE3014B06.g, PACEPE3007G06.g, PACEPE3023H10.g and PACEPS7017D03.g. These sequences are expected to encode proteins with 231, 158, 254, 237, and 222 amino acids, respectively. The deduced CHS proteins showed more than 80% similarity to CHSs of other plant species (Table 1). To determine the phylogenetic relationship of different CHSs, we aligned protein sequences from a diverse range of plant species (moss, ferns, gymnosperms and angiosperms), cyanobacterium (*Synechococcus* sp.) and *Passiflora* representatives of the CHS superfamily (Figure 2). The phylogenetic tree was resolved in three clades. These three clades were highly supported with 100% bootstrap values. The *Passiflora* proteins were consistently positioned into different clades.

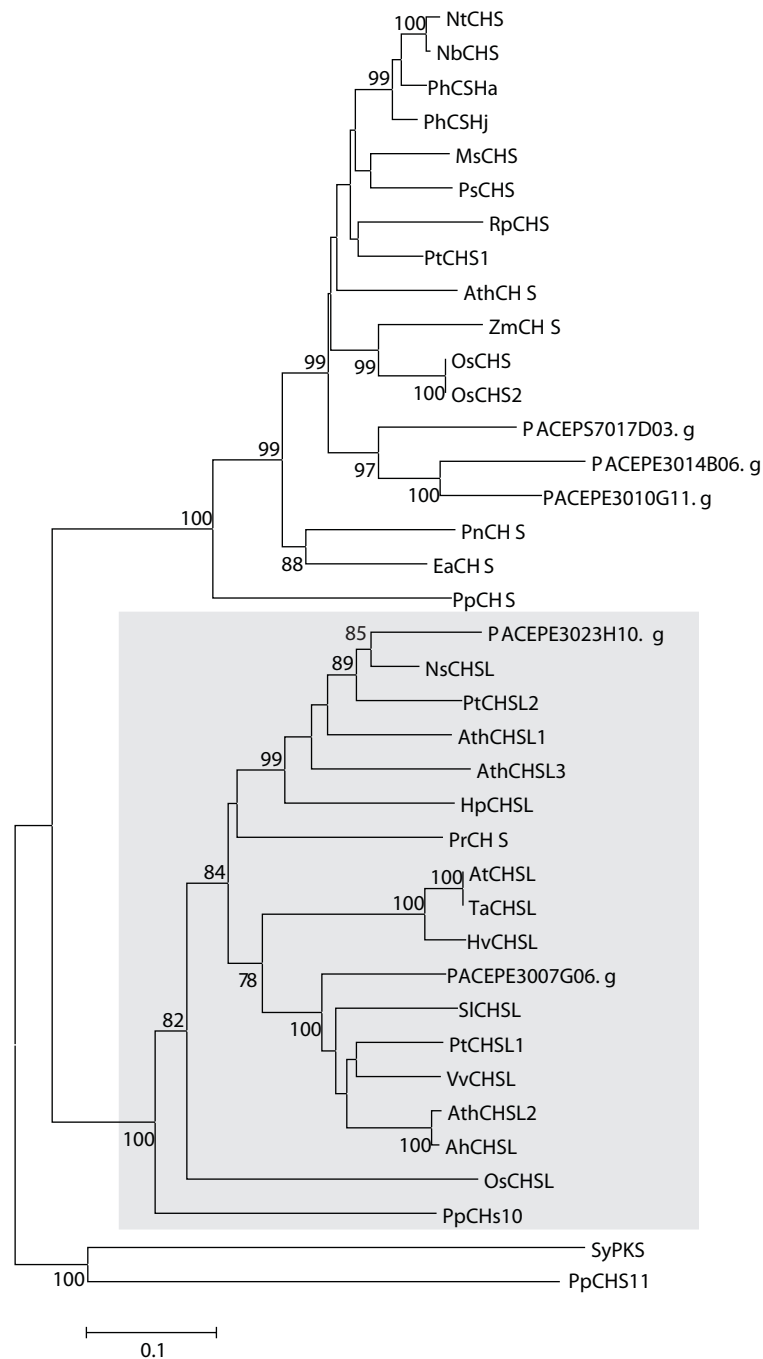


FIGURE 2: A Neighbor-joining phylogenetic tree of chalcone synthase (CHS) amino acid sequences. The cluster containing all anther-specific CHS-like enzymes is highlighted. Bootstrap values from 1,000 replicates were used to assess the robustness of the trees. Only bootstrap values above 75% are indicated at the nodes. Accession numbers for genes from other species are given in Supplementary data.

One of these monophyletic clades (highlighted in Figure 2) contains all the anther-specific CHS-like genes (ASCLs; [40, 41]). The remaining sequences, including three *Passiflora* members, were clustered in the other sister clade together with all CHS genes from seed plants.

**3.1.2. Dihydroflavonol 4-Reductases (DFR).** A single *Passiflora* cDNA sequence of 850 bp encoding a predicted protein

of 204 amino acids showed significant  $e$ -value ( $1e^{-95}$ ) and 94% similarity to a *Populus* DFR sequence (Table 1). Figure 3 shows an alignment of the deduced amino acid sequence of the *Passiflora* DFR with some other plant sequences containing an NADP-binding domain, considered the region of substrate preference of DFR enzymes [42, 43]. Additionally, the *Passiflora* DFR showed an aspartic acid residue at position 134, as it is observed for the *Petunia* and *Populus* proteins,

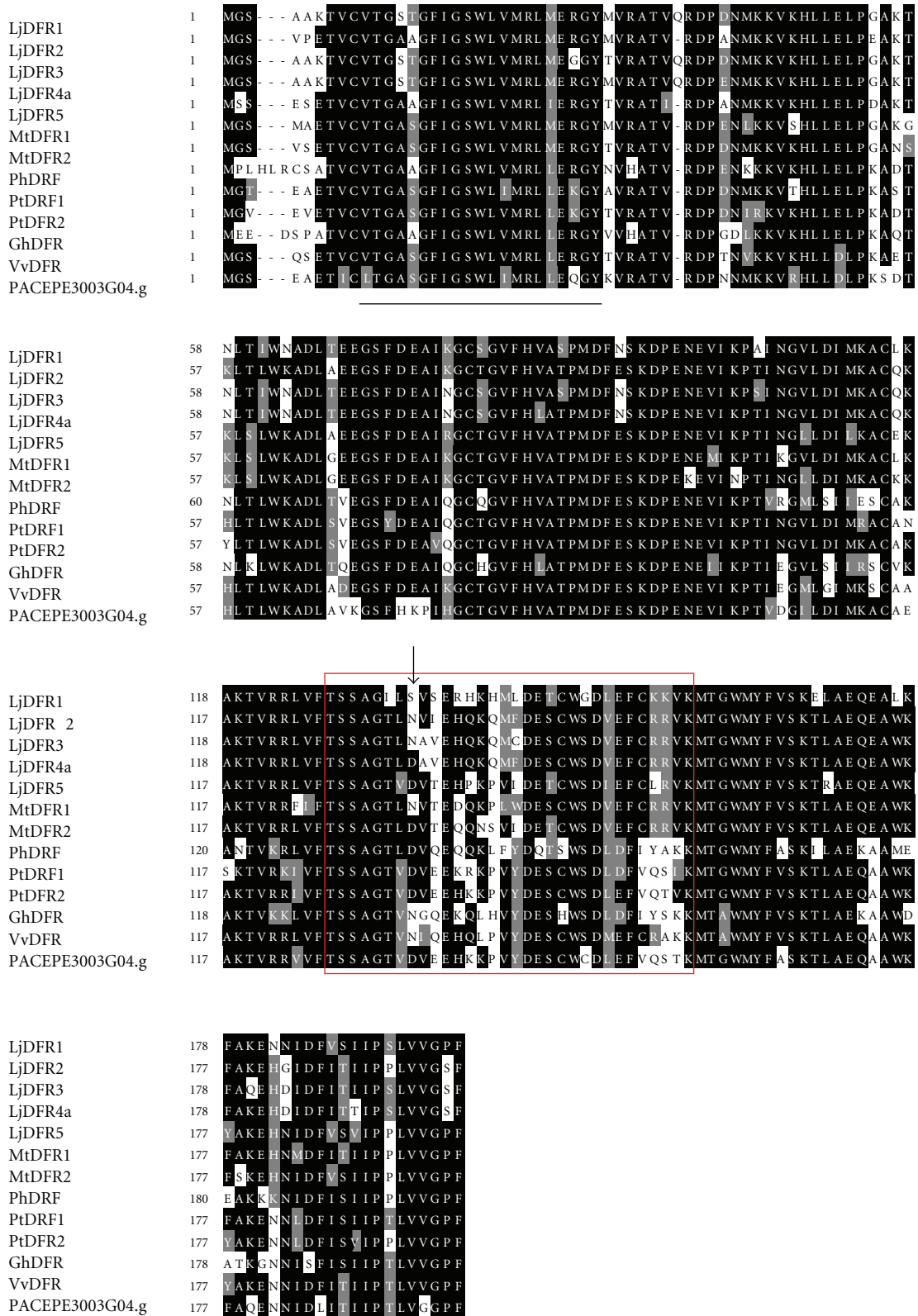


FIGURE 3: Multiple sequence alignment of the *Passiflora* sequence with other plant DFR sequences. The identical and similar residues are highlighted on a black and gray background, respectively. NADP-binding domain is underlined. Boxed amino acids have been considered to control the substrate specificity of DFR enzyme [40], and the amino acid residue (indicated by an arrowhead) is especially important for this specificity [41]. The alignment was performed using CLUSTALX and BOXSHADE program.

