UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITY OF PADOVA

DOCTORAL SCHOOL OF CROP SCIENCE

AGROBIOTECHNOLOGY

XXVI CYCLE

ASSOCIATION OF THE EXPRESSION LEVELS OF TRANSCRIPTION FACTORS WITH THE PHENOTYPES AND GENOTYPES OF PEACH FRUITS THAT DIFFER IN THEIR QUALITATIVE CHARACTERISTICS

Director: Ch.mo Prof. Antonio Berti

Supervisor: Ch.mo Prof. Livio Trainotti

PhD student: Md Abdur Rahim

31 January 2014

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Md Abdur Rahim, 31/01/2014

A copy of the thesis will be available at http://paduaresearch.cab.unipd.it/

Dichiarazione

Con la presente affermo che questa tesi è frutto del mio lavoro e che, per quanto io ne sia a conoscenza, non contiene materiale precedentemente pubblicato o scritto da un'altra persona né materiale che è stato utilizzato per l'ottenimento di qualunque altro titolo o diploma dell'università o altro istituto di apprendimento, a eccezione del caso in cui ciò venga riconosciuto nel testo.

Md Abdur Rahim, 31/01/2014

Una copia della tesi sarà disponibile presso http://paduaresearch.cab.unipd.it/

Acknowledgements

All praises to almighty Allah (the creator) for giving me patience, strength and ability to complete this PhD thesis successfully.

I would like to express my deeply-felt thanks to my supervisor Prof. Livio Trainotti for allowing me to be part of his research group to learn Plant Molecular Biology. I am grateful to him for his continuous inspiration, overall resourceful guidelines, scientific contribution, reviewing, editing of my PhD thesis and supporting my research expenses during entire PhD period.

I would like to acknowledge and express my heartfelt gratitude to Prof. Giorgio Casadoro for his encouragement, innovative science discussion, motivation and kind suggestions that enriched me about plant sciences.

I would like to thank to Prof.ssa Patrizia Torrigiani, Department of Fruit Tree and Woody Plant Sciences, University of Bologna, Italy for providing primers for peach anthocyanin biosynthetic genes.

I would like to express my gratitude to my good friend Nicola Busatto for helping me through the difficult moments, emotional support and the fun we have had during my stay in Italy.

A special thanks to my labmates, Alessandro Lovisetto, Anna Pavanello for creating pleasant lab atmosphere, being always ready to help, and inspiring me.

I would like to acknowledge the "*CARIPARO Foundation*" for providing me fellowship during PhD program.

I am thankful to Shere-e-Bangla Agricultural University (SAU) for providing me study leave during entire PhD program.

Finally, I want to thank my parents, my parents-in-law, my wife, my lovely daughter (Noshaba), my brother and my sister. It would not have been possible to finish my PhD work without their blessings, sacrifices never ending affection.

Md. Abdur Rahim
31 January, 2014
University of Padova, Padova

Index

Riassunto

Le antocianine sono composti vegetali bioattivi che controllano il colore di molte verdure, fiori, frutti, radici e altre parti di piante. Nel pesco [*Prunus persica* (L.) Batsch], il colore è un fattore determinante per la qualità del frutto ed è regolato dai flovonoidi tra cui gli antociani. Queste molecole bioattive hanno benefici effetti per la salute umana, compresa la protezione contro il cancro, le malattie cardiovascolari, infiammatorie e altre malattie croniche. I fattori di trascrizione (TF) R2R3-MYB controllano l'espressione dei geni della sintesi delle antocianine con l'aiuto di co-attivatori appartenenti alle famiglie basic-helix-loop-helix (bHLH) e WD40. Nel genoma di pesco, attraverso questo studio sono stati identificati tre geni di tipo MYB10-like (MYB10.1, MYB10.2 e MYB10.3) e tre bHLH-like (bHLH3, bHLH33 e GL3) quali migliori candidati ad essere i regolatori dell'accumulo degli antociani in pesca. Tra questi *MYB10*s, *MYB10.1* è il gene più espresso nei frutti maturi. L'accumulo di antociani in frutti di pesco è più alto nella buccia, seguito dal mesocarpo attorno all'endocarpo e minima nel resto del mesocarpo. L'espressione dei geni *MYB10.1* e *MYB10.3* è correlata con i livelli di antocianine delle diverse parti della pesca. Essi sono correlati positivamente con i principali geni implicati nella via di sintesi delle antocianine, come la *calcone sintasi* (*CHS*), *flavanone-3β-idrossilasi* (*F3H*), *diidroflavonol 4-reduttasi* (*DFR*) e *UDP-glucose:flavonoid-3-O-glucosiltrasferase* (*UFGT*). Il ruolo dei *MYB10* di pesco è stato testato in tabacco e si è visto che nel caso vengano co-espressi con bHLHs si attivano alcuni geni chiave nella via di sintesi delle antocianine. Solo le sovraespressioni di *MYB10.1/bHLH3* e *MYB10.3/bHLH3* attivano la produzione di antociani regolando i geni *NtCHS*, *NtDFR* e *NtUFGT* mentre altre combinazioni non sono state efficaci, confermando anche dal punto di vista funzionale la maggior importanza di *MYB10.1* rispetto a *MYB10.3* e *MYB10.2*. Per meglio comprendere il ruolo di *MYB10.1* nella regolazione della biosintesi delle antocianine nella pesca in maturazione, la porzione codificante del gene è stata fatta sovra-esprimere in modo costitutivo in piante transgeniche di tabacco. La sua sovra-espressione induce la tipica pigmentazione dovuta ad antocianine nelle parti riproduttive di linee di tabacco transgeniche attraverso l'induzione dei livelli di trascrizione di molti importanti geni della via biosintetica degli

antociani, come *fenilalanin-ammonio-liase* (*NtPAL*), *NtCHS*, *calcone isomerasi* (*NtCHI*), *NtF3H*, *NtDFR*, *antocianina sintasi* (*NtANS*) e *NtUFGT*. L'accumulo dei pigmenti è sempre stato limitato alle parti riproduttive e mai a quelle vegetative come fusto e foglie. Accanto alla regolazione della biosintesi di antocianine, l'espressione ectopica di *MYB10.1* conduce a due diversi fenotipi nel tabacco rispetto al WT. Nel fenotipo forte, detto Tipo I, si ha forma irregolare della foglia e dimensione della pianta ridotta in altezza, ridotta lunghezza del filamento dello stame, antere non-deiscenti, ridotta lunghezza del pistillo, nessuna formazione di nettari e ridotto sviluppo della capsula; inoltre, parti del fiore, tra cui androceo, gineceo e petali, sono più pigmentate. Nel fenotipo blando, detto di tipo II, la crescita delle piante è regolare così come lo sviluppo e la maturazione del fiore; tuttavia i tegumenti di ovuli e semi sono viola. Sorprendentemente, la sovra-espressione di *MYB10.1* non solo ha fatto aumentare la sintesi di antociani, ma ha anche fortemente alterato lo sviluppo fiorale in linee di tabacco transgeniche. In piante di tipo I, la sovra-espressione di *MYB10.1* porta alla riduzione dell'espressione di *NtMYB305*, gene necessario per lo sviluppo del fiore, simile, ma non ortologo, a *MYB10.1*. Inoltre, *MYB10.1* induce l'espressione di geni coinvolti nella biosintesi (*NtAOS*) e segnalazione (*NtJAZd*) dell'acido jasmonico (JA), oltreché di quella dell'etilene (*NtACO*). Inoltre, *NtNCE1*, gene di cui è noto il coinvolgimento nella formazione della ghiandola del nettario, è represso nella sua trascrizione, che è noto essere regolata da *NtMYB305*. Al contrario, piante transgeniche di tipo II hanno mostrato risultati opposti per *NtMYB305*, *NtAOS*, *NtJAZd* e *NtNEC1*. Pertanto, si può concludere che la sovra- espressione di *MYB10.1* di pesco nel tabacco non solo regola la biosintesi dei flavonoidi nelle parti riproduttive ma colpisce anche altri processi come lo sviluppo vegetativo e riproduttivo. Questi nuovi risultati promuoveranno ulteriori indagini per chiarire i molteplici ruoli dei TF di tipo R2R3-MYB. Inoltre, le maggiori e migliori conoscenze sulla regolazione della biosintesi degli antociani aiuteranno i coltivatori a produrre frutta più colorata introducendo nuove cultivar di pesco con più alto livello di antiossidanti.

Abstract

Anthocyanins are plant bioactive compounds that control color of many vegetables, flowers, fruit, roots and other plant parts. In peach [*Prunus persica* (L.) Batsch], color is a key determinant for fruit quality and depends by flovonoids including anthocyanins. These bioactive molecules have potential benefits to human health including protection against cancer, cardiovascular diseases, inflammation and other chronic diseases. The R2R3-MYB transcription factors (TFs) control the expression of genes for the synthesis of anthocyanin pigments with the help of co-activators belonging to the basic-helix-loop-helix (bHLH) and WD40 repeat family. In the peach genome, through this study three MYB10-like (MYB10.1, MYB10.2 and MYB10.3) and three bHLH-like (bHLH3, bHLH33 and GL3) TFs were identified as best candidates to be the regulators of the anthocyanin accumulation in peach. Among these *MYB10*s, *MYB10.1* is the highly expressed gene in ripe peach fruits. The accumulation of anthocyanins in yellow flesh peach fruits is highest in the peel, abundant in the part of the mesocarp surrounding the stone and lowest in the mesocarp. The expression of *MYB10.1* and *MYB10.3* genes were correlated with anthocyanin levels of different peach parts. They have also positive correlation with the key structural genes of the anthocyanin pathway, like *chalcone synthase* (*CHS*), *flavanone-3β-hydroxylase* (*F3H*), *dihydroflavonol 4-reductase* (*DFR*) and *UDP-glucose:flavonoid-3-O-glucosyltransferase* (*UFGT*). Functions of peach *MYB10s* were tested in tobacco and shown to activate key genes in the anthocyanin pathway when *bHLH*s were co-expressed as partners. Over-expression of *MYB10.1*/*bHLH3* and *MYB10.3*/*bHLH3* activated anthocyanin production by up-regulating *NtCHS, NtDFR* and *NtUFGT* while other combinations were not effective.

To better understand the role of *MYB10.1* in anthocyanin synthesis during peach ripening, its coding sequence has been constitutively expressed in tobacco transgenic plants. Its over-expression induce anthocyanin pigmentation in the reproductive parts of transgenic tobacco lines through the up-regulation of transcript levels of many important genes involved in the anthocyanin biosynthetic pathway like *phenylalanineammonia-lyase* (*NtPAL*), *NtCHS*, *chalcone isomerase* (*NtCHI*), *NtF3H*, *NtDFR*, *anthocyanin synthase* (*NtANS*) and *NtUFGT*. The pigment accumulation was always limited to reproductive parts and never present

in the vegetative part like stem and leaves. Besides the regulation of anthocyanin biosynthesis, the ectopic expression of *MYB10.1* leads to two different types of phenotype in transgenic as compared to WT tobacco. The strong phenotype, called type-I, is characterized by irregular leaf shape and size, reduced plant height, stamens with reduced filament length and non-dehiscing anthers, reduced pistil length, no nectary gland formation and reduced capsule development but pigmented reproductive parts including androecium, gynoecium and petals. On the contrary, so-called type-II plants, have a mild phenotype with regular plant growth and development, but purple seed coats. Surprisingly, the over-expression of *MYB10.1* altered the floral development in transgenic tobacco lines. In type-I transgenic plants, the over-expression of peach *MYB10.1* leads to down-regulation of *NtMYB305*, a gene similar, but not orthologous, to *MYB10.1*, which is required for floral development. Moreover, *MYB10.1* up-regulates the expression of jasmonic acid (JA) biosynthesis (*allene oxide synthase*, *NtAO*S) and signaling (*NtJAZd*) pathway genes as well as *1-aminocyclopropane-1-carboxylate oxidase* (*NtACO*), a key gene in ethylene synthesis. Furthermore, *NECTARIN1* (*NtNCE1*) that is known to be involved in nectary gland formation, was repressed since its transcription is regulated by *NtMYB305*. On the contrary, type-II transgenic plants showed opposite results for *NtMYB305*, *NtAOS*, *NtJAZd* and *NtNEC1*. Therefore, it can be concluded that the over-expression of peach *MYB10.1* in tobacco not only regulates flavonoid biosynthesis in the reproductive parts but also affects (directly and/or indirectly) other process like vegetative and reproductive development. These new findings will promote further investigation to elucidate more diverse role of R2R3-MYB TFs. Furthermore, they will promote more and better insights on the regulation of anthocyanin biosynthesis and will aid fruit breeders to introduce new peach cultivars with higher antioxidant level as well as colored fruit.