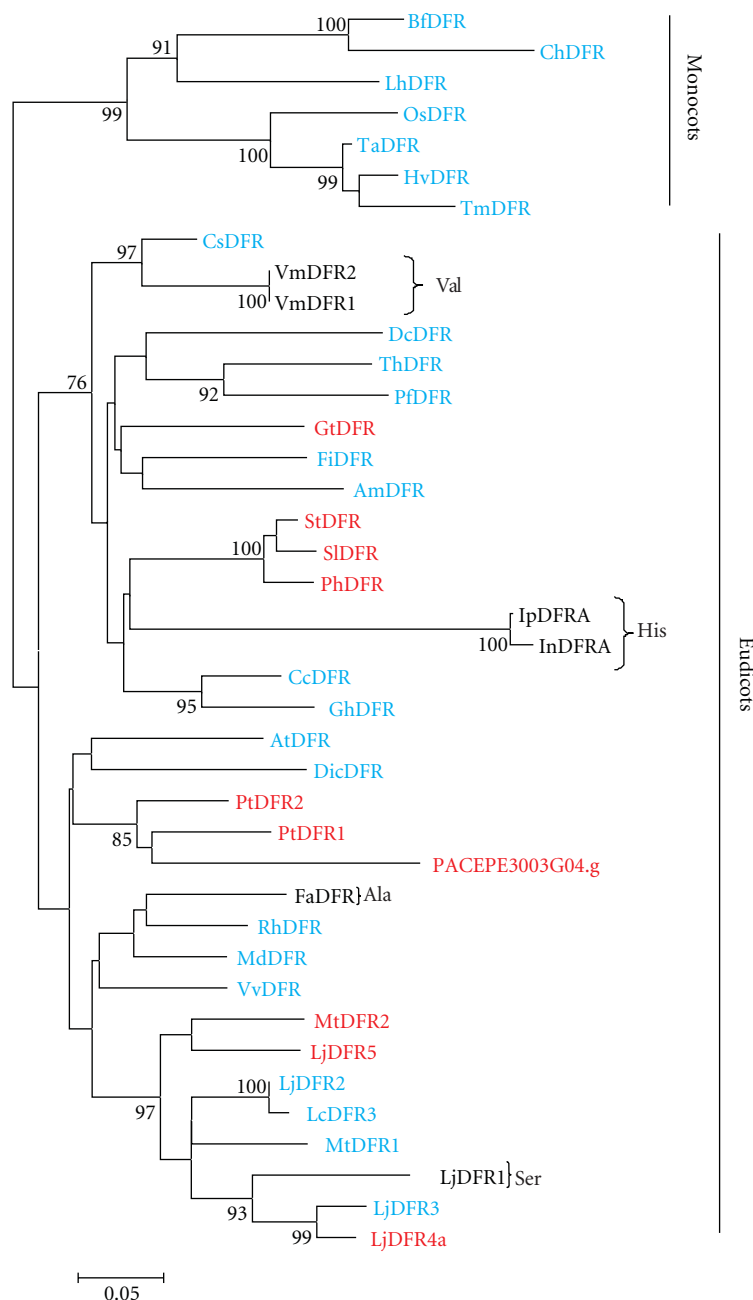


FIGURE 4: A Neighbor-joining phylogenetic tree of dihydroflavonol 4-reductase (DFR) amino acids sequences. Bootstrap values from 1,000 replicates were used to assess the robustness of the trees. Only bootstrap values above 75% are indicated at the nodes. Asn-type DFRs, Asp-type DFRs, and DFRs of neither Asn nor Asp-type are indicates in blue, red, and black, respectively [42]. Accession numbers for genes from other species are given in Supplementary data which are available online at doi:10.4061/2011/37157.

whereas *Gerbera* and some *Lotus* DFR show an asparagine residue at the same position (Figure 3). We adopted the terminology suggested by Shimada and coworkers [44] to designate the conserved motifs present in the DFR sequence.

A neighbor-joining tree was constructed based on the alignment DFR sequences shown in Figure 3. The monocots and eudicots DFRs were positioned separately. While monocot DFR genes formed one clade, the eudicot DFR sequences diverged into two clades. Clearly, Asn-type DFRs are found

in a larger number of species. On the other hand, Asp-type DFRs are restricted to some species, including *Passiflora* and *Populus* (Figure 4).

**3.1.3. Glucosyltransferases (GT).** We identified two *Passiflora* EST clones, PACEPE3030G03.g and PACEPS7021H02.g, encoding proteins with sequence similarity to *Ricinus communis* glucosyltransferases (Table 1). The first cDNA sequence contained an ORF specifying a 124 amino acid

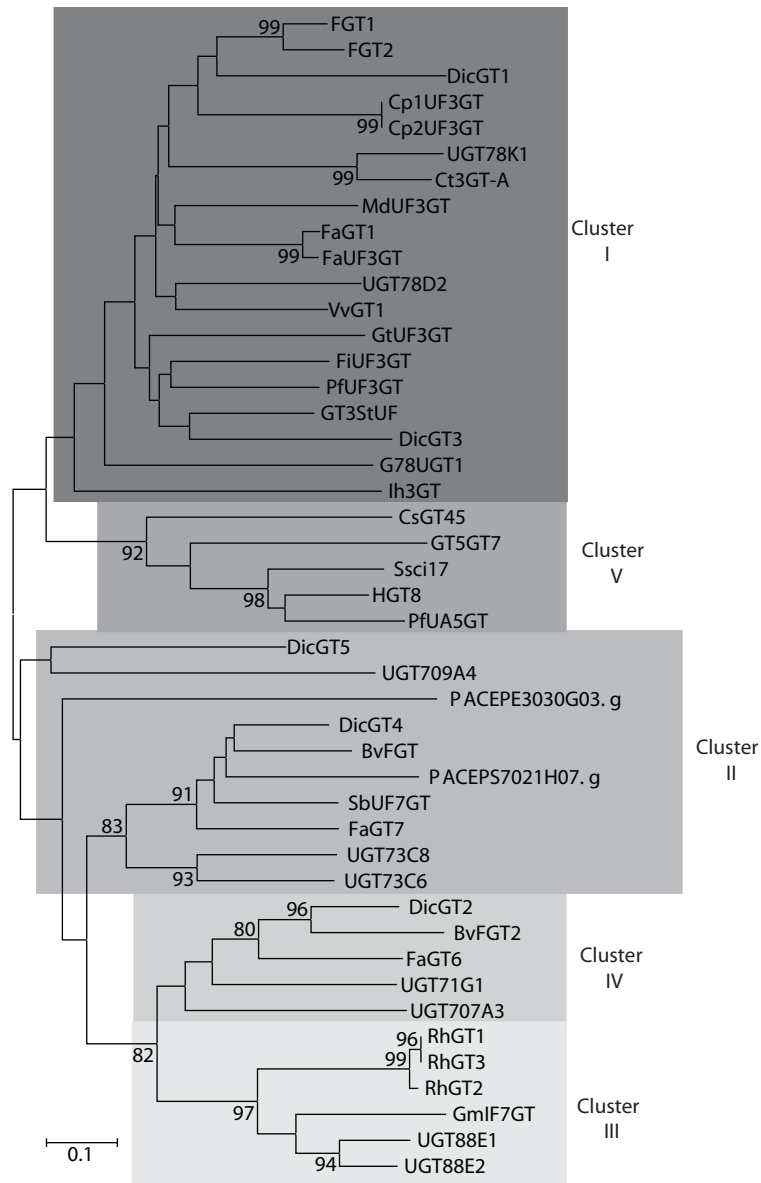


FIGURE 5: A Neighbor-joining phylogenetic tree of glucosyltransferase (GT) amino acid sequences. Bootstrap values from 1,000 replicates were used to assess the robustness of the trees. Only bootstrap values above 75% are indicated at the nodes. Accession numbers for genes from other species are given in Supplementary data.

protein, and the second cDNA encoded a protein of 200 amino acid residues. These putative *Passiflora* GT proteins were compared with those GT enzymes described by Kovinick and colleagues [45] and retrieved from the NCBI database. The obtained phylogenetic tree resulted in five clades, according to their *in vitro* substrate specificities [45]. Phylogenetic analysis revealed that the *Passiflora* sequences were positioned within the Cluster II proteins (Figure 5).

**3.1.4. Glutathione S-Transferases (GSTs).** We have identified five *Passiflora* sequences representing putative members of the GST family. Each member was represented by a single EST sequence. Comparison of these deduced GST protein sequences with those in the GenBank database revealed

homology with multifunctional GSTs from *Populus*, *Ricinus*, and *Glycine* spp (see Table 1). Phylogenetic relationships among the putative *Passiflora* GSTs and family members of other plant species were established (Figure 6). Based on sequence similarity, the five *Passiflora* putative GSTs were grouped into three clades. PACEPE3018F08.g, PACEPS4006H06.g, and PACEPS7023B03.g are type I GSTs, PACEPE3007A05.g is a type II GST, and PACEPE3013H01.g is a type III GST [46].

We could not find any putative homologs to chalcone isomerases (CHI), flavanone 3-hydroxylases (F3H), and anthocyanidin synthases (ANS; see Figure 1) in the PASSIOMA database. Three EST sequences were identified corresponding to a putative flavonoid 3-O-hydroxylase (F3'H) gene, and



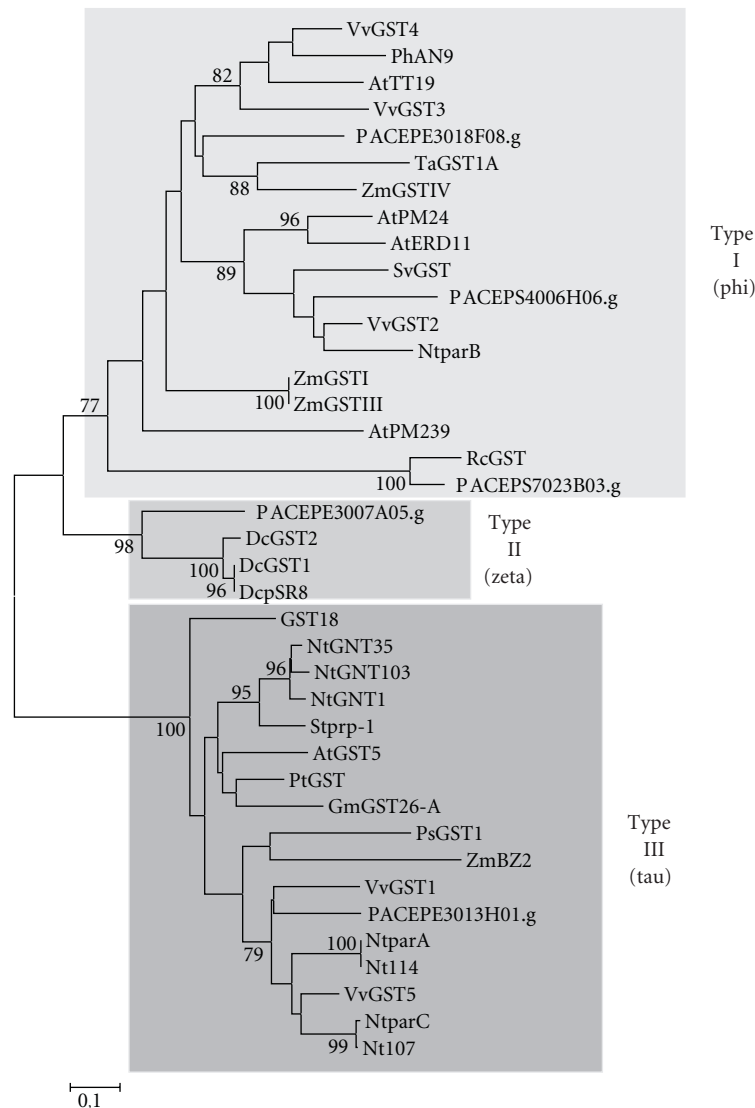


FIGURE 6: A Neighbor-Joining phylogenetic tree of glutathione S-transferase (GST) amino acids sequences with three types representing phi, tau, and zeta classes. Phi and tau are plant-specific GSTs. Bootstrap values from 1,000 replicates were used to assess the robustness of the trees. Only bootstrap values above 75% are indicated at the nodes. Accession numbers for genes from other species are given in Supplementary data.

one sequence was found that showed significant homology to genes encoding flavonoid 3-5-O-hydroxylases (F3'5'H; data not show). As these sequences were incomplete at their 5' end, they were not considered in our analyses.

**3.2. Identification and Phylogenetic Analysis of Passionflower Genes Potentially Involved in Spatially and Temporally Patterning Anthocyanin Deposition.** Based on the searches in the PASSIOMA database, we identified one potential homolog for an MYB transcription factor of the R2R3 class. The *P. suberosa* cDNA clone PACEPS7022E07.g encodes a protein of 132 amino acids showing 91% similarity to the *Ricinus communis* R2R3 MYB. On the other hand, PACEPE3007G07.g is a putative *P. edulis* WD40 gene of 886 bp encoding 291 amino

acid residues showing 96% similarity to an *R. communis*, WD40 (Table 1).

Figure 7 shows an alignment of the deduced PACEPS7022E07.g protein sequence with 17 other plant anthocyanin-related R2R3-MYB, indicating the presence of a conserved DNA-binding domain, designated as the R2R3 domain. All sequences analyzed also contained a second conserved amino acid motif in the R3 repeat (red box), important for the interaction between MYB and bHLH proteins in *Arabidopsis* [48]. The four specific residues required for this interaction in maize [49] are also indicated by the arrows in Figure 7. The third conserved motif appears to be ANDV (blue box) in the R3 repeat of all eudicot R2R3-MYB proteins related to anthocyanin biosynthesis.