Chapter I

1 Introduction

1.1 The fruit

Fruits are typical of flowering plants and, in general, they derive from an ovary following an event of fertilization. However, in some species the fruit can also develop from extracarpellary floral parts like receptacle (strawberry), bracts (pineapple), calyx (mulberry), floral tube (apple, pear) or complete inflorescence (fig) and they are known as false fruit or pseudocarp (Giovannoni 2004). Higher plants produce either dry or fleshy fruits to protect their seeds and to facilitate their dispersal into the environment by means of animal, water and wind or gravity. There are two types of dry fruits: dehiscent in which the seeds are contained in a seedpod that splits at maturity to release the seeds, while indehiscent fruits normally contain one seed and thus it is the whole fruit that can be released into the environment. In fleshy fruits (e.g. tomato, peach, grapes and others) the juicy pericarp has three distinct layers such as the outer layer (exocarp), the middle layer (mesocarp), and the inner layer (endocarp). They are mostly edible and colorful so they attract the attention of frugivorous animals that carry out the dispersal of seeds through their dejections (Montiel and Montaña 2000). From now onward only fleshy fruits will be considered.

The ripening of fleshy fruits consists of a series of biochemical and physiological changes that occur during the last stages of the fruit development. These changes generally include alteration of cell wall structure and texture, change of color, accumulation of sugars, increased susceptibility to pathogens, and accumulation of flavor and aromatic volatiles (Giovannoni 2001). Fleshy fruits can further be divided into two types based on the presence or not of a respiration peak during ripening: climacteric and non-climacteric fruit, respectively. Climacteric fruits have a high respiration rate that is paralleled by a dramatic increase of ethylene production, while non-climacteric fruits do not produce any significant

ethylene amount. Climacteric fruits include apple, peach, pear, banana, tomato and others, whereas strawberry, grapes and citrus are examples of non-climacteric fruits. Due to their importance in the human diet, studies on fruits have mostly dealt with fleshy fruits, although dry fruits represent the majority (Giovannoni 2004).

1.2 The rosaceous fruits

The Rosaceae family comprises 95 genera with 2830 species, some economically important for the production of fruit, nuts, ornamental flowers, and some for being woody perennials (Stevens 2001). From the ancient periods, rosaceous plants populated native human habitats and were considered important sources of food. Many species of this family provide nutritious foods as well as industrial and aesthetic products (Shulaev et al. 2008). The most economically important fruits of this family are apples, peaches, apricots, plums, cherries, pears, quinces, raspberries, and strawberries. *Prunus* is the largest genus of subfamily Amygdaloideae (Takhtadzhian 1997) which consists of over 200 species (Rehder 1949) including peaches, plums, cherries, apricots and almonds. During ripening, most of these fruits accumulate different types of bioactive phytochemicals such as flavonoids (including anthocyanins, flavones and flavonols) that are beneficial to human health. Color is an important qualitative trait of these fruits and is mostly determined by anthocyanins.

1.3 Consumer preferences and fruit quality

The economic value of the fruit depends highly on its quality. The consumers' preferences are based on some combinations of observable characteristics such as fruit size, color, shape, and amount of defects, and also some non-observable qualities such as sweetness, juiciness, texture, and flavor (Parker et al. 1991). The rosaceous fruits like apple, peach, pear, strawberry, cherry, and plum play an important role in human health as they are rich in vitamins (C: ascorbic acid, B1: thiamine, B3: niacin, B6: pyridoxine, B9: folic acid), minerals and dietary fibers, and different types of bioactive phytochemicals such as flavonoids (including anthocyanins, flavones and flavonols) and carotenoids (Craig and Beck 1999,

Wargovich 2000, Rao and Rao 2007). Therefore, fruit quality is of paramount importance to the orchard owners and the fruit breeders. The fruit development, growth, ripening, and senescence occur through biochemical processes and the genetic dissection of these complex processes may aid in maximizing and maintaining fruit quality from production to post-harvest processing (Ogundiwin et al. 2009). The fruit flavor is also an important quality for consumers and depends upon several factors such as firmness, juiciness, sweetness, acidity, aroma, and texture (Shulaev et al. 2008; Shulaev et al. 2010). In fleshy fruits, sweetness and acid content are the most important factors and sweetness depends on the sugar content which can be measured as soluble solid content (SSC). In particular, in ripe peach the most abundant sugar is sucrose followed by glucose, fructose, and sorbitol (Brooks et al. 1993). After genome sequencing, the discovery of high density molecular markers in the chromosomes became much easier and this can be used for future improvement of genetic maps useful for the improvement of quality traits through marker assisted selection (MAS). The genetic map is also an important tool to identify genes and to map qualitative and quantitative trait loci (QTL) as well as for comparative mapping among species. The genetic maps of several fruit species are now available to be used in fruit breeding program such as apple (Hemmat et al. 1994, Liebhard et al. 2003, Kenis and Keulemans 2005), pear (Yamamoto et al. 2002), peach (Dirlewanger et al. 1998, Lambert et al. 2004, Dirlewanger et al. 2006), grapes (Troggio et al. 2007), strawberry (Sargent et al. 2006), tomato (Sharma et al. 2009), melon (Perin et al. 2002), papaya (Chen et al. 2007), banana (Hippolyte et al. 2010), sweet cherry (Stockinger et al. 1996), kiwifruit (Fraser et al. 2009). In *Prunus davidiana*, Quantitative Trait Loci (QTLs) analyses were carried out by Quilot et al. (2004) with 24 physical and biochemical traits and they found alleles from *P. davidiana* with agronomically favorable effects related to fruit and stone sizes, sugar and acid concentrations and red flesh coloration. Dirlewanger et al. (2004a) developed a microsatellite genetic linkage map of myrobalan plum and an almond-peach hybrid and identified *Ma* marker linked to root knot nematode resistance. Furthermore, Dirlewanger et al. (2004b) showed the position of 28 major loci affecting agronomic characters found in different species (Figure 1.1). In the genus *Prunus* a fruit quality map was developed by Ogundiwin et al. (2009) who reported 133 genes involved in fruit texture, pigmentation, flavor, and chilling

injury resistance. Thus, the application of molecular markers could be useful for improving the quality of fruits.

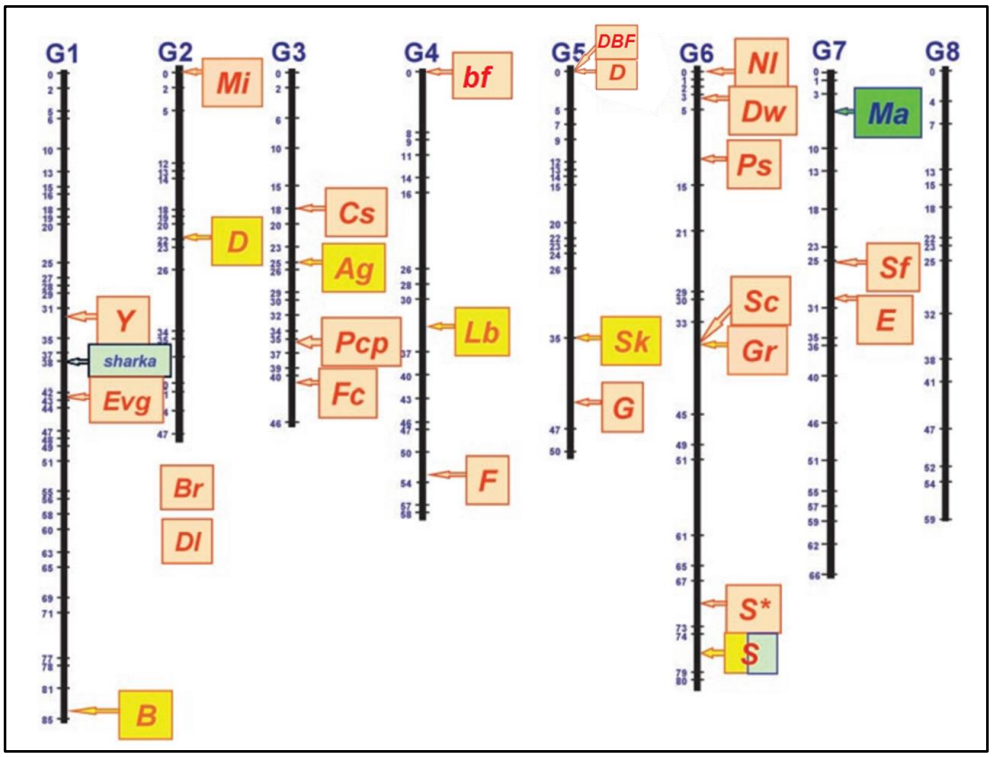

Figure 1.1 Position of major genes mapped in different populations of apricot (blue background), peach (orange background), almond or almond × peach (yellow background), and Myrobalan plum (green background) on the *Prunus* reference map. Gene abbreviations correspond to: *Y*, peach flesh color; *sharka*, plum pox virus resistance; *B*, flower color in almond × peach; *Mi*, nematode resistance from peach; D, almond shell hardness; *Br*, broomy plant habit; *Dl*, double flower; *Cs*, flesh color around the stone; *Ag*, anther color; *Pcp*, polycarpel; *Fc*, flower color; *bf*, blood-flesh (Gillen and Bliss 2005); *Lb*, blooming date; *F*, flesh adherence to stone; *D*, non-acid fruit in peach, *DBF*, dominant blood-flesh (Shen et al. 2013); *Sk*, bitter kernel; *G*, fruit skin pubescence; *Nl*, leaf shape; *Dw*, dwarf plant; *Ps*, male sterility; *Sc*, fruit skin color; *Gr*, leaf color; *S**, fruit shape; *S*, self-incompatibility (almond and apricot); *Ma*, nematode resistance from Myrobalan plum; *E*, leaf gland shape; *Sf*, resistance to powdery mildew. Genes *Dl* and *Br* are located on an unknown position of G2 (modified from Dirlewanger et al., 2004b).

1.4 Ripening of fleshy fruit

The major changes that occur during ripening of fruits are: i) organic acids (e.g. malic, citric, tartaric) are partially used either for the formation of carbohydrates or for respiration; ii) tannins disappear by complete oxidation without forming carbohydrates; iii) the starch is transformed into sugar; iv) the glucidic substrates partly disappear by oxidation (EFS 1898). Ripening of fleshy fruits is a complex and highly coordinated developmental process that coincides with seed maturation (Klee and Giovannoni, 2011). During fruit ripening, the different physiological and

biochemical changes occur through differential expression of various genes that are developmentally regulated and up and/or down regulation of these genes contribute to various changes in the fruit that make it visually attractive and edible (Bapat et al. 2010). Tomato is a climacteric fruit that has increased ethylene production and respiration during ripening. It has a small and functionally well characterized genome, simple propagation technique either through seed or clonal, efficient sexual hybridization, different ripening phenotypes, a short life cycle, and year round production (Giovannoni, 2001). A number of molecular tools are also available for tomato genomics, such as genetic and physical map, molecular markers, cDNA libraries, EST resources (120892 ESTs), microarrays, and RNA-seq (Alba et al. 2005, Karlova et al. 2011, Zuo et al. 2012). In addition to these, the process of ripening is well characterized in respect of metabolic changes impacting softening, accumulation of sugars and acids, chlorophyll degradation, lycopene accumulation, and increases in ethylene and flavor volatiles (Chung et al. 2010). Therefore, tomato is used as a model system for climacteric fruit ripening (Figure 1.2). The phytohormone ethylene plays an important role for tomato fruit ripening as well as for other climacteric fruits. For example, in the tomato mutant *never ripe* (*Nr*), fruits do not undergo ripening and abscission even if treated with exogenous ethylene as a result of a change of a single amino acid in the NR ethylene receptor (Klee and Giovannoni 2011). This mutant demonstrates the significance of ethylene in climacteric fruit ripening. The *RIN* (*RIPENING INHIBITOR*) gene and the *CNR* (*COLORLESS NON-RIPENING)* gene encode transcription factors that are important for the regulation of ripening since the tomato plants harboring either *rin* or *cnr* mutations are unable to ripe their fruits (Vrebalov et al. 2002; Manning et al. 2006). Besides, another transcription factor gene, *APETALA2a*, regulates ripening of tomato fruit via regulation of ethylene biosynthesis and acts together with *COLORLESS NON-RIPENING* (*CNR*) in a negative feedback loop (Karlova et al. 2011). *AP2a* RNAi fruits showed alteration in the expression of genes involved in several metabolic pathways as those for phenylpropanoid and carotenoid synthesis, as well as in hormone synthesis and perception.