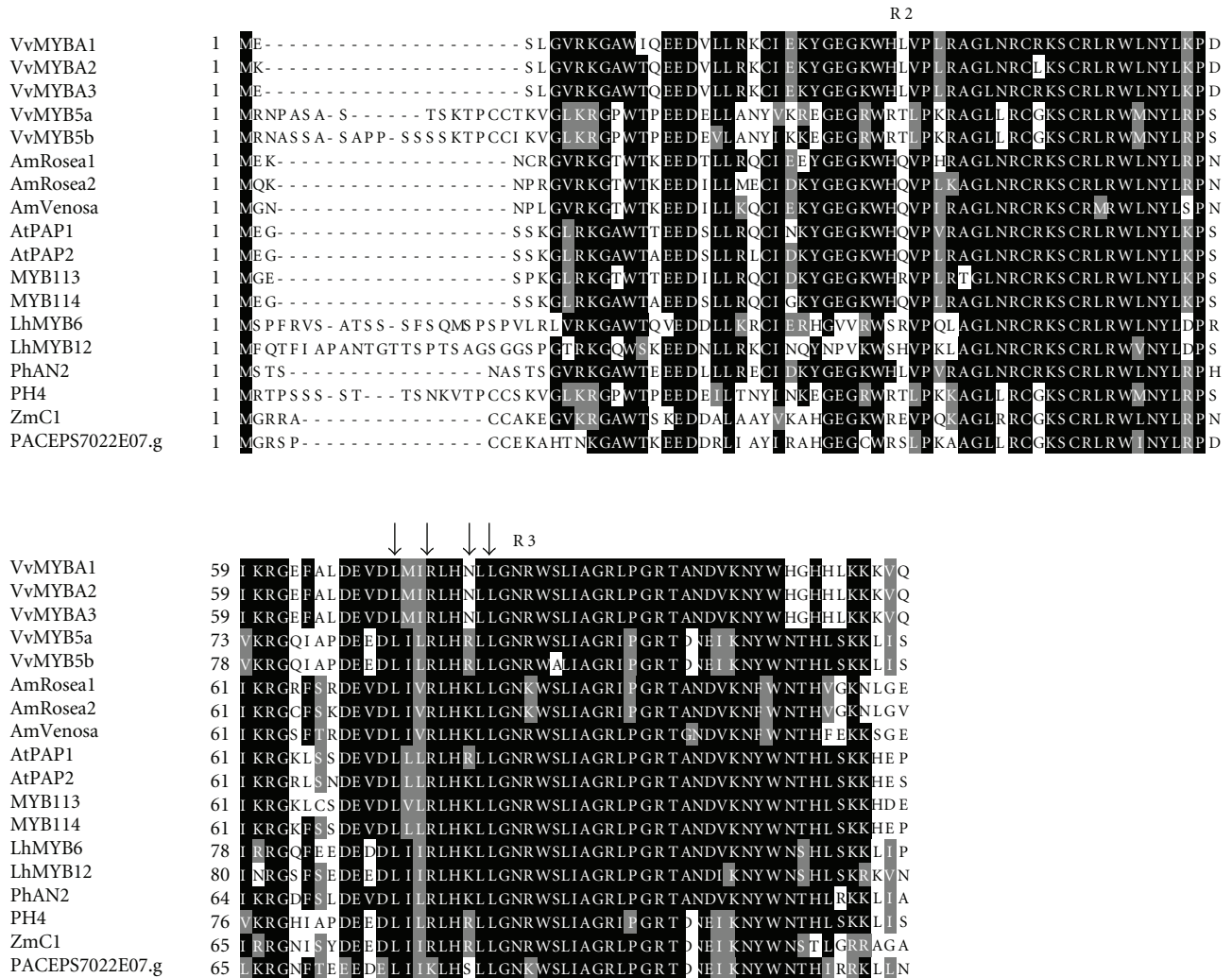


FIGURE 7: Multiple sequence alignment of the R2R3 MYB domains involved in anthocyanin production including the deduced amino acid sequence of *Passiflora suberosa*. R2R3 repeats refer to two imperfect repeats of the MYB domain. The identical and similar residues are highlighted on a black and gray background, respectively. Red box shows the R/B like bHLH interacting motif in the R3 repeat [45], and arrows indicate four specific residues of maize C1 required for interaction with a bHLH cofactor R [46]. Blue box shows a conserved motif in the R2R3 repeats for eudicots MYB related to the anthocyanin pigments [47]. The alignment was performed using CLUSTALX and BOXSHADE program.

A phylogenetic tree of selected plant R2R3-MYB transcription factors, including PACEPS7022E07.g, was constructed using the alignment of the conserved R2R3 repeats (Figure 8). The *Passiflora* sequence was placed within the clade including ZMC1 (*Zea mays*), PhPH4 (*Petunia hybrida*), VvMYB5a, and VvMYB5b (*Vitis vinifera*), which are known to be involved in the regulation of the anthocyanin pathway in these species [49–51].

Sequence comparison of selected plant WD40 proteins with the sequence obtained from *P. edulis* indicated that the four WD repeats are highly conserved among all species analyzed (Figure 9). Phylogenetic analysis of these amino acid sequences confirmed that *P. edulis* WD40 grouped together with *Ricinus communis* WD40 and found to be more related to other dicot proteins (Figure 10).

No putative homologs to bHLH transcription factors were found in the PASSIOMA database.

#### 4. Discussion

Flavonoid pathway results in the production of a range of flavonoid compounds, including anthocyanins (Figure 1). CHS is the first enzyme in the phenylpropanoid pathway and is encoded by members of a plant-specific multigene family of polyketide synthases. Nevertheless, genes belonging to the CHS family have been recently described to occur in some microorganisms (*Azotobacter vinelandii*; [52] and *Neurospora crassa*; [53]) and, therefore, indicate CHS functions might have evolved previous to the divergence of land plants. Thus, the biological functions of some of

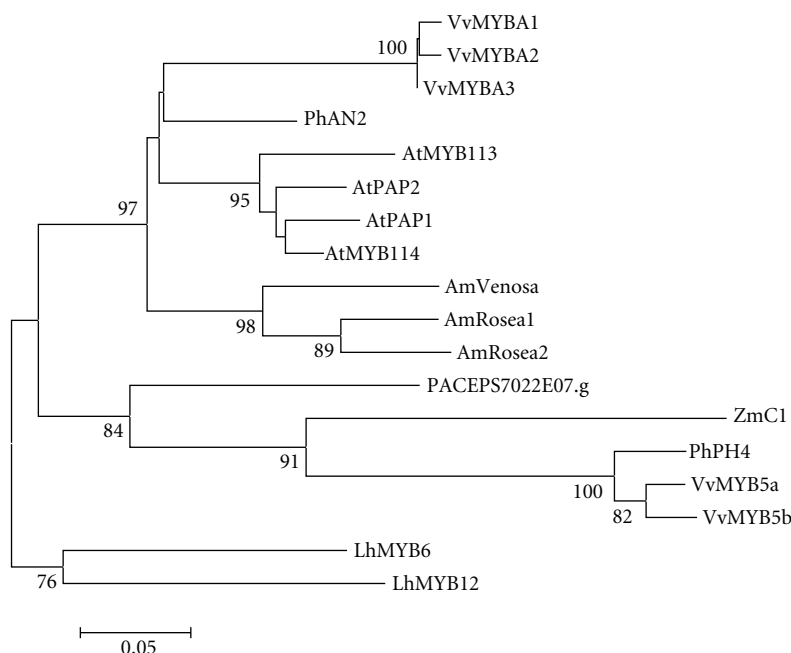


FIGURE 8: A Neighbor-joining phylogenetic tree of plant R2R3 MYB sequences. Bootstrap values from 1,000 replicates were used to assess the robustness of the trees. Only bootstrap values above 75% are indicated at the nodes. Accession numbers for genes from other species are given in Supplementary data.

the CHS superfamily members are clearly important to plant adaptation. CHS proteins are collectively linked to the biosynthesis of different plant products with diverse functions such as UV protection, defense against pathogens, pigment biosynthesis, and pollen fertility [54, 55].

Sequence analysis indicated that two *Passiflora* CHS deduced proteins belong to a small distinct group of chalcone synthases that includes angiosperm and gymnosperms homologs to anther-specific chalcone synthase-like genes (ASCLs; highlighted in Figure 2). Furthermore, all ASCLs form a monophyletic clade. Recently, ASCLs transcripts were detected within the tapetum cells during microspore stage in wheat [56]. These genes apparently have important roles in anther development and in pollen fertility [40, 41, 56].

The remaining three *Passiflora* CHSs were clustered together in a sister clade containing all seed plant CHS genes. Their products are considered key in the biosynthesis of flavonoids. These include *CHSA* and *CHSJ* genes, known to be expressed in floral tissues, and involved in floral pigmentation in petunia [30, 31, 57]. Moreover, two nonchalcone genes, divergent from the typical CHSs, formed a separate clade. The *SyPKS* gene from cyanobacterium encodes an enzyme of the thiolase superfamily [58], whereas the function of the *PpCHS11* gene (from *Physcomitrella patens*) may resemble more the most recent common ancestor of all plant CHSs than do other members of the plant CHS superfamily [55].

We do not have identified putative genes encoding CHI enzymes. Besides the general limitations and drawbacks of the EST-based approach, another possible explanation may be because the rapid isomerization of chalcone to form

naringen and the fact that even in the absence of a functional CHI enzyme, chalcone can spontaneously isomerize to form naringenin [15].

DFR is an enzyme catalysing the reduction of three dihydroflavonols: dihydromyricetin (DHM), dihydroquercetin (DHQ), and dihydrokaempferol (DHK) into colorless leucoanthocyanidins. These are further converted to delphinidin, cyaniding, and pelargonidin (Figure 1). The synthesis of three different anthocyanidins is mainly determined by the enzymes activities of two hydroxylases: F3'OH and F3'5'OH. The first converts DHK to DHQ and F3'5'OH converts DHK to DHM [15].

In some plant species, DFR displays distinct substrate specificity in according to the hydroxylation pattern of anthocyanin molecule [30]. A hypothesis to determine substrate specificity was proposed based on the amino acid sequence alignment of *Petunia* DFR with others plants. The alignment indicated a variable region that controls substrate recognition. Naturally, *Petunia hybrida* does not produce orange flowers, because the DFR enzyme cannot use dihydrokaempferol as substrate to produce pelargonidin, due to an aspartic acid residue at the 134th position [30, 42], as it was also observed for *Passiflora* (Figure 3), thus converting dihydroquercetin to leucocyanidin and, more efficiently, the reduction of dihydromyricetin to leucodelphinidin [30, 59]. On the other hand, some *Gerbera* genotypes have an asparagine residue at this same position and can utilize three dihydroflavonols as substrates of DFR, consequently producing orange to red colored flowers [9, 30]. Thus, the flower color is partly determined by alteration of a single amino acid that changes the substrate specificity of the DFR enzyme.