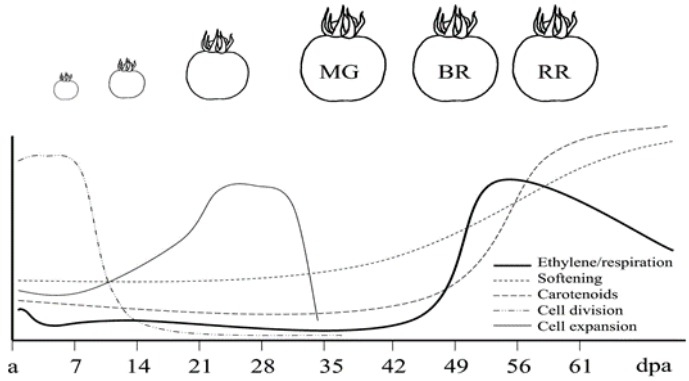

Figure 1.2 The major developmental changes during tomato fruit development and ripening. Abbreviations correspond to a, anthesis period; dpa, days after anthesis; MG, fully expanded mature green; BR, breaker; and RR, red ripe (adapted from Giovannoni, 2004).

1.5 The peach fruit

Peach is one of the most economically important fruit species. It belongs to the order Rosales, family Rosaceae, sub-family Prunoidae, genus *Prunus*, and species *Prunus persica* (L.) Batsch. Peaches are originated and domesticated from China before the Zhou archeological period (3300-2500 B.C.). Thereafter, peaches spread from China to Iran via the silk roads and from Iran they were introduced to Europe by the Romans during 2nd century B.C. to beginning of the Christian era (Faust and Timon 1995). Finally, peaches were brought to the USA by the Spaniards during 16th century (Bassi and Monet 2008). In 2011, peaches and nectarines were cultivated on a total of 1.57 million ha of land and the production was 21.53 million tons in the world (http://faostat3.fao.org). China is the largest world producer of peaches followed by Italy (Figure 1.3). In Europe, Italy is the main exporter of peaches.

Peach is a diploid species (2n=16) and it is a model fruit crop due to its small sized genome (265 Mb), self-compatibility and functionally well characterized agronomic traits. The peach (*Prunus persica*) genome sequencing was announced at the "Plant and Animal Genome XV Meeting" on 16th January 2007 by Jerry Tuskan from the "Joint Genome Institute (JGI)". Subsequently, the "International Peach Genome Initiative (IPGI)" was formed to sequence peach genome cooperatively under the direction of Drs Bryon Sosinski (NC State University), Ignazio Verde (Consiglio per la Ricerca e la Sperimentazione in Agricoltura) and Daniel Rokhsar (DOE Joint Genome Institute), included numerous researchers from

countries around the world including the USA, Italy (Drupomics), Spain and Chile. The first draft of peach genome cv 'Lovell' (peach v1.0) (assembled and annotated by IPGI) was released on 1st April 2010. Then, the peach genome browser was housed at JGI, the Genome Database for Rosaceae (GDR), and the Italian version housed at "Istituto di Genomica Applicata (IGA)". Finally, the high-quality draft genome was published on March 2013 (Verde, 2013). There are 27852 protein-coding genes and 28689 protein-coding transcripts present in the peach genome.

Peaches are low in calories and do not contain saturated fats. They are a source of antioxidant (vitamin C, vitamin A, ß-carotene), they are rich in minerals (potassium, fluoride and iron) and contain health promoting flavonoid, and antioxidant carotenoids (lutein, zeaxanthin and ß-cryptoxanthin) (Table 1.1). Peach is a small to medium-sized deciduous tree growing maximum 8m. It has long generation time (usually 3-5 years) like other fruit trees. Leaves are simple, long, glossy green, lanceolate, serrrate. The flowers are pure white to dark red, hermaphrodite, perigynous, self-pollinated and the blooming period is March to May. Peaches are stone fruits (drupes). Like in other stone fruits (cherry, plum, apricot), peach seeds are enclosed by a lignified endocarp called stone (inner part), surrounded by a fleshy mesocarp (middle part) and a thin exocarp (outer part) called skin (Figure 1.4). The peach fruit shape is either round or elongate with elongate dominant over round (Bassi and Monet 2008). The fruit can also be flat, and in flat peaches (also called "saucer", "pan-tao", or "peento"), both the fruit and the pith are flattened to opposite poles. This trait is monogenic and dominant over round/elongate (S/s) (Lesley 1940). Fruits are divided into two categories based on skin texture and pubescence such as peaches (fuzzy and dull) and nectarines (smooth and shine). Both peaches and nectarines are further classified as *freestone* (pit relatively free of the flesh) or *clingstone* (pit adheres to flesh). The colors of the peach flesh are either yellow or white and white (Y) is dominant over yellow (y) (Dirlewanger et al. 2004b). The "bloody flesh" peaches accumulate high amount of anthocyanins in their flesh and this trait is controlled by one locus *bf* (blood-flesh) and associated with reduced tree height (Werner et al. 1998). Peach flesh texture can be melting or non-melting. Melting peaches are good for fresh market as they soften quickly after harvest, while non-melting peaches remain firm and are mainly used for canning purposes.

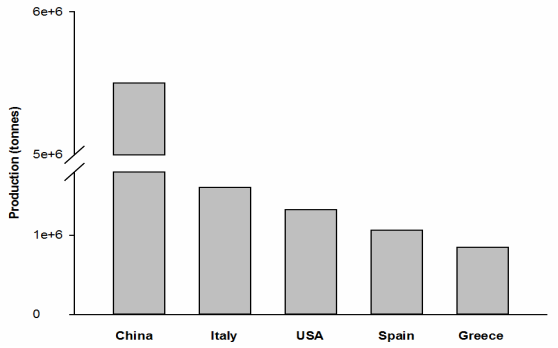

Figure 1.3 Top five peaches and nectarines producing countries with estimated production in tonnes (FAOstat, updated 2013).

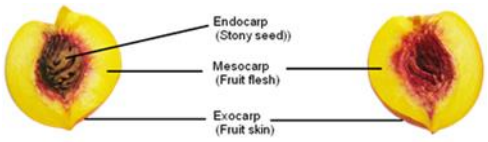

Figure 1.4 Cross section of peach fruit (cv 'Fantasia').

Table 1.1 Nutritive value per 100g fresh peaches

Principle	Nutrient Value	% of RDA	Principle	Nutrient Value	% of RDA
Energy	39 Kcal	2	Vitamin E	0.73 mg	5
Carbohydrates	9.54 g	7	Vitamin K	2.6 µg	2
Protein	0.91 g	1.5	Potassium	190 mg	4
Total Fat	0.25 g	1	Calcium	6 mg	0.6
Cholesterol	0 mg	0	Copper	0.068 mg	7.5
Dietary Fiber	1.5 g	4	Iron	0.25 mg	3
Folates	4 µg	1	Magnesium	9 mg	2
Niacin	0.806 mg	5	Manganese	0.61 mg	3
Pantothenic acid	0.153 mg	3	Phosphorus	11 mg	2
Pyridoxine	0.025 mg	2	Zinc	0.17 mg	1.50
Riboflavin	0.031 mg	2.5	ß-Carotene	162 µg	--
Thiamin	0.024 mg	2	ß-Cryptoxanthin	67 µg	--
Vitamin A	326 IU	11	Lutein-zeaxanthin	91 µg	--
Vitamin C	6.6 mg	11			

(Source: http://www.nutrition-and-you.com/peaches.html)

1.5.1 Peach ripening

Peach is a climacteric fruit and has four distinct developmental stages: S1, S2, S3 and S4 (Masia et al. 1992, Zanchin et al. 1994). The fruit development shows a double-sigmoid growth pattern (Chalmers and Ende 1975, Masia et al. 1992) and S1 and S3 exhibit exponential growth while growth is slow at S2 and S4. In S1 cellular multiplication and distension occur; during S2 the pith hardening takes place to form the stone, during S3 fruit size increases mainly due to cell

expansion and in S4 the fruit reaches its final dimensions and ripens (Figure 1.5). During peach ripening there is a significant increase in ethylene production that changes the expression level of a number of different genes, many of them regulated by the hormone (Trainotti et al. 2007). Moreover, ethylene directly regulates the transcriptional level of softening-related genes like *endopolygalacturonase*, an *α-L-arabinofuranosidase*/*β-xylosidase*, and *expansin* (Hayama et al. 2006). Such coordinated and programmed modulation of gene expression leads to several alterations which ultimately influence fruit quality (Trainotti et al. 2006). The fruit quality highly depends on physiological maturity of the fruit (Busatto 2012). There are many changes that occur in pigment composition during maturity and ripening, and they include chlorophyll degradation, and carotenoid and anthocyanin accumulation. In peach, the red peel color is mainly due to the accumulation of anthocyanins which are one of the most important factors determining consumers' acceptance.

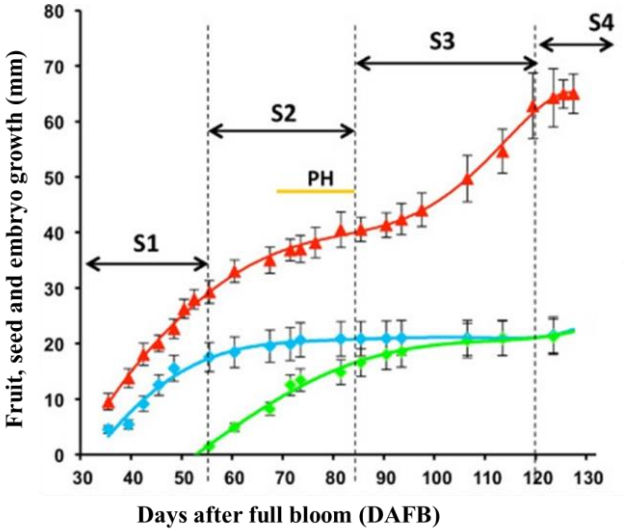

Figure 1.5 Fruit and seed growth pattern of peach (cv. Fantasia). Fruit growth (red) is expressed as cross diameter while length is used for seed (blue) and embryo (green). The S1, S2, S3 and S4 are the various fruit developmental stages. The yellow bar labeled with PH indicates pith hardening stage (adapted from Bonghi et al. 2011).

1.6 Plant pigments

The main classes of plant pigments, besides the chlorophylls, are the betalains, the carotenoids, and the anthocyanins (Grotewold, 2006; Tanaka et al., 2008). Betalains are water-soluble, nitrogen-containing pigments synthesized from tyrosine. This class of pigments is responsible for yellow, orange, red and purple colors of flowers,

grains, vegetables, fruits and roots that belong to the Chenopodiniae suborder and never co-occur in the same plants with anthocyanins (Azeredo 2009, Mosco 2012, Stafford 1994, Strack et al. 2003). The most important commercially exploited betalain crop is the red beet root (*Beta vulgaris* subsp. *vulgaris*) that contains two major soluble pigments, betanin (red) and vulgaxanthine I (yellow) (Azeredo 2009, Strack et al. 2003). Betalains have also anti-oxidant and radical scavenging properties (Pedreno and Escribano 2000). Carotenoids are isoprenoid, plastid-synthesized molecules and are essential for plant life, providing important photoprotective functions during photosynthesis (Grotewold 2006). There are two types of carotenoids: carotenes such as lycopene and β-carotene, and xanthophylls such as lutein and zeaxanthin (Bramley 2002). The carotenoids confer yellow (β-carotene and xanthophylls), orange and red (lycopene) colors in flower, fruit and roots (Alkema and Seager 1982). The antioxidant properties of carotenoids may help to reduce the risks of cardiovascular disease (Gaziano et al. 1995), colorectal adenomas (Jung et al. 2013), and breast cancer of women (Eliassen et al. 2012). Anthocyanins are the plant bioactive compounds which are an important determinant for color of many flowers, fruits and vegetable. They belong to the flavonoid class of plant pigments and are synthesized from phenylalanine. Anthocyanins are water-soluble, synthesized in the cytosol, and localized in vacuoles. They are broadly distributed in flowering plants and provide a wide range of colors such as orange, red, purple, or blue color according to the molecule and vacuolar pH, co-pigments and metal ions. These plant pigments serve as visible signals to attract insects, birds and animals for pollination and seed dispersal, and also protect plants from damage caused by UV and visible light (Tanaka et al. 2008).

1.6.1 Flavonoid compounds in plants

Flavonoids are polyphenols derived from plant secondary metabolism and are abundant in many flowers, fruits and leaves (Taylor and Grotewold 2005). In the majority of higher plants, there are six major subgroups of flavonoids including chalcones, flavones, flavonols, flavandiols, anthocyanins, and condensed tannins or proanthocyanidins (Winkel-Shirley 2001). Among flavonoids, anthocyanins

represent the largest class of polyphenols comprising over 9000 identified compounds (Welch et al. 2008).

1.6.2 Anthocyanins

Anthocyanins are an important determinant for the color of many flowers, fruits, seeds, vegetables, leaves, roots and other storage organs (Escribano-Bailón et al. 2004, Espley et al. 2007, Mano et al. 2007, Deluc et al. 2008, Welch et al. 2008, Chiu et al. 2010, Kim et al. 2010, Lin-Wang et al. 2010). They are water soluble and the color produced by anthocyanins vary as red, purple or blue according to the molecules and the pH of the vacuole. In higher plants there are six common anthocyanidins (Figure 1.6): pelargonidin (Pg), peonidin (Pn), cyanidin (Cy), malvidin (Mv), petunidin (Pt) and delphinidin (Dp) with Cy, Dp and Pg as the most widespread in nature, being present in 80% of pigmented leaves, 69% of fruits and 50% of flowers (Kong et al. 2003). Excluding chlorophylls, anthocyanins are the most important group of visible plant pigments (Kong et al. 2003). They attract animals for pollination and seeds dispersal hence they are of considerable value in the co-evolution of these plant-animal interactions (Kong et al. 2003). They also play key role in signaling between plants-microbes, in male fertility of some species, in defense as antimicrobial agents and feeding deterrents, and protect plants against UV-B irradiation as well as abiotic and biotic stresses (Winkel-Shirley 2001). During ripening, many fruits accumulate different types of bioactive chemicals, including anthocyanins, that give protection in human health against cancer, cardiovascular diseases, neurodegenerative disease and other chronic diseases (Butelli et al. 2008, Rao and Rao 2007, Singh et al. 2008). Besides, anthocyanins increase antioxidant levels in serum (Mazza et al. 2002), cholesterol distribution (Xia et al. 2007), restoration of vision disorders (Matsumoto et al. 2003), help in reducing obesity (Tsuda et al. 2003) as well as protect human red blood cells (RBCs) from oxidative damage (Tedesco et al. 2001). Several studies have been carried out about the health properties of anthocyanin. For example, a cancer-susceptible *Trp53*$^{-/-}$ mice was fed a purple tomato powder with a high anthocyanin content and it significantly extended its life span (Butelli et al. 2008). In another experiment, rats were fed with a high-cholesterol diet together with the

purple potato flake diet and it was found that the anthocyanins present in the purple potato flake could protect the adverse effects related to oxidative damage cause by the high-cholesterol diet (Han et al. 2007). These experimental evidences showed that plant anthocyanins were beneficial to animal health. The visible reddish color in the skin and flesh of peach fruit is due to accumulation of anthocyanins (Kataoka and Beppu 2004). The major phenolic compounds in ripening peaches are chlorogenic acid, catechin, epicatechin, rutin and cyanidin-3-glucoside (Andreotti et al. 2008) but cyanidin-3-glucoside is only detectable in the peel (Chang et al. 2000, Tomás-Barberán et al. 2001).

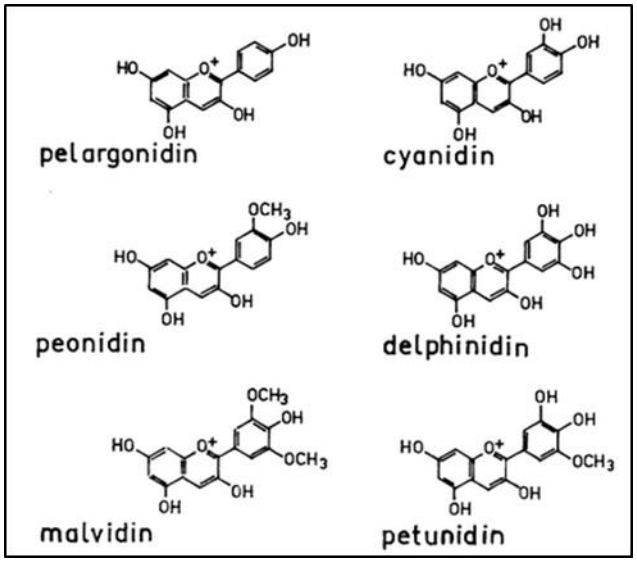

Figure 1.6 Structures of major anthocyanidins (adapted from Brouillard 1982).

1.7 Anthocyanin biosynthetic pathway

The anthocyanin biosynthetic pathway is the best characterized secondary metabolic pathway in plants (Dixon and Steele 1999, reviewed in Grotewold 2006). It is a specific branch of the flavonoid pathway, which starts from the key amino acid phenylalanine to produce 4-coumaroyl CoA by phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H) and 4 coumarate CoA ligase (4CL). The main precursors for all flavonoids are 4-coumaroyl CoA and three molecules of malonyl CoA that produces chalcones (Figure 1.7) containing two phenyl rings by chalcone synthase (CHS) (Dixon and Steele 1999). After that, the pathway is catalyzed by a number of enzymes to yield flavanones (via chalcone isomerase, CHI), dihydroflavonols (via flavanone 3-hydroxylase, F3H), leucoanthocyanidins

(via dihydroflavonol 4-reductase, DFR), anthocyanidins (via leucoanthocyanidin dioxygenase, LDOX) and anthocyanins (via UDP-flavonoid glucosyl transferase, UFGT). The pathway leads to synthesis of anthocyanins along with branches for synthesis of flavonols by flavonol synthase (FLS) and synthesis of condensed tannins by leucoanthocyanidin reductase (LAR) and anthocyanidin reductase (ANR) (Takos et al. 2006).

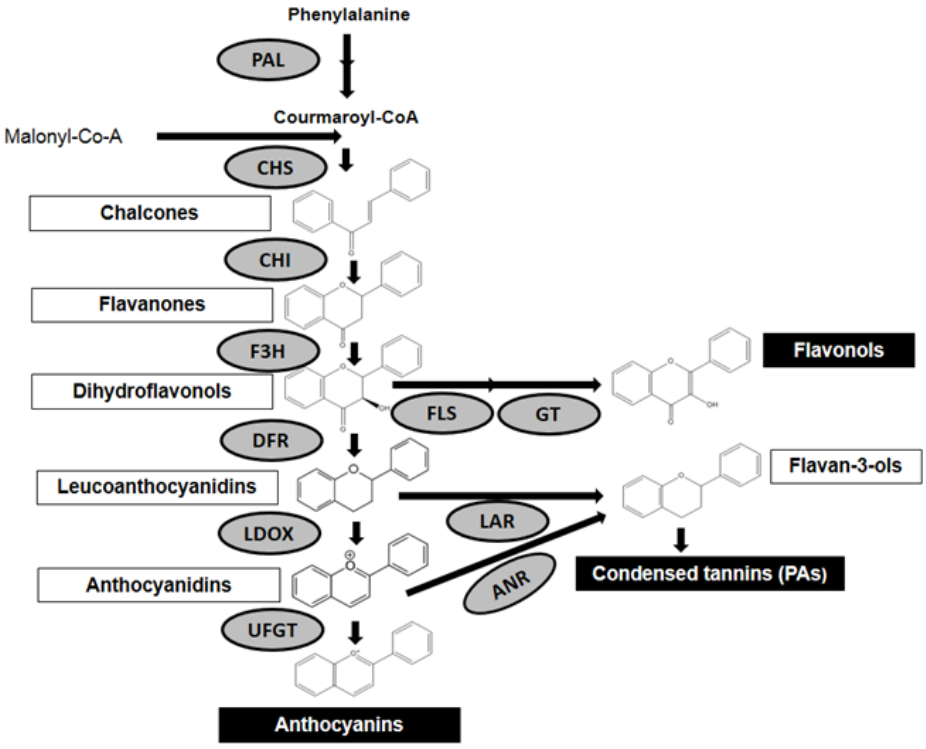

Figure 1.7 Schematic representation of the anthocyanin biosynthetic pathway. The enzymes shown in grey circles are as follows. PAL, phenylalanine ammonia-lyase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone-3β-hydroxylase; FLS, flavonol synthase; GT, unidentified enzyme encoding a glycosyl transferase for flavonol glycone synthesis; DFR, dihydroflavonol 4-reductase; LAR, leucoanthocyanidin reductase; LDOX, leucoanthocyanidin dioxygenase; ANR, anthocyanidin reductase; UFGT, UDP-glucose:flavonoid-3-O-glucosyltransferase (modified from Takos et al. 2006).

The transcription of the structural genes is differentially regulated in monocot and dicot species. The flavonoid biosynthetic genes are regulated by a MYB-bHLH-WD40 (MBW) complex of regulators in monocots while in dicots the early biosynthetic genes (CHS, CHI, F3H and FLS)] are regulated by R2R3-MYB transcription factor (TF) without co-regulators while late biosynthetic genes (DFR,

LAR, LDOX, UFGT and ANR) need a MBW complex (reviewed by Petroni and Tonelli 2011).

1.8 Transport and localization of anthocyanins

Anthocyanins are synthesized in the cytosolic surface of the endoplasmic reticulum (Sun et al. 2012). Then, they are transported via glutathionine S-transferase (GST)-like proteins and localized permanently in the vacuole (Mueller et al. 2000, Lepiniec et al. 2006). The color produced by the anthocyanin depends on the synthesized compound and vacuolar pH (Ohno et al. 2011). This transport mechanism has been shown in maize (Marrs et al. 1995, Goodman et al. 2004), *Petunia* (Mueller et al. 2000) and *Arabidopsis* (Kitamura 2006, Sun et al. 2012).

1.9 Transcription factors (TFs)

The properties of a cell depend on genetic information on its genome (Hobert 2008). The transcription of genes is controlled by several *cis*-acting regulatory sequences that are found adjacent to the structural portion of a gene and are required for gene expression. The transcription of a gene is started at the TATA box in the promoter region (reviewed from Zawel and Reinberg 1995) and is regulated by TFs either alone or with other proteins in a complex, by promoting (activator), or blocking (repressor) the recruitment of RNA polymerase to specific gene (reviewed from Lee and Young 2000). In addition, some other proteins such as co-activators, chromatin remodelers, histone acetylases, deacetylases, kinases, and methylases, may also play important roles in gene regulation (Brivanlou and Darnell 2002). TFs can regulate multiple genes. One such TF is *AtMYB75/PAP1* that regulates many structural genes in the anthocyanin biosynthetic pathway (Tohge et al. 2005). In plant transcription factor database (http://planttfdb.cbi.pku.edu.cn), there are 53574 transcription factors classified into 58 groups from 49 plant species are available (Zhang et al. 2011). The plant specific TFs are APETALA2 (AP2)/ethylene-responsive-element binding protein (EREBP), NAC, WRKY, auxin response factors (ARF) and DOF. Some of the others are MYB, bHLH, MADS and basic-region leucine-zipper (bZIP), which, are not particularly abundant in animals or

yeast and have undergone significant amplification in plants (Riechmann and Ratcliffe 2000). The regulation of flavonoid biosynthesis is accomplished by the combinatorial action(s) of transcriptional activators belonging to different TF families such as R2R3-MYB, bHLH, WD40, WRKY, bZIP and MADS-box (Koes et al. 2005, Ramsay and Glover 2005, Lepiniec et al. 2006).

1.9.1 The MYB TFs

The MYB family of TFs is represented in all eukaryotes and is one of the largest in plants (Dubos et al. 2010, Feller et al. 2011, Tombuloglu et al. 2013). The first MYB TF was identified from avian myeloblastosis virus (AMV), which contains an oncogene *v-myb* (Klempnauer et al. 1982, Klempnauer et al. 1986). The first plant MYB gene, called *C1*, was identified in maize where it was involved in anthocyanin pigmentation (Paz-Ares et al. 1987). According to the transcription factor database (http://planttfdb.cbi.edu.cn/), in total 3485 MYB and 2754 MYB-related sequences are found in plants (Zhang et al. 2011). The functions of MYB proteins in plants are diverse including regulation of primary and secondary metabolism, cellular morphogenesis, cell cycle and regulation of meristem formation, responses to biotic and abiotic stresses (Martin and Paz-Ares 1997, Dubos et al. 2010). The characteristic of MYB proteins is the MYB-domain, consisting of up to four imperfect tandem repeats (R) of about 52 amino acids, that binds DNA in a sequence-specific manner (Ogata et al. 1995, Jin and Martin 1999, Dubos et al. 2010). Each repeat contains a helix-turn-helix (HTH) variation motif (forming three α-helices) and three regularly spaced conserved tryptophan (W) residues participate in forming a hydrophobic core structure. The second and the third helices form a HTH structure that bind to their promoter targets (Lipsick 1996, Jia et al. 2004). Furthermore, the third helix is a recognition helix that is in direct contact with DNA and intercalates in the major groove (Williams and Grotewold 1997, Jia et al. 2004, Dubos et al. 2010). MYB proteins are classified into four different classes depending on the number of MYB-domain repeats (Figure 1.8). Therefore, MYB TFs are called 1R-MYB, R2R3-MYB, 3R-MYB and 4R-MYB having one, two, three and four MYB-domains, respectively (Stracke et al. 2001, Dubos et al. 2010). Among these four classes of MYB TFs, the R2R3-MYB (proteins with two repeats)

family is specific to plant and yeast but predominant in plants (Jin and Martin 1999). The single repeat R3-MYB TFs, which include *TRIPTYCHON (TRY), CAPRICE (CPC)* and *MYBL2*, are likely to have evolved from *R2R3-MYB* genes (reviewed from Dubos et al. 2010). In *Arabidopsis*, there are total six R3-MYBs and they function in trichome and root hair patterning (Wang et al 2010; Gan et al. 2011). Some of these R3-MYB TFs are shown to function as transcriptional repressor and negative regulators of anthocyanin biosynthesis. The R3-MYB TF CPC (CAPRICE) negatively regulate anthocyanin biosynthesis by competing with a R2R3-MYB (AtPAP1) TF in *Arabidopsis* (Zhu et al. 2009). Moreover, Matsui et al. (2008) showed that also AtMYBL2 is a transcriptional repressor and negatively regulates the pigment biosynthesis. Similarly, also Aharoni et al. (2001) reported that a MYB TF, *FaMYB1*, functions as a repressor of anthocyanin biosynthetic genes in ripening strawberry, but in this case the genes codes for a R2R3-MYB.