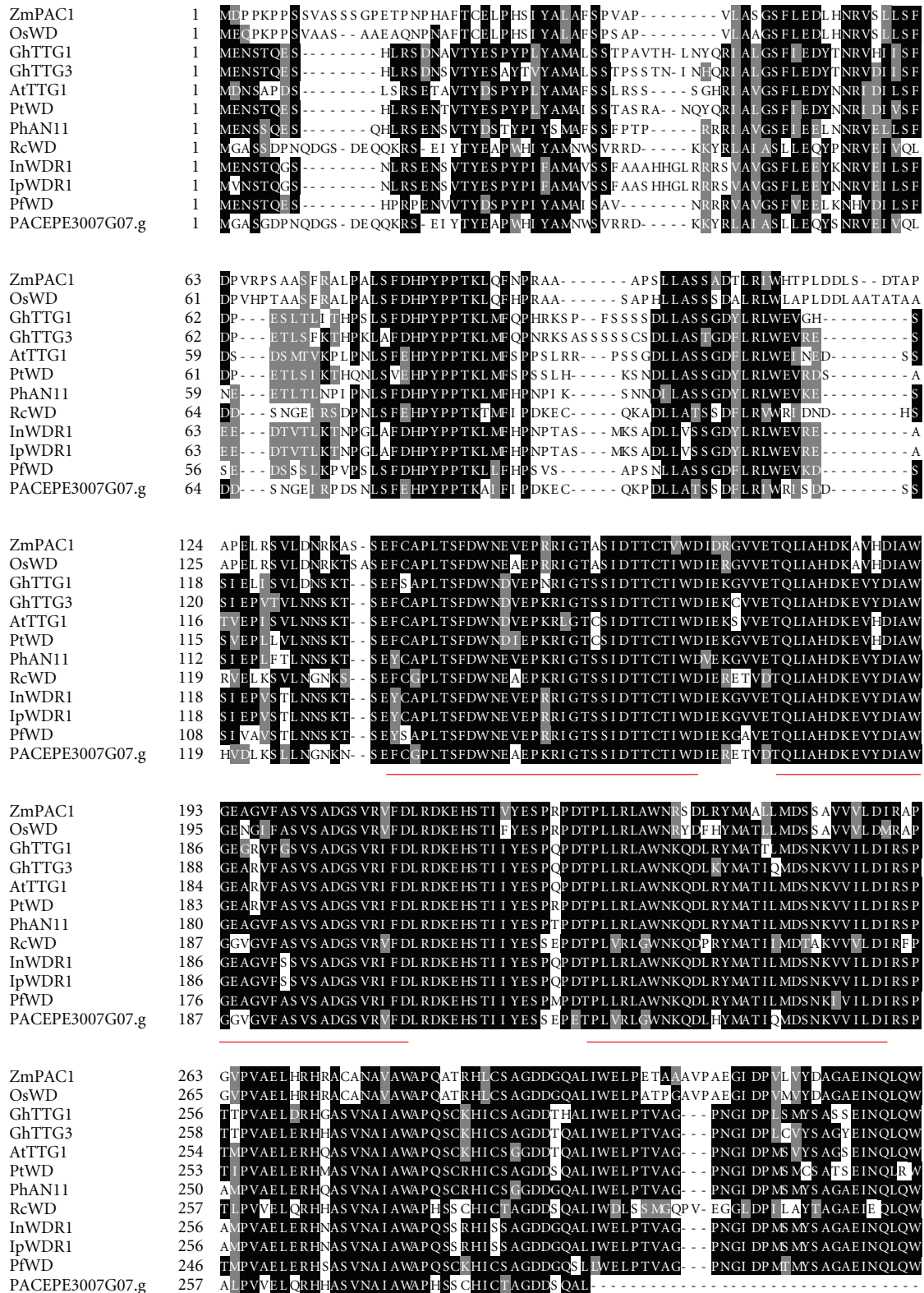


FIGURE 9: Multiple sequence alignment of the WD40 proteins involved in anthocyanin production, including the deduced amino acid sequence of the *Passiflora edulis* WD40. The identical and similar residues are highlighted on a black and gray background, respectively. Four conserved WD repeat domain are underlined in red. The alignment was performed using CLUSTALX and BOXSHADE program.

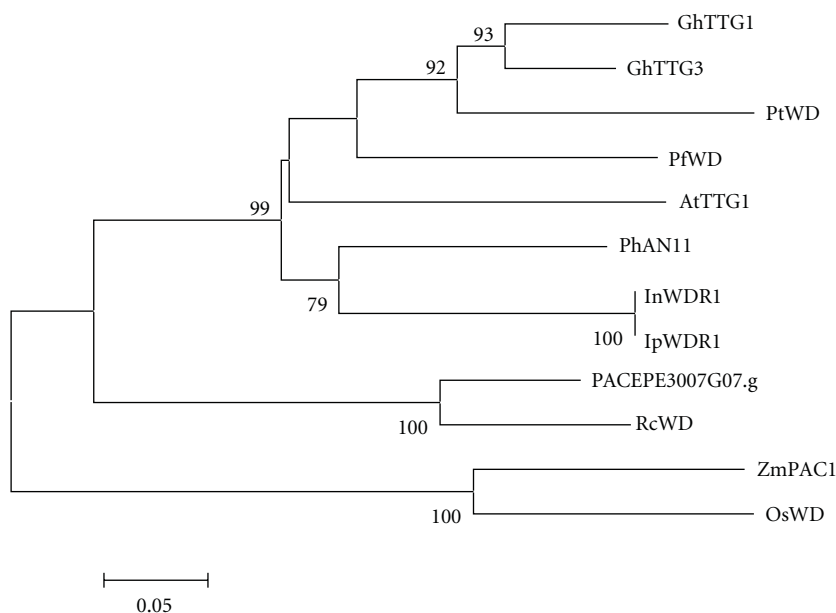


FIGURE 10: A Neighbor-joining phylogenetic tree of plant WD 40 proteins. Bootstrap values from 1,000 replicates were used to assess the robustness of the trees. Only bootstrap values above 75% are indicated at the nodes. Accession numbers for genes from other species are given in Supplementary data.

Almost all anthocyanidins undergo several modifications, which vary across species and involve enzymes of the glucosyltransferase, methyltransferase, and acyltransferase families. The most common is glycosylation of the 3-position of anthocyanidins (represented in Figure 1) to produce stable anthocyanin molecules [15, 30, 31, 60]. UDP-glucose:flavonoid 3-O-glucosyltransferase (3GT) belongs to a large multigene glucosyltransferases (GTs) family, representing the final step in anthocyanin biosynthesis.

In this work, we adopted the classification of the GTs into clusters according to Kovinic and colleagues [45]. Cluster I groups includes 3GTs enzymes. Cluster II includes GTs with multiples substrates preferences, generally for chalcones, flavones and flavonols but not anthocyanidins. Enzymes from Cluster III have isoflavone 7-O and anthocyanidin 3,5-O-GT activities. Cluster IV glycosylates flavonol and isoflavonol substrates and Cluster V have anthocyanin 5-O and/or flavone 7-O-UGT enzymes [45]. Our results indicated that the obtained *Passiflora* glucosyltransferase gene sequences were grouped in Cluster II, together with other family members that show a high catalytic specificity for more than one class of flavonoid substrates (Figure 5). DicGT5 (from *Dianthus caryophyllus*) glycosylates a chalcononaringenin 2'-O-glucosyltransferase [61], whereas the *Beta vulgaris* GT has a favonoid-7, 4'-O-betanidin-5-O-glucosyltransferase activity [62]. Both GTs have non-anthocyanidin substrate specificity. Despite these results, obviously neither GT substrate specificity, nor *in vivo* function of the *Passiflora* GTs can be predicted solely based on amino acid sequence similarities and must be experimentally determined.

Anthocyanin biosynthesis has been demonstrated to occur predominantly in the cytosol, but these pigments are exclusively accumulated in the vacuole of epidermal

cells [20]. Transport of pigments to the vacuoles requires a glutathione S-transferase and a specific carrier protein localized in the vacuolar membrane. GSTs are multifunctional proteins encoded by a large familiar present in all cellular organisms. Plants GSTs are classified on the basis of sequence identity into four classes: phi, tau, theta, and zeta [46]. The two small zeta and theta classes include GSTs from animals and plants, while the phi and tau classes are plant-specific. Several studies have confirmed the involvement of GSTs in the vacuolar transport of anthocyanins. PhAN2 (from *Petunia*), ZmBZ2 (from maize), and AtTT19 (from *Arabidopsis*) are GST proteins involved in anthocyanin transport [30–32, 63–65].

To characterize their phylogenetic relationships, the deduced amino acid sequences from the *Passiflora* putative GSTs were compared with other plant GST sequences, including the ones mentioned above. Figure 6 shows that the *Passiflora* GSTs are included into three different clades: three sequences were positioned in the same clade of PhAN9 and AtTT19 (phi class), whereas one sequence was grouped together with ZmBZ2 (tau class; [66]). Although of these known proteins belong to distinct GST clades, they perform similar functions [63–65].

Interestingly, PACEPE3007A05.g was clustered with carnation (*Dianthus caryophyllus*) GST type II (zeta class) which is associated to petal senescence in response to ethylene [67, 68].

At the moment, we can classify the *Passiflora* GSTs into type I (phi), type II (zeta), and type III (tau). At least, four of them might be involved in the anthocyanin pathway and PACEPE3007A05.g might be related to other biological processes related to flower development such as those observed for the carnation GST.

In all analyzed species, the spatial and temporal expression of the structural genes of the anthocyanin biosynthetic pathway is controlled by regulatory genes, which interfere with the intensity and pattern of anthocyanin biosynthesis [15]. MYBs, basic helix-loop-helix (bHLH) transcription factors and WD40 proteins form a transcriptional complex for the activation of the structural genes [4, 12, 20, 47, 69, 70]. MYBs and bHLHs proteins are coded by large multigene families, and those associated with anthocyanin biosynthesis are characterized by a conserved DNA-binding domain consisting of two imperfect repeats (named R2R3), and a specific bHLH domain, respectively. These two gene families have been extensively studied in model plants such as *Arabidopsis* and maize [48, 49, 71].

A multiple sequence alignment of the R2R3 domains of selected MYB proteins known to be involved in anthocyanin biosynthesis regulation, and the deduced amino acid sequence of PACEPS7022E07.g confirmed the presence of the conserved R2R3-MYB domain in this *P. suberosa* sequence (Figure 7) as well as that of a second conserved domain in the R3 repeat (red box, Figure 7), which is known to be necessary for the interaction between MYB and bHLH transcription factors [48, 49]. Additionally, a third motif in the R3 repeat (ANDV, blue box in Figure 7) represents a conserved motif shared among all eudicot MYBs involved in the anthocyanin biosynthesis [72].

The phylogenetic tree obtained using the alignment shown in Figure 7 is presented in Figure 8 and indicates LhMYB6 and LhMYB12 clustered outside the eudicot clade. These two genes regulate anthocyanin biosynthesis in the flowers of lily (*Lilium* hybrid), a monocot [73]. One clade is formed exclusively by eudicot anthocyanin regulators (PhAn2, AtPAP1, AtPAP2, AmROSEA1, and AmROSEA2; [12, 71, 74–77]. Curiously, one regulator of the anthocyanin in maize (a monocot), ZmC1 was positioned in the same clade of other dicot members such as PhPH4 (from *Petunia*), VvMYB5a, and VvMYB5b (from *Vitis*), as well as the *Passiflora* R2R3-MYB sequence. PhPH4 is expressed in the petal epidermis and activates vacuolar acidification in petunia [50]. VvMYB5a and VvMYB5b genes are involved in the regulation of anthocyanin biosynthesis during grape berry development [51].

WD40 proteins are highly conserved and can be found in organisms that do not biosynthesize anthocyanins as algae, fungi, and animals [78, 79]. In plants, these proteins are involved in a plethora of developmental and biochemical functions. As an example, the *Arabidopsis* TRANSPARENT TESTA GLABRA 1 (TTG1), which is a WD40 protein, is involved in regulating trichome formation, anthocyanin biosynthesis, seed coat pigmentation, and seed coat mucilage production. A common feature of WD40 repeat proteins is that they facilitate protein-protein interactions between the MYB and bHLH proteins [22, 79].

The alignment of the *Passiflora* WD40 protein sequence with other known WD40s from different plant species revealed the presence of conserved WD40 motifs in the C-terminal region (Figure 9). The phylogenetic tree constructed based on this alignment is shown in Figure 10. The results indicated that the monocot sequences ZmPAC1 and OsWD

clustered together, whereas the eudicot WD40s known to function as anthocyanin regulators were grouped into a different clade, with *Passiflora* WD40 being closely related to the *Ricinus communis* protein (RcWD, Table 1 and Figure 10). Although WD40 proteins are required to regulate anthocyanins and proanthocyanidin together with MYB and bHLH transcription factors, their potential involvement in other biological processes is enormous, therefore, it is premature to say what functions PACEPE3007G07.g might perform in *Passiflora*.

The fact that no putative homologs to bHLH transcription factors were found in the PASSIOMA database may reflect the high degree of novelty of most of the libraries of the PASSIOMA project indicating that full gene expression spectra was not completely achieved [27]. Perhaps a more deep sequencing effort would reveal that such homologs are indeed expressed in *Passiflora* flowers, as these elements are generally essential to MYB-WD40 protein complex stability [30–32].

## 5. Conclusions and Perspectives

We took the first steps toward the understanding of the molecular processes involved in the biosynthesis of anthocyanins in *Passiflora* that could account for the differences in pollinator preferences found in the genus. We identified 15 putative coding sequences derived from two distinct *Passiflora* species (*P. edulis* and *P. suberosa*) expressed in developing flower buds and potentially involved in the anthocyanin biosynthetic pathway. Comparisons of deduced amino acid sequences from the 15 *Passiflora* cDNAs with selected sequences from other plant species revealed strong similarity with genes that encode key elements involved in the biosynthesis (8 sequences), transcriptional regulation (2 sequences), and transport (5 sequences) of anthocyanin molecules.

Needed research concerning the determination of temporal and spatial expression patterns of all these *Passiflora* putative anthocyanin-related genes presented here are already ongoing in our group. We expect that future work on the manipulation of their expression patterns, using transgenic approaches, will help us to unravel important aspects relating anthocyanin biosynthesis, flower pigmentation, and flower pollination in rapidly changing tropical environments.

## Acknowledgment

The authors acknowledge FAPESP and CNPq (Brazil) for financial support.

## References

- [1] J. B. Harborne, *The Flavonoids: Advances in Research Since 1986*, Chapman & Hall/CRC, New York, NY, USA, 1st edition, 1994.
- [2] E. Grotewold, “The genetics and biochemistry of floral pigments,” *Annual Review of Plant Biology*, vol. 57, pp. 761–780, 2006.