To date, it is evident in plants that R2R3-MYB TF are the main regulators for anthocyanin biosynthesis and are located in the upstream cascade (Martin and Paz-Ares 1997, Nesi et al. 2001, Ban et al. 2007) of structural genes encoding enzymes for the anthocyanin biosynthesis pathway (Stracke et al. 2007). In *Arabidopsis*, there are 126 R2R3-MYB TFs (Yanhui et al. 2006) and some of them regulate flavonoid biosynthesis. In particular, *PRODUCTION OF ANTHOCYANIN PIGMENT1* (AtPAP1/AtMYB75), *PAP2* (AtPAP2/AtMYB90), *PAP3* (AtPAP3/AtMYB113) and *PAP4* (AtPAP4/AtMYB114) are involved in anthocyanin production (Borevitz et al. 2000, Nesi et al. 2001, Ramsay and Glover 2005, Stracke et al. 2007, Gonzalez et al. 2008, Heppel et al. 2013), and *AtMYB123/TRANSPARENTTESTA2* (*TT2*) regulates the proanthocyanidin (PA) biosynthesis (Lepiniec et al. 2006). In addition to *Arabidopsis*, anthocyanin promoting MYB TFs have been studied in many species, for instance in tomato: *ANT1* (Mathews et al. 2003); *Petunia*: *AN2* (Quattrocchio et al. 1999); *Capsicum*: *A* (Borovsky et al. 2004); Grape: *MYB1a* (Kobayashi et al. 2002); Maize: *P* (Grotewold et al. 1991); sweet potato: *IbMYB1* (Mano et al. 2007); snapdragon: *ROSEA1*, *ROSEA2* and *VENOSA* (Schwinn et al. 2006); Gerbera: *GhMYB10* (Elomaa et al. 2003); *Malus* × *domestica MdMYB10*, *MdMYB1*/*MdMYBA* (Takos et al. 2006, Ban et al. 2007, Espley et al. 2007, Lin-Wang et al. 2010); strawberry:

FaMYB10, *FaMYB1* and *FcMYB1* (Lin-Wang et al. 2010; Salvatierra et al. 2013); Chinese bayberry: *MrMYB1* (Huang et al. 2013b); and Pear: *PcMYB10* (Feng et al. 2010). The R2R3-MYB proteins that regulate anthocyanin biosynthesis in dicot plants have a conserved amino acid sequence motif [A/S/G]NDV in the R2R3 domain (Lin-Wang et al. 2010). Many R2R3-MYB TFs depend on a co-activator called basic helix-loop-helix (bHLH) to function (Borevitz et al. 2000, Zimmermann et al. 2004). The R3-domain of R2R3-MYB TFs has a conserved amino acid sequence signature [DE]Lx$_2$[RK]x$_3$Lx$_6$Lx$_3$R that specify interaction between MYB and R/B-like bHLHs (Zimmermann et al. 2004, Fraser et al. 2013).

Some R2R3-MYB TFs are flower specific, for example MYB305-related proteins play an important role in the activation of phenylpropanoid biosynthetic genes in tobacco and snapdragon flowers (Sablowski et al. 1994). Another study showed that tobacco MYB305 has specific role in regulating major nectarin genes and flavonoid biosynthetic genes in the flower (Liu et al. 2009).

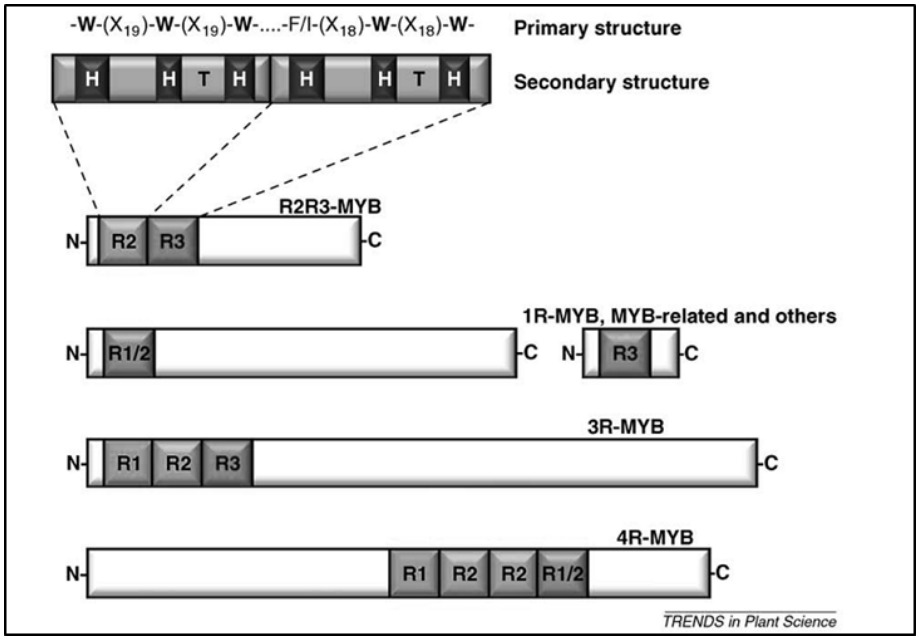

Figure 1.8 The various types of MYB TFs found in plants. The letters R, H, T, W and X are corresponding to the MYB domain, helix, turn, tryptophan and any amino acid, respectively (adapted from Dubos et al. 2010).

1.9.2 The family of basic helix-loop-helix (bHLH) TFs

The bHLH was first identified by (Murre et al. 1989) in animal system. The bHLH proteins belong to a super family of transcriptional regulators that bind as dimers to

specific DNA target sites and found in three eukaryotic kingdoms but well characterized in non-plants eukaryotes (Toledo-Ortiz et al. 2003). There are 162 bHLH TFs in *Arabidopsis thaliana* (Bailey et al. 2003), grouped into 12 major groups, some of which further divided into subfamilies (Heim et al. 2003, Toledo-Ortiz et al. 2003). This TF is known to be evolved from the ancestral group B class of bHLH genes (Heim et al. 2003). In plants, bHLH TFs have functional diversity in conjugation with MYB TFs (reviewed from Allan et al. 2008). The bHLH proteins are characterized by a domain, called bHLH, that consists of about 18 hydrophilic and six basic amino acids residues at the N-terminal end (Figure 1.9) and is involved in DNA binding, followed by two distinct regions of hydrophobic residues which form α-helices separated by an intervening loop at the C-terminal end of the protein (Murre et al. 1994, Atchley et al. 1999). The C-terminal HLH region functions as a dimerization domain (Ferré-D'Amaré et al. 1993). In addition, some bHLH proteins bind to sequences containing a consensus core element like the E-box (5′-CANNTG-3′), and the G-box (5′-CACGTG-3′). Like non-plant bHLH proteins, many plant bHLH proteins have specific amino acid residues at 5, 9, and 13 (H-E-R) which are the most critical for DNA binding (Heim et al. 2003). Sub-groups IIId, IIIe, and IIIf have a conserved stretch of amino acids positioned towards the N-terminal end which is specific for plant bHLH proteins. In *Arabidopsis*, the subgroup IIIf comprises AtTT8 (AT4G09820), AtMYC1 (AT4G00480), AtGL3 (AT5G41315), and AtEGL1 (AT1G63650). Among these bHLH TFs (subgroup IIIf), AtTT8 and AtMYC1 are involved in flavonoid/anthocyanin biosynthesis while AtGL3 is involved in trichome initiation (Heim et al., 2003). The R/B-like bHLH proteins of sub-group IIIf interact with MYB proteins and a member of this sub-group has a possible functional involvement in *TTG1*-dependent regulatory processes (Zimmermann et al. 2004). In several species the functionality of bHLH proteins acting as co-activators with anthocyanin related MYB proteins have been tested, such as B and R in maize; AN1 and JAF13 in *Petunia*; TT8 with PAP1 in *Arabidopsis*; MYC-F3G1 and MYC-GP/RP in *Perilla*; and DELILA with ROSEA1, ROSEA2, and VENOSA (MYB TFs) in snapdragon (Goff et al. 1992, Spelt et al. 2000, Springob et al. 2003, Schwinn et al. 2006).

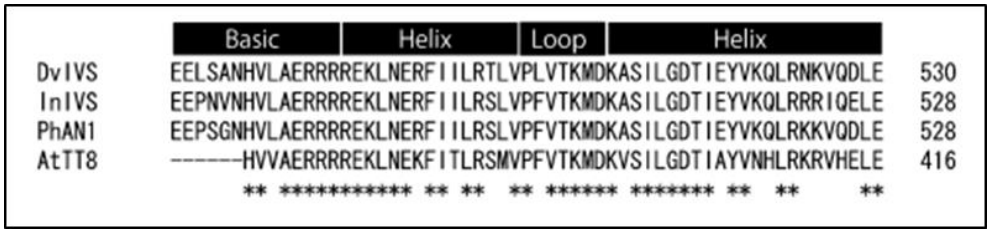

Figure 1.9 The bHLH domain of known bHLH proteins. Asterisks indicates amino acids that are fully conserved in each of the proteins (adapted from Ohno et al. 2011).

1.9.3 WD-40 repeats

The WD-40 repeat (WDR), also called β-transducin, is found in a multitude of eukaryotic proteins and is associated with a range of critical functions from signal transduction, transcription, autophagy and apoptosis (Stirnimann et al. 2010). These proteins are characterized by the presence of 44-60 residue sequence which contain glycine-histidine (GH) dipeptide at the N-terminus and the tryptophan-aspartate (WD) dipeptide at the C-terminus end (reviewed by Ramsay and Glover, 2005). The WDR proteins have 4 to 16 repeating units, all of which are thought to form a circularized beta-propeller structure that makes a stable platform to form reversible complexes with several proteins (Smith et al. 1999, Li and Roberts 2001). The repeating unit acts as a site for protein-protein interactions. The best characterized WDR protein is the Gβ subunit of heterotrimeric G proteins, which forms a tight dimer (Gβγ) with Gγ subunits. In *Arabidopsis*, most of the WDR proteins are strongly conserved and involved in a range of plant-specific processes (van Nocker and Ludwig 2003). Several WDR involved in anthocyanin biosynthesis have been described in different plants like TRANSPARENT TESTA GLABRA 1 (TTG1) in *Arabidopsis* (Berger et al. 1998, Walker et al. 1999), ANTHOCYANIN11 (AN11) in *Petunia* (de Vetten et al. 1997), Perilla frutescens WD (PFWD) in *Perilla* (Sompornpailin et al. 2002), PALE ALEURONE COLOR1 (PAC1) in maize (Carey et al. 2004), *Malus domestica* TRANSPARENT TESTA GLABRA 1 (MdTTG1) in apple (An et al. 2012) and *Vitis vinifera* WDR1/2 (VvWDR1/2) in grape (Matus et al. 2010). In *Arabidopsis*, the WDR proteins are central to all aspects (Figure 1.10) of epidermal cell fate, positively regulating trichome formation, anthocyanin production, seed-coat pigmentation and seed-coat mucilage production, and negatively regulating hypocotyl stomatal-cell identity and root-hair formation (Zimmermann et al. 2004, Ramsay and Glover 2005, Heppel et al. 2013).

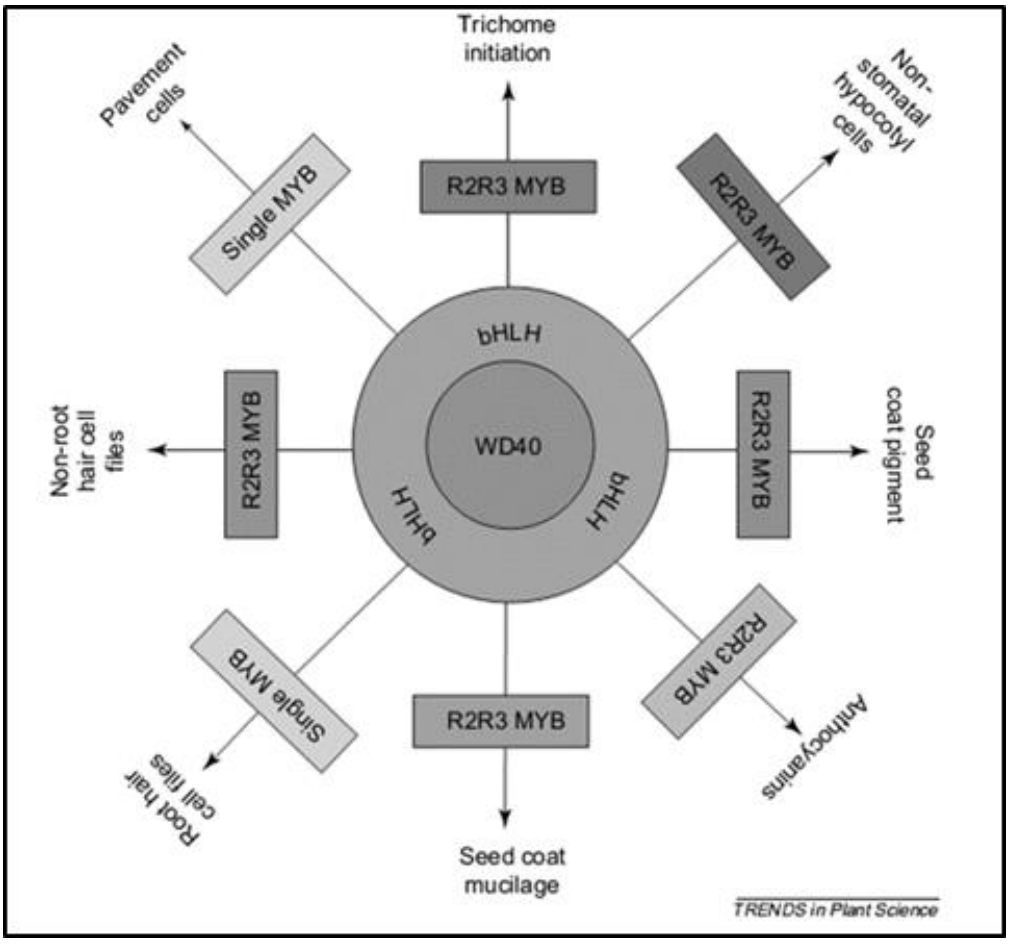

Figure 1.10 Schematic presentation of interactions between WD40, bHLH and MYB transcription factors in *Arabidopsis* resulting in different aspects of cell differentiation and metabolic regulation (adapted from Ramsay and Glover 2005).

1.10 The MBW complex

In general, a protein complex, called "MBW" and comprised of a MYB, a bHLH and a WDR protein (Baudry et al. 2004) regulates the expression of various distinct target genes for anthocyanin and proanthocyandin biosynthesis in a wide range of plant species (Larkin et al. 1999, Grotewold et al. 2000, An et al. 2012, Schaart et al. 2013, Zhao et al. 2013). This MBW transcription complex was found in *Arabidopsis*, snapdragon, Japanese morning glory, *Petunia*, apple and grape (de Vetten et al. 1997, Quattrocchio et al. 1999, Spelt et al. 2000, Zhang et al. 2003, Morita et al. 2006, Schwinn et al. 2006). One such MBW complex model was proposed by Koes et al. (2005) for the transcriptional modulation of structural genes for anthocyanin synthesis (Figure 1.11). In a MBW protein complex, the DNA

binding specificity is determined by the specific amino acid residues present in the MYB domain (Carr et al. 1996, Heppel et al. 2013). A large number of MYB TFs has overlapping functions due to differences in the target genes and tissue specificity (Jin and Martin 1999). Another member of the complex is the bHLH which usually shows a more functional specificity (Zimmermann et al. 2004). There are several evidences of interaction between MYB and bHLH proteins for the transcriptional regulation of target genes of anthocyanin biosynthesis. For example ZmC1 (MYB TF) and members of *R* gene family (bHLH TFs) in maize; AN2 (MYB TFs) and AN1/JAF13 (bHLH TFs) in *Petunia;* and ROSEA1, ROSEA2 and VENOSA (MYB TFs) and MUTABILIS and DELILA (bHLH TFs) in Snapdragon; MdMYB10 (MYB TF) and MdbHLH3/MdbHLH33 (bHLH TFs) in apple (Ludwig and Wessler 1990, Goff et al. 1992, Goodrich et al. 1992, Mol et al. 1998, Schwinn et al. 2006, Espley et al. 2007). In addition to these two members, WDR is involved as a third factor in MBW complex for anthocyanin biosynthesis. Carey et al. (2004) showed that PAC1 encodes a WDR that is required for anthocyanin accumulation in the aleurone and scutellum of the maize seed. In *Arabidopsis*, TTG1 (WDR) protein in combination with PAP1 (MYB) and GLABRA3/ENHENCER OF GLABRA3 (GL3/EGL3, bHLH) proteins regulate the transcription of target genes for anthocyanin biosynthesis while it also activate transcription of PA biosynthesis together with TRANSPARENT TESTA2 (TT2, MYB) and TRANSPARENT TESTA8 (TT8, bHLH) proteins (Zhang et al. 2003, Baudry et al. 2004, Zimmermann et al. 2004, Gonzalez et al. 2008). There are reports on different plant species that endogenous bHLH TFs and WD40 proteins are sufficient for forming a complex with exogenous R2R3-MYB TFs to induce the anthocyanin biosynthetic genes in transgenic plants (Mathews et al. 2003, Espley et al. 2007).

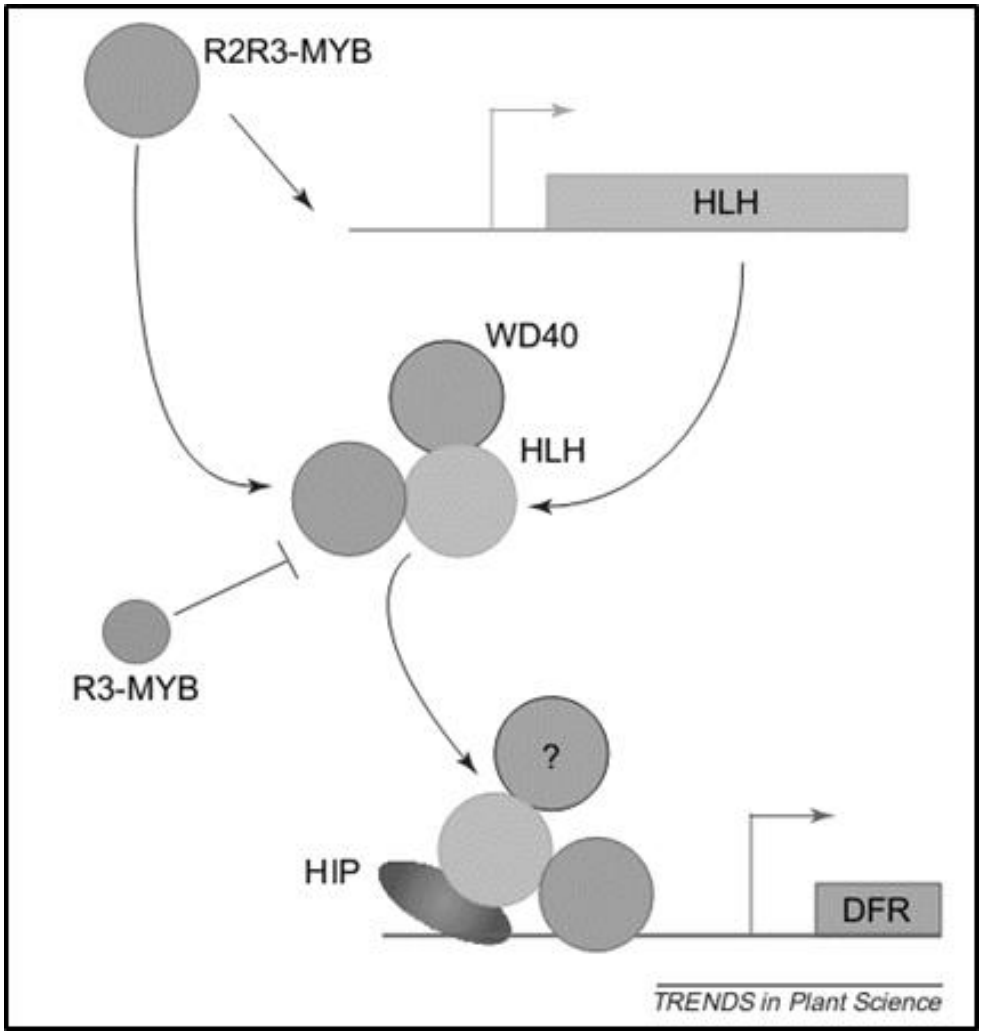

Figure 1.11 A model for coordinated role of MYB, HLH and WD40 regulators in transcription activation of a structural pigmentation gene (DFR). Genes, including upstream sequences are indicated by rectangles. (Some) of the MYBs control transcription of (some) of the HLH factors, and subsequently form a complex that also involves the WD40 protein. Whether the WD40 protein is also present in the transcription complex on the promoter of the structural gene (DFR) is unclear, indicated by the question mark. The MYB factors contact the DNA directly, whereas the HLH proteins probably bind indirectly via a hypothetical HLH INTERACTING PROTEIN (HIP). The small R3-MYB acts as an inhibitor, probably by sequestration of the HLH protein into an inactive complex (adapted from Koes et al. 2005).

1.12 Aim of the work

Fleshy fruits undergo a series of biochemical and physiological changes that are especially relevant during the final part of their development, when they have to ripe. In general, fleshy fruits signal ripening by changing their appearance from green to more trumpery colors, by loosening firmness (from hard to soft), by acquiring particular flavors (synthesis of aroma) and taste (acidic to sweet). Therefore, ripe fruits are normally sweeter, more colorful and fragrant, thus becoming attractive to animals. By eating the fleshy fruits, frugivorous animals disperse the enclosed seeds through their dejections. Since animals move they can disperse seeds for long distances, thus increasing the evolutionary success of plants. It is believed that this process may have led to the evolution of a large variety of fleshy fruits. Like any other animal, also mankind is attracted by the color of fruits. In many fleshy fruits the color (red, purple and blue) is developed through accumulation of anthocyanin pigments. Besides being attractive, these pigments also have anti-oxidant properties so they may help protecting the human body from different chronic diseases and improve the immunity level. In peach, color is a key quality trait for the market acceptance. In particular, the red color, both in the peel and in the flesh, is due to the accumulation of anthocyanins. The most important aim of this work was to identify genes involved in the regulation of the anthocyanin biosynthesis in peach and to functionally characterize them. Hopefully, the achieved results will be transferred to peach breeders to improve the anthocyanin's level in the fruit.

Chapter II

2 Materials and methods

2.1 Chemical solutions and media

TAE 1X : Tris-Acetate 40 mM
EDTA 1 mM pH 8.0

TE : Tris-HCl 10 mM
EDTA 1 mM pH 8.0

DNA extraction buffer

Sorbitol : 0.35 M
Tris : 0.1 M
EDTA : 5 mM
pH : 8.0

Lysis buffer

Tris : 0.2M
EDTA : 0.8M
NaCl : 2 M
CTAB : 2%

RNA extraction buffer

CTAB : 2%
PVP K30 : 2%
Tris-HCl pH 8.0 : 100 mM
EDTA pH 8.0 : 25 mM
NaCl : 2 M
Spermidin : 0.5 g/L
β-mercaptoethanol : 2%

Total soluble protein extraction buffer

$NaPO_4$ pH 7.0	:	50 mM /L
EDTA 0.5 M pH 8.0	:	10 mM/L
β-mercaptoethanol	:	2 mM/L (freshly added)
Triton X-100	:	0.1%

Reaction buffer for enzymatic GUS assay

$NaPO_4$ pH 7.0	:	50 mM /L
EDTA 0.5 M pH 8.0	:	10 mL/L
β-mercaptoethanol	:	2 mM/L (freshly added)
TritonX-100	:	0.1%
MUG	:	0.77 mg/L

Stop buffer for enzymatic GUS assay

Na_2CO_3	:	0.2 M/L

Reaction buffer for histochemical GUS assay

$NaPO_4$	:	20 mM pH 7.0
EDTA	:	10 mM pH 8.0
TritonX-100	:	0.1%
$K_3Fe(CN)_6$	:	0.5 mM
$K_4Fe(CN)_6$	:	0.5 mM
Methanol	:	20%
X-Gluc	:	0.521 g/L

LB medium

Tryptone	:	10 g/L
Yeast extract	:	5 g/L
NaCl	:	10 g/L
Agar (for solid media)	:	15 g/L
pH	:	7.0

YEB medium

Beef extract	:	5 g/L
Yeast extract	:	5 g/L
Tryptone	:	1 g/L
Sucrose	:	5 g/L
Agar (for solid medium)	:	15 g/L
$MgSO_4$	:	0.049 g/L

SOC broth medium

Tryptone	:	20 g/L
Yeast extract	:	5 g/L
NaCl (10 mM)	:	0.5 g/L
KCl (10 mM)	:	0.19 g/L
$MgCl_2$ (10 mM)	:	0.95 g/L
$MgSO_4$ (10 mM)	:	1.2 g/L
Glucose (20 mM)	:	3.6 g/L

Murashige and Skoog (MS) medium

MS salts including vitamins	:	4.4 g/L
Sucrose	:	30 g/L
Plant agar	:	6 g/L
pH	:	5.6-5.8

TAB1 medium

MS medium plus:

6-Benzylaminopurine (6-BAP)	:	1 mg/L
Indole acetic acid (IAA)	:	0.2 mg/L
Sucrose	:	30 g/L
Plant agar	:	6 g/L
pH	:	5.6-5.8

TAB2 medium

TAB1 medium plus:

Kanamicyn	:	200 mg/L
Cefotaxime	:	500 mg/L
Hygromycin	:	15 mg/L

TAB3 medium

MS medium plus:

Kanamicyn	:	200 mg/L
Cefotaxime	:	500 mg/L
Hygromycin	:	15 mg/L

MMA medium

MS salts including vitamins	:	4.4 g/L
MES	:	2.13 g/L
Sucrose	:	20 g/L
Acetosyringone	:	200 µM
pH	:	5.6

Infiltration medium for stable transformation

MS salts including vitamins	:	4.4 g/L
6-BAP	:	1 mg/L
Sucrose	:	20 g/L
pH	:	5.6

2.2 Bacterial strains

Agrobacterium tumefaciens	:	GV3101
Escherichia coli	:	DB3.1

2.3 Plant materials

Tobacco (*Nicotiana tabacum*)	:	Cultivar (cv) ‘Samsung *NN*’
Peach (*Prunus persica*)	:	cvs ‘Fntasia’, ‘Roza’, ‘Redhaven’, ‘Stark Red Gold’, and ‘Spring Crest’

Peach (*Prunus persica*) fruit were collected at S4 (ripening) stage from the experimental farm of the University of Padova, Agripolis, Legnaro, Padova, Italy. After collection samples were frozen in liquid nitrogen and stored at -80°C until use.

2.5 DNA extraction

The DNA was extracted from plant materials as described by Fulton et al. (1995). The samples were ground by micro pestle in 1.5 mL micro-centrifuge tube using 750 µL extraction buffer (1 volume of DNA extraction buffer, 1 volume of lysis nuclei buffer, 0.4 volume of 5% sarkosyl and 3-5 mg/mL of $NaHSO_3$) and kept in a water bath at 65°C for 20 minutes. After that, 750 µL of chloroform/3-methyl-1-butanol (24:1 v:v) were added and placed on a shaker for 5 minutes, and centrifuged at 10000 *rpm* for 10 minutes. The DNA was precipitated by isopropanol and centrifuged at 10000 *rpm* for 10 minutes. The DNA pellet was washed with 70% ethanol and dried. The DNA pellets were resuspended in 50 µL Milli-Q water.

2.4 Total RNA extraction

Total RNA was extracted from leaves, flowers and fruits according to Chang et al. (1993). For fruit samples, 2-4 g tissues were ground in liquid N_2. After that, the powder was transferred into a clean RNase-free tube containing 20 mL CTAB extraction buffer with 400µL β-mercaptoethanol preheated at 65°C. Then the powder was mixed with extraction buffer by vigorous shaking and an equal volume of chloroform/isoamyl-alcohol [chloroform : isoamyl alcohol (3-methyl-1-butanol) :: 24 : 1, v : v] was added. Samples were placed on a shaker for 15 minutes and centrifuged at 4000 *rpm* for 15 minutes. The chloroform/3-methyl-1-butanol (24:1 v:v) extraction was repeated 3 times. The RNA precipitation was done with the addition of LiCl (final concentration 2M) keeping tubes overnight at 4°C. On the following day samples were centrifuged at 4000 rpm for 1hour at 4°C and the RNA pellets were washed with 5 mL of cold RNase-free 80% ethanol. Finally, pellets were dried and resuspended in Milli-Q water treated with 0.1% Diethylpyrocarbonate (DEPC). The RNA extraction from tobacco leaf disc and floral part were carried out by scaling down the previous protocol for micro-

extractions with 700 µL of CTAB extraction buffer in a 1.5 mL micro-centrifuge tube.

2.6 Quantification of nucleic acids

The concentration and purity of nucleic acids (DNA and RNA) were checked by UV absorption spectra (Perkin-Elmer UV/Vis Lambda) with four different wavelengths: 230 nm, 260 nm, 280 nm, 320 nm. The quantification was carried out in a quartz cuvette and H_2O mQ was used as blank. The concentrations of nucleic acids were expressed in µg/µL:

- For DNA: OD_{260} * 50 * dilution factor
- For RNA: OD_{260} * 40 * dilution factor

The peak of absorption of proteins is at 280 nm, while at 230 nm is for sugars, so the ratio with OD_{260} allows to understand the purity of the sample. If $OD_{260}/OD_{280} > 1.8$, the sample is free of protein contaminations. Similarly, if $OD_{260}/OD_{230} > 1.8$, the sample is free of sugar contaminations. The integrity of nucleic acids was determined by agarose gel electrophoresis with TAE 1X buffer followed by ethidium bromide staining.

2.7 cDNA synthesis

The concentration of RNA was determined using a NanoDrop Spectrophotometer (NanoDrop Technologies, USA). Then cDNAs were synthesized from 4 µg of total RNA, pre-treated with 1.5 units of DNaseI. Random primers were used in the reaction together with "High Capacity cDNA Archive Kit" (Applied Biosystem) following the manufacturer's instruction.

2.8 Real time PCR

The quantitative real time PCR was used to amplify and quantify the relative transcription level of target genes. Gene specific primers (Table 2.1) were designed using "Lasergene" software package (DNASTAR). The primer selection criteria

were Tm between 55ºC and 65ºC, an amplicon length between 90 and 150 bp and primer length not longer than 24 bases. cDNAs were used as the starting template in the reaction. Reactions were carried out in a total volume of 10 µL using the "Syber green PCR master mix" (Applied Biosystems), with 0.05 pmoles of each forward and reverse primers. The instrument used was the "7500 Real-Time PCR System" (Applied Biosystems) using the following PCR conditions.

50°C	:	2 minutes	
95°C	:	10 minutes (to activate the enzyme)	
95°C	:	15 seconds	} 40 times
60°C	:	15 seconds	
65°C	:	15 seconds	

At the end of PCR reaction, the dissociation curve was performed at 60°C. After completion of reaction, the Ct values were analyzed by "Q-gene" software (Muller et al. 2002) by averaging three independently calculated normalized expression values for each sample. The *PpN1*, *NtUbc2* coding for tobacco ubiquitin-conjugating enzyme E2 and *GUS* genes were used as internal control for peach, the first, and tobacco, the latters.

2.9 PCR amplification

The routine PCR was performed using 50 µl reactions with the following reagents (EuroClone) and conditions.

PCR reagents

DNA	:	120 ng for genomic DNA, 5 ng for cDNA
Reaction buffer 10x	:	5.0 µL
MgCl2 50 mM	:	2.5 µL
dNTPs 10 mM	:	1.0 µL
Primer_forward 10 µM	:	1.0 µL
Primer_reverse 10 µM	:	1.0 µL
Taq 5 U/µL (polymerase enzyme)	:	0.2 µL
mQ H_2O	:	upto 50.0 µL volume

To maximize the efficiency and reduce the error rate of PCR for gene cloning a blend of two different polymerase enzymes (Taq from Euroclone 5 U/μL, 0.20 μL, and Pfu from Promega, 3 U/μl, 0.05 μL) was used.

Table 2.1 List of Primers used for qRT-PCR

Gene name	Primer	Sequence (5'→3')	Accession No.
MYB10.1	F	CAGGAAGGACAGCGAATGATG	ppa026640m
	R	TCGGGGTTGAGGTCTTATTACG	
MYB10.2	F	AGGACTGCGAATAATGTGAAAA	ppa016711m
	R	GGTTGTGGCCTTATTATGGTCT	
MYB10.3	F	GTACGCCATCACAAACATCACC	ppa020385m
	R	CATCATCAACCCAAAAACTCGT	
bHLH3	F	TCTTGTTCAGCGTTCCGTTCCT	ppa002884m
	R	TTGGCGCTGAGCTCATCTTGTG	
HLH33	F	GGCCTGTGACATTGACGAAACTG	Ppa002645m
	R	CTATAAGGGCATCTCATCTCTA	
GL3	F	ATGTCTAGCAGAAGGTCGAGGC	ppa002762m
	R	GGACCTTAGATGCTGATACC	
*CHS**	F	CAGAGATACCCAAAGGTTGGAAGGC	ppa006888m
	R	AACCATCCTTCCCGACAGCGAT	
*CHI**	F	TGAAGACCTCAAGGAACTTCTCAATGG	ppa011276m
	R	ACACAGGTGACAACGATACTGCCACT	
*F3H**	F	TCCGAGGGCAGAGCGAAGAAC	ppa007636m
	R	TTGTGGAGGCTTGTGAGGATTGG	
*DFR**	F	GGTCGTCCAGGTGAACATACTGCC	ppa008069m
	R	ATTTCTCATGCCATCCATGCCAC	
*LDOX**	F	AAGTGGGTCACTGCCAAGTGTGTTC	ppa007738m
	R	GTGGCTCACAGAAAACTGCCCAT	
*UFGT**	F	CCGCTGCCTCTCCCAACACTC	ppa005162m
	R	CCATCAGCCACATCAAACACCTTTAT	
*ANR**	F	ACTTCAAGGCTAAGGGGCTGCTG	ppa008295m
	R	CCAAGCCAGATAAACGCCAATCAC	
*LAR**	F	CATCCACGGGGAAATTCACCTG	ppa007994m
	R	ACCCTTCCCAGAGTTACCATCACTGA	
*FLS1**	F	GTTTTCTGACGGCAACGTTACGAA	ppa008322m
	R	CCCAACCCTAGCGATAGGAGCC	
MYB24	F	CAAGTGGGGAAACAGGTGGTC	ppa011751m
	R	CTGCTTGCTTAATGTGCTTTTGAA	
PpN1	F	CCAGGAGAATCGGTGAGCAGAAAA	ppa009483m
	R	TCGAGGGTGGAGGACTTGAGAATG	
NtMYB305	F	GGACAAGGATTCAGAAGCACATA	EU111679.1
	R	GTTGGACCAGCAGACGACATA	
NtNEC1	F	AAGCCAGGAGCCACAAACAACAAA	AF132671
	R	TCAATTCGAGCGAGGGATACACCA	
NtNEC5	F	CGACGAAAACCCCTAAACCTCTTG	AF503441
	R	ATGCCCTCCACTACGAACCCTAAC	
NtPAL	F	CCCCTTCGCGGCACCATCAC	D17467.1
	R	TGCTTTAGAATTAGGCCGACCAGT	

*The primers for peach anthocyanin biosynthetic genes were kindly provided by Prof.ssa Patrizia Torrigiani, Department of Fruit Tree and Woody Plant Sciences, University of Bologna, Italy.

Table 2.1 Continued.

Gene name	Primer	Sequence (5'→3')	Accession No.
NtCHS	F	CACGGTCATGGCTATCGGAACA	AF311783.1
	R	GCTCAACCTTATGCTCGCTATTA	
NtCHI	F	GCCGGCGCAGGAAATAGAGGT	AB213651
	R	TAGCGGCGAGAAAAGGAAGAGCA	
NtF3H	F	CTTACCCTTGGACTGAAACGACAC	AB289450
	R	CAACGGGCTGAACAGTAATCCA	
NtDFR	F	CAGAGAAGGCCGCAATGGAAGC	AB289448
	R	GGTGGGAATGTAGGCGTGAGGAAT	
NtANS	F	TTAACTACTACCCCAAATGTCC	AB289447
	R	TGCCGTTACCCACTGTCCTTC	
NtAN2	F	GAAGAAAGGTGCATGGACTG	FJ472647
	R	TCTGCAGCTCTTTCTGCATC	
NtAN1b	F	CTTGAACACTTCTCAAACCGA	HQ589209
	R	TGCTAGGGCACAATGTGAAG	
NtUbc2	F	TGAGAACAAGCGCGAATACAACAG	AB026056
	R	AACAGATTAAGAGTGCGGGAGATG	

2.10 Isolation of peach MYB and bHLH genes

The full length coding sequence (CDS) of ppa026640m, ppa002884, ppa002645m and ppa002762m, named *MYB10.1*, *bHLH3*, *bHLH33* and *GL3*, respectively, were amplified by PCR from fruit cDNA of peach cv 'Stark Red Gold' while gene fragments encompassing the CDS with introns of ppa016711m and ppa020385m, named *MYB10.2*, *MYB10.3*, respectively, were amplified from gDNA of 'Redhaven', as it was not possible to get any amplicon from cDNA. The primers used to amplify those TFs are presented in Table 2.2.