- [3] M. D. Rausher, "Evolutionary transitions in floral color," *International Journal of Plant Sciences*, vol. 169, no. 1, pp. 7–21, 2008.
- [4] F. Quattrocchio, J. Wing, K. van der Woude et al., "Molecular analysis of the *anthocyanin2* gene of *Petunia* and its role in the evolution of flower color," *Plant Cell*, vol. 11, no. 8, pp. 1433–1444, 1999.
- [5] C. Spelt, F. Quattrocchio, J. N. M. Mol, and R. Koes, "*Anthocyanin1* of *Petunia* encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes," *Plant Cell*, vol. 12, no. 9, pp. 1619–1631, 2000.
- [6] W. Heller, G. Forkmann, L. Britsch, and H. Grisebach, "Enzymatic reduction of (+)-dihydroflavonols to flavan-3,4-cis-diols with flower extracts from *Matthiola incana* and its role in anthocyanin biosynthesis," *Planta*, vol. 165, no. 2, pp. 284–287, 1985.
- [7] K. Stich, T. Eidenberger, F. Wurst, and G. Forkmann, "Enzymatic conversion of dihydroflavonols to flavan-3,4-diols using flower extracts of *Dianthus caryophyllus* L. (carnation)," *Planta*, vol. 187, no. 1, pp. 103–108, 1992.
- [8] K. M. Davies, J. M. Bradley, K. E. Schwinn, K. R. Markham, and E. Podivinsky, "Flavonoid biosynthesis in flower petals of five lines of *lisianthus* (*Eustoma grandiflorum* Grise.)," *Plant Science*, vol. 95, no. 1, pp. 67–77, 1993.
- [9] Y. Helariutta, P. Elomaa, M. Kotilainen, P. Seppänen, and T. H. Teeri, "Cloning of cDNA coding for dihydroflavonol-4-reductase (DFR) and characterization of dfr expression in the corollas of *Gerbera hybrida* var. Regina (Compositae)," *Plant Molecular Biology*, vol. 22, no. 2, pp. 183–193, 1993.
- [10] E. Grotewold, B. J. Drummond, B. Bowen, and T. Peterson, "The myb-homologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset," *Cell*, vol. 76, no. 3, pp. 543–553, 1994.
- [11] J. M. Hernandez, G. F. Heine, N. G. Irani et al., "Different mechanisms participate in the R-dependent activity of the R2R3 MYB transcription factor C1," *The Journal of Biological Chemistry*, vol. 279, no. 46, pp. 48205–48213, 2004.
- [12] K. Schwinn, J. Venail, Y. Shang et al., "A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the Genus *antirrhinum*," *Plant Cell*, vol. 18, no. 4, pp. 831–851, 2006.
- [13] Y. Morita, M. Saitoh, A. Hoshino, E. Nitasaka, and S. Iida, "Isolation of cDNAs for R2R3-MYB, bHLH and WDR transcriptional regulators and identification of c and ca mutations conferring white flowers in the Japanese morning glory," *Plant and Cell Physiology*, vol. 47, no. 4, pp. 457–470, 2006.
- [14] K. I. Park, N. Ishikawa, Y. Morita, J. D. Choi, A. Hoshino, and S. Iida, "A bHLH regulatory gene in the common morning glory, *Ipomoea purpurea*, controls anthocyanin biosynthesis in flowers, proanthocyanidin and phytomelanin pigmentation in seeds, and seed trichome formation," *Plant Journal*, vol. 49, no. 4, pp. 641–654, 2007.
- [15] T. A. Holton and E. C. Cornish, "Genetics and biochemistry of anthocyanin biosynthesis," *Plant Cell*, vol. 7, no. 7, pp. 1071–1083, 1995.
- [16] R. Koes, W. Verweij, and F. Quattrocchio, "Flavonoids: a colorful model for the regulation and evolution of biochemical pathways," *Trends in Plant Science*, vol. 10, no. 5, pp. 236–242, 2005.
- [17] K. Davies, "Plant pigments and their manipulation," *Annual Plant Reviews*, vol. 14, p. 368, 2004.
- [18] H. S. Lee and V. Hong, "Chromatographic analysis of anthocyanins," *Journal of Chromatography*, vol. 624, no. 1–2, pp. 221–234, 1992.
- [19] J. B. Harborne, T. J. Mabry, and H. Mabry, *The Flavonoids*, Academic Press, New York, NY, USA, 1975.
- [20] J. Mol, E. Grofelowd, and R. Koes, "How genes paint flowers and seeds," *Trends in Plant Science*, vol. 3, no. 6, pp. 212–217, 1998.
- [21] B. Winkel-Shirley, "Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology," *Plant Physiology*, vol. 126, no. 2, pp. 485–493, 2001.
- [22] N. A. Ramsay and B. J. Glover, "MYB-bHLH-WD40 protein complex and the evolution of cellular diversity," *Trends in Plant Science*, vol. 10, no. 2, pp. 63–70, 2005.
- [23] L. Lepiniec, I. Debeaujon, J. M. Routaboul et al., "Genetics and biochemistry of seed flavonoids," *Annual Review of Plant Biology*, vol. 57, pp. 405–430, 2006.
- [24] T. Ulmer and J. M. MacDougal, *Passiflora, Passion Flowers of the World*, Timber Press, Cambridge, UK, 2004.
- [25] M. M. Halim and R. P. Collins, "Anthocyanins of *Passiflora quadrangularis*," *Bulletin of the Torrey Botanical Club*, vol. 97, no. 5, pp. 247–248, 1970.
- [26] L. KidØy, A. M. Nygård, Ø. M. Andersen, A. T. Pedersen, D. W. Aksnes, and B. T. Kiremire, "Anthocyanins in fruits of *Passiflora edulis* and *P. suberosa*," *Journal of Food Composition and Analysis*, vol. 10, no. 1, pp. 49–54, 1997.
- [27] M. C. Dornelas, S. M. Tsai, and A. P. M. Rodriguez, "Expressed sequence tags of genes involved in the flowering process of *Passiflora* spp.," in *Floriculture, Ornamental and Plant Biotechnology*, J. A. Teixeira da Silva, Ed., pp. 483–488, Global Science Books, London, UK, 2006.
- [28] I. G. Varassin, J. R. Trigo, and M. Sazima, "The role of nectar production, flower pigments and odour in the pollination of four species of *Passiflora* (Passifloraceae) in south-eastern Brazil," *Botanical Journal of the Linnean Society*, vol. 136, no. 2, pp. 139–152, 2001.
- [29] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.
- [30] T. Nakatsuka, Y. Abe, Y. Kakizaki, S. Yamamura, and M. Nishihara, "Production of red-flowered plants by genetic engineering of multiple flavonoid biosynthetic genes," *Plant Cell Reports*, vol. 26, no. 11, pp. 1951–1959, 2007.
- [31] M. Hanumappa, G. Choi, S. Ryu, and G. Choi, "Modulation of flower colour by rationally designed dominant-negative chalcone synthase," *Journal of Experimental Botany*, vol. 58, no. 10, pp. 2471–2478, 2007.
- [32] S. Martens, A. Preuß, and U. Matern, "Multifunctional flavonoid dioxygenases: flavonol and anthocyanin biosynthesis in *Arabidopsis thaliana* L.," *Phytochemistry*, vol. 71, no. 10, pp. 1040–1049, 2010.
- [33] X. Huang and A. Madan, "CAP3: a DNA sequence assembly program," *Genome Research*, vol. 9, no. 9, pp. 868–877, 1999.
- [34] T. A. Hall, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT," *Nucleic Acids Symposium Series*, vol. 41, no. 41, pp. 95–98, 1999.
- [35] S. Hunter, R. Apweiler, T. K. Attwood et al., "InterPro: the integrative protein signature database," *Nucleic Acids Research*, vol. 37, no. 1, pp. D211–D215, 2009.
- [36] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins, "The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by