Table 2.2 Primers used for cloning of peach MYB and bHLH TFs.

Gene	Primer	Sequence (5'→3')	Accession No.
MYB10.1	F	ATGGAGGGCTATAACTTGGGTGT	ppa026640m
	R	TTAATGATTCCAAAAGTCCACGTT	
MYB10.2	F	ATGGAGGGTTATGACTTGAGTGT	ppa016711m
	R	TTACTTTCTATATTCTTCATTTGAAT	
MYB10.3	F	ATGGGGGGAAATAACTTGGATGT	ppa020385m
	R	TTATTCTTCTTTTGAATGATTCCA	
bHLH3	F	ATGGCTGCACCGCCAAGTAGCA	ppa002884m
	R	CTAGGAATCAGATTGGGGAATT	
bHLH33	F	ATGGCTAATGGGACTCAAAACCAT	Ppa002645m
	R	TCAACACTTACCGGCAATTTTCCA	
GL3	F	ATGGCTACTAGGCTTCAGAACCA	ppa002762m
	R	TCAACAGCTTCTGGCGATTCG	

2.11 Purification of the PCR product

The PCR products were purified by column purification using the EuroGOLD Gel Extraction Kit. PCR reactions were mixed with equal volume of binding buffer and loaded in a PerfectBind DNA column followed by centrifugation at 10000 rpm for 1 minute. Subsequently, the columns were washed with 750 µL of wash buffer and centrifuged twice. Finally, 50 µL of elution buffer was added in the column, then centrifuged at 5000 rpm for 1 minute to allow the DNA elution.

2.12 TOPO TA cloning of peach MYB and bHLH genes

The cloning reactions for peach *MYB* and *bHLH* genes were carried out using TOPO TA Cloning® kit (Invitrogen) following manufacturer's instructions. After that *E. coli* cells were transformed by electroporation with an electric discharge of 1500 V (Invitrogen Electroporator II, capacity 50 µF). Then bacteria were transferred into a vial containing 1 mL of SOC media and incubated at 37°C for 45 minutes. Aliquots of 30 µl and 100 µl of the mixture were plated on LB plates supplemented with 10 µg/mL of spectinomycin. The plates were incubated at 37°C overnight. On the following day, a single colony was inoculated in 3 mL of LB broth with the proper antibiotic and was grown over night at 37°C in a rotary incubator. The right colonies were selected by PCR and the identities of the cloned inserts were confirmed through DNA sequencing.

2.13 Isolation of plasmid DNA

Overnight liquid cultures from single colony were moved into a clean eppendorf tube (about 1.5 mL) and pelleted by centrifugation for 2 minutes at maximum speed. The pellet was resuspended with 200 µL of P1 solution (100 µg/mL RNAse A; 50 mM Tris-HCl, 10 mM EDTA, pH 8.0). Then 200 µL of P2 lysis solution (0.2 M NaOH, SDS 1%) and cold P3 neutralization solution (3.0 M KAc, pH 5.5) were added quickly. The sample was mixed and centrifuged at 4°C with maximum speed for 10 minutes. The supernatant was moved into a clean tube and an equal volume of phenol/chloroform/3-methyl-1-butanol (25:24:1, v:v:v) was added. After mixing and centrifugation, the DNA was precipitated from the aqueous phase with the

addition of 2.5 volumes of ethanol (EtOH) 100% and incubating the tube at -20°C for 15 minutes followed by 15 minutes centrifugation at 16000 rpm at 4°C. The pellet was washed with 1 mL of EtOH 70% and dried. Finally, the pellet was resuspended in 50 μL of mQ H_2O.

2.14 DNA sequencing and analysis

DNA sequencing was performed at the BMR-genomics sequencing facility using a PCR based dideoxynucleotide terminator protocol and an automated sequencer. Sequences were determined on both strands using, when necessary, chemically synthesized oligonucleotides. Sequence manipulations, analyses, and alignments were performed using the 'Lasergene' software package (DNASTAR, USA). Multiple sequence alignments were performed with the ClustalW online program (http://embnet.vital-it.ch/software/ClustalW.html). Peach sequences used to construct the sequence alignments were deduced from the nucleic acid sequences, obtained from www.phytozome.net, while those of other species were obtained from public databases (accession numbers are given in the figure legends).

2.15 The over-expression system

The transcription factors (TFs) have some overlapping and redundant functions (Riechmann and Ratcliffe 2000). For example, several MADS-box genes are showing functional redundancy in *Arabidopsis* (Pinyopich et al. 2003). Similar functional redundancy and antagonistic effects among R2R3-MYBs and bHLH were also revealed by several old and recent studies (Martin and Paz-Ares 1997, Stracke et al. 2007, Dubos et al. 2010, Bomal et al. 2013, Payne et al. 2000). Therefore, over-expression of transcription factor genes by the CaMV 35S constitutive promoter may lead to difficulties for functional analysis. Such strong promoter (35S) can be detrimental to plant growth, development, and reproduction and may lead to pleiotropic effects. To overcome these barriers, a dual components transcriptional activation system was used. It was developed by Moore and his co-workers (Moore et al. 1998, 2006) and the vectors were modified by Busatto (2012) in Prof. Trainotti's Laboratory. This over-expression system includes two modular

activities: a DNA-binding activity fused with a transcriptional activation and a cis-acting module with a minimal promoter driving the expression of the gene of interest, each of which used to prepare specific transgenic lines called the driver lines and the responder lines, respectively. The modified vector system is based on the transactivation between the chimeric TF resident on plasmid 35S::LhG4^{At0} and the gene of interest cloned into pH-TOP::GWA. The vectors are derived from pGreen, (Hellens et al. 2000) whose plasmid copy number is low and contributes less T-DNA molecules during the transformation process (Busatto 2012). For these reasons, a dual component promoter system based on transactivation was used for both transient (*MYB10*s and *bHLH*s) and stable (*MYB10.1*) transformations.

2.16 Binary plasmid construction

After sequencing, the cloned DNAs were moved into a binary vector modified in house from pH-TOP (Craft et al., 2005) in order to contain *attR* sites and a GUS reporter gene interrupted by a plant intron (Vancanneyt et al, 1990). The cloning reaction mixtures contained 1 µL of destination vector, 2 µL of LR Clonase II, 100 ng of insert (DNA fragment to be cloned), and final volume up to 10 µL with TE Buffer (pH 8.0). The reaction mixtures were incubated at room temperature for 1 h and the reactions were terminated by adding 1 µL of proteinase K (for each reaction) followed by an incubation at 37°C for 10 min. Finally, 2 µL of these mixtures were electroporated into competent *E. coli* cells strain DB3.1 which is resistant to ccdB gene. The transformed cells were selected both by antibiotic positive selection and by ccdB (control of cell death) negative selection. The ccdB gene is maintained in non-recombinant vectors and it leads lethal effect in most *E. coli* strains. Then plasmid DNA was isolated only from the positive colonies and used to transform *Agrobacterium tumefaciens*. The CDSs in the final expression vectors were under the control of a pOp promoter, recognized by the synthetic transcription factor LhG4, cloned on an independent plasmid under the control of the CaMV 35S promoter (construct LhG4 in Craft et al. 2005).

2.17 Transformation of *Agrobacterium tumefaciens*

The binary vectors harboring the desired constructs were transferred into *A. tumefaciens* strain GV3101 (competent *Agrobacterium* cells containing pSOUP with tetracycline resistance) by mixing 0.5-1 μg of plasmid DNA followed by incubation for 5 min on ice, 5 min in liquid nitrogen and 5 minutes at 37°C. The mixture was diluted with 1 mL of YEB and it was shaken for 4 hours at 28°C and bacteria were plated on YEB plates with proper antibiotics (kanamycin 50 mg/L, gentamycin 25mg/L, rifampicin 100 mg/L).

2.18 Stable transformation of *Nicotiana tabacum*

The stable tobacco transformations were carried out according to a previously described method (Fisher and Guiltinan 1995). For the transformation, the transformed *Agrobacteria* were grown overnight in liquid culture of YEB medium supplemented with proper antibiotics at 28°C. Cultures were centrifuged at 3000 x g for 20 min at 22°C and pellets were resuspended in infiltration medium to a final OD_{600} of 0.8. Then healthy, green and undamaged tobacco leaves were collected from plants grown in vitro and parallel cuts were realized with a knife on the leaf surface. The leaf petioles were also removed. These leaves were soaked in the *Agrobacterium* culture for 10 minutes, dried on sterile chromatography paper and then placed on TAB1 (without antibiotics) co-cultivation medium. Then they were allowed to grow two days in the dark at 25°C; thereafter, leaves were washed by immersion in MS liquid medium, dried on sterile chromatography paper and finally moved into TAB2 plates (shooting medium with selection by kanamycin and hygromycin) and grown in a climate chamber at 25°C with a photoperiod of 16 h light and 8 h dark until shoot formation. When shoots were of 1-3 cm in height they were slashed with a cut of 45° and were transferred into plates with TAB3 (rooting medium with selection by kanamycin and hygromycin) medium. The individual shoot was labeled with a serial number and the obtained plantlets were transferred to soil and left the in the climate chamber for hardening when they started to root. After 3-4 day of hardening, they were placed in the greenhouse. Before transferring to soil, each independent clone was genotyped by PCR to confirm the transgene presence.

2.19 Transient transformation of tobacco

Wild type (WT) tobacco 'Samsun *NN'* plants were grown in the glasshouse at the Department of Biology, University of Padova for transient transformation. The growth and induction of *Agrobacterium* were carried out according to Kapila et al. (1997) with modifications described by Spolaore et al. (2001). During the transactivation experiments, the driver and responder strains were mixed in a 1:1 ratio (or 1:0.5:0.5 when both the MYB and the bHLH were tested together) for the agro-infiltration. The *Agrobacterium* suspension was infiltrated in the abaxial surface of the tobacco leaves with a sterile 2 mL injection syringe without needle.

2.20 Transient transformation of peach fruit

The peach cv 'Springcrest' was used for agro-infiltration. Fruits between S3II and S4I stage of development were collected still attached to their branch from the experimental farm of the University of Padova, Italy. The bacterial suspension was evenly injected into single halves of fruits by means of a sterile 1 mL hypodermic syringe and the injected area was labelled (Spolaore et al. 2001). After agro-infiltration, branches with fruits were placed in a pot containing water. Finally, they were moved to a growth chamber at 22 °C for 72 h with a 16 h light and 8 h dark conditions until sample collection.

2.21 Seeds sterilization

Tobacco seeds were surface-sterilized by washing them with 5% bleach and Tween-20 0.1% for 20 min followed by four rinses with sterile mQ H_2O of 10 min each.

2.22 Histochemical GUS assay

The *uidA* gene (also called *GUS*) encodes a β-glucuronidases enzyme which is commonly used as a reporter gene in plants since they do not have intrinsic *glucuronidas*e activity (Jefferson et al. 1987). The histochemical assay was done by immersing plant tissues in histochemical GUS-buffer which contains X-Gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronide), a substrate of β-glucuronidase. Then samples were infiltrated with vacuum to enhance the penetration of the GUS-buffer

into the tissues and incubated at 37°C overnight to carry on the reaction. During the reaction, X-Gluc is cleaved to produce glucuronic acid and chloro-bromoindigo. The chloro-bromoindigo dimerizes to produce the insoluble blue precipitate dichloro-dibromoindigo when oxidized. The blue staining intensity is related with the promoter activity. Thereafter, samples were bleached with a solution of methanol and glacial acetic acid (1:3 ratio) and treated plant tissues were preserved in 70% ethanol.

2.23 Enzymatic GUS assay

The activity of the β-glucuronidase reporter enzyme was assayed with the substrate 4-methylumbelliferyl-ß-D-glucuronide (MUG). The β-glucuronidases catalyze the hydrolysis of β-D-glucuronic acid residues and release the fluorescent molecule 4-methylumbelliferone. Soluble proteins were extracted from frozen