- quality analysis tools," *Nucleic Acids Research*, vol. 25, no. 24, pp. 4876–4882, 1997.
- [37] S. Kumar, K. Tamura, and M. Nei, "MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment," *Briefings in Bioinformatics*, vol. 5, no. 2, pp. 150–163, 2004.
- [38] N. Saitou and M. Nei, "The neighbor-joining method: a new method for reconstructing phylogenetic trees," *Molecular Biology and Evolution*, vol. 4, no. 4, pp. 406–425, 1987.
- [39] M. J. Moore, P. S. Soltis, C. D. Bell, J. G. Burleigh, and D. E. Soltis, "Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 10, pp. 4623–4628, 2010.
- [40] A. Ageez, Y. Kazama, R. Sugiyama, and S. Kawano, "Male-fertility genes expressed in male flower buds of *Silene latifolia* include homologs of anther-specific genes," *Genes and Genetic Systems*, vol. 80, no. 6, pp. 403–413, 2005.
- [41] C. Jiang, C. K. Schommer, S. Y. Kim, and D. Y. Suh, "Cloning and characterization of chalcone synthase from the moss, *Physcomitrella patens*," *Phytochemistry*, vol. 67, no. 23, pp. 2531–2540, 2006.
- [42] M. Beld, C. Martin, H. Huits, A. R. Stuitje, and A. G. M. Gerats, "Flavonoid synthesis in *Petunia hybrida*: partial characterization of dihydroflavonol-4-reductase genes," *Plant Molecular Biology*, vol. 13, no. 5, pp. 491–502, 1989.
- [43] E. T. Johnson, S. Ryu, H. Yi, B. Shin, H. Cheong, and G. Choi, "Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase," *Plant Journal*, vol. 25, no. 3, pp. 325–333, 2001.
- [44] N. Shimada, R. Sasaki, S. Sato et al., "A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome," *Journal of Experimental Botany*, vol. 56, no. 419, pp. 2573–2585, 2005.
- [45] N. Kovinich, A. Saleem, J. T. Arnason, and B. Miki, "Functional characterization of a UDP-glucose:flavonoid 3-O-glucosyltransferase from the seed coat of black soybean (*Glycine max* (L.) Merr.)," *Phytochemistry*, vol. 71, no. 11–12, pp. 1253–1263, 2010.
- [46] D. P. Dixon, A. Laphorn, and R. Edwards, "Plant glutathione transferases," *Genome Biology*, vol. 3, no. 3, pp. 1–10, 2002.
- [47] Y. Borovsky, M. Oren-Shamir, R. Ovadia, W. De Jong, and I. Paran, "The A locus that controls anthocyanin accumulation in pepper encodes a MYB transcription factor homologous to Anthocyanin2 of *Petunia*," *Theoretical and Applied Genetics*, vol. 109, no. 1, pp. 23–29, 2004.
- [48] I. M. Zimmermann, M. A. Heim, B. Weisshaar, and J. F. Uhrig, "Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins," *Plant Journal*, vol. 40, no. 1, pp. 22–34, 2004.
- [49] E. Grotewold, M. B. Sainz, L. Tagliani, J. M. Hernandez, B. Bowen, and V. L. Chandler, "Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 25, pp. 13579–13584, 2000.
- [50] F. Quattrocchio, W. Verweij, A. Kroon, C. Spelt, J. Mol, and R. Koes, "PH4 of *petunia* is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway," *Plant Cell*, vol. 18, no. 5, pp. 1274–1291, 2006.
- [51] L. Deluc, J. Bogs, A. R. Walker et al., "The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries," *Plant Physiology*, vol. 147, no. 4, pp. 2041–2053, 2008.
- [52] N. Funa, H. Ozawa, A. Hirata, and S. Horinouchi, "Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in *Azotobacter vinelandii*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 16, pp. 6356–6361, 2006.
- [53] N. Funa, T. Awakawa, and S. Horinouchi, "Pentaketide resorcylic acid synthesis by type III polyketide synthase from *Neurospora crassa*," *The Journal of Biological Chemistry*, vol. 282, no. 19, pp. 14476–14481, 2007.
- [54] M. L. Durbin, B. McCaig, and M. T. Clegg, "Molecular evolution of the chalcone synthase multigene family in the morning glory genome," *Plant Molecular Biology*, vol. 42, no. 1, pp. 79–92, 2000.
- [55] P. K. H. Koduri, G. S. Gordon, E. I. Barker, C. C. Colpitts, N. W. Ashton, and D. Y. Suh, "Genome-wide analysis of the chalcone synthase superfamily genes of *Physcomitrella patens*," *Plant Molecular Biology*, vol. 72, no. 3, pp. 247–263, 2010.
- [56] S. Wu, S. J. B. O'Leary, S. Gleddie, F. Eudes, A. Laroche, and L. S. Robert, "A chalcone synthase-like gene is highly expressed in the tapetum of both wheat (*Triticum aestivum* L.) and triticale (x *Triticosecale* Wittmack)," *Plant Cell Reports*, vol. 27, no. 9, pp. 1441–1449, 2008.
- [57] R. E. Koes, C. E. Spelt, P. J. M. van den Elzen, and J. N. M. Mol, "Cloning and molecular characterization of the chalcone synthase multigene family of *Petunia hybrida*," *Gene*, vol. 81, no. 2, pp. 245–257, 1989.
- [58] C. Jiang, S. Y. Kim, and D. Y. Suh, "Divergent evolution of the thiolase superfamily and chalcone synthase family," *Molecular Phylogenetics and Evolution*, vol. 49, no. 3, pp. 691–701, 2008.
- [59] G. Forkmann and B. Ruhnau, "Distinct substrate specificity of dihydroflavonol-4-reductase from flowers of *Petunia hybrida*," *Zeitschrift für Naturforschung—Section C: Biosciences*, vol. 42, pp. 1146–1148, 1987.
- [60] S. Rudd, "Expressed sequence tags: alternative or complement to whole genome sequences?" *Trends in Plant Science*, vol. 8, no. 7, pp. 321–329, 2003.
- [61] J. Ogata, Y. Itoh, M. Ishida, H. Yoshida, and Y. Ozeki, "Cloning and heterologous expression of cDNAs encoding flavonoid glucosyltransferases from *Dianthus caryophyllus*," *Plant Biotechnology*, vol. 21, no. 5, pp. 367–375, 2004.
- [62] J. Isayenkova, V. Wray, M. Nimtz, D. Strack, and T. Vogt, "Cloning and functional characterisation of two regioselective flavonoid glucosyltransferases from *Beta vulgaris*," *Phytochemistry*, vol. 67, no. 15, pp. 1598–1612, 2006.
- [63] M. R. Alfenito, E. Souer, C. D. Goodman et al., "Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione S-transferases," *Plant Cell*, vol. 10, no. 7, pp. 1135–1149, 1998.
- [64] K. A. Marrs, M. R. Alfenito, A. M. Lloyd, and V. Walbot, "A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*," *Nature*, vol. 375, no. 6530, pp. 397–400, 1995.
- [65] S. Kitamura, N. Shikazono, and A. Tanaka, "TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*," *Plant Journal*, vol. 37, no. 1, pp. 104–114, 2004.
- [66] L. Loyall, K. Uchida, S. Braun, M. Furuya, and H. Frohnmeyer, "Glutathione and a UV light-induced glutathione S-transferase are involved in signaling to chalcone synthase in cell cultures," *Plant Cell*, vol. 12, no. 10, pp. 1939–1950, 2000.



- [67] R. C. Meyer, P. B. Goldsbrough, and W. R. Woodson, "An ethylene-responsive flower senescence-related gene from carnation encodes a protein homologous to glutathione S-transferases," *Plant Molecular Biology*, vol. 17, no. 2, pp. 277–281, 1991.
- [68] H. Itzhaki, J. M. Maxson, and W. R. Woodson, "An ethylene-responsive enhancer element is involved in the senescence-related expression of the carnation glutathione-S-transferase (GST1) gene," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 19, pp. 8925–8929, 1994.
- [69] P. Elomaa, A. Uimari, M. Mehto, V. A. Albert, R. A. E. Laitinen, and T. H. Teeri, "Activation of anthocyanin biosynthesis in *Gerbera hybrida* (Asteraceae) suggests conserved protein-protein and protein-promoter interactions between the anciently diverged monocots and eudicots," *Plant Physiology*, vol. 133, no. 4, pp. 1831–1842, 2003.
- [70] H. Mathews, S. K. Clendennen, C. G. Caldwell et al., "Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport," *Plant Cell*, vol. 15, no. 8, pp. 1689–1703, 2003.
- [71] R. Stracke, M. Werber, and B. Weisshaar, "The R2R3-MYB gene family in *Arabidopsis thaliana*," *Current Opinion in Plant Biology*, vol. 4, no. 5, pp. 447–456, 2001.
- [72] K. Lin-Wang, K. Bolitho, and K. Grafton, "An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae," *BMC Plant Biology*, vol. 10, no. 1, pp. 50–67, 2010.
- [73] M. Yamagishi, Y. Shimoyamada, T. Nakatsuka, and K. Masuda, "Two R2R3-MYB genes, homologs of petunia AN2, regulate anthocyanin biosyntheses in flower tepals, tepal spots and leaves of asiatic hybrid Lily," *Plant and Cell Physiology*, vol. 51, no. 3, pp. 463–474, 2010.
- [74] C. Jiang, X. Gu, and T. Peterson, "Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of *Arabidopsis* and *Oryza sativa* L. ssp. indica," *Genome Biology*, vol. 5, no. 7, p. R46, 2004.
- [75] S. Kobayashi, M. Ishimaru, K. Hiraoka, and C. Honda, "Myb-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis," *Planta*, vol. 215, no. 6, pp. 924–933, 2002.
- [76] A. R. Walker, E. Lee, J. Bogs, D. A. J. McDavid, M. R. Thomas, and S. P. Robinson, "White grapes arose through the mutation of two similar and adjacent regulatory genes," *Plant Journal*, vol. 49, no. 5, pp. 772–785, 2007.
- [77] A. Gonzalez, M. Zhao, J. M. Leavitt, and A. M. Lloyd, "Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings," *Plant Journal*, vol. 53, no. 5, pp. 814–827, 2008.
- [78] N. de Vetten, F. Quattrocchio, J. Mol, and R. Koes, "The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals," *Genes and Development*, vol. 11, no. 11, pp. 1422–1434, 1997.
- [79] J. Brueggemann, B. Weisshaar, and M. Sagasser, "A WD40-repeat gene from *Malus × domestica* is a functional homologue of *Arabidopsis thaliana* TRANSPARENT TESTA GLABRA1," *Plant Cell Reports*, vol. 29, no. 3, pp. 285–294, 2010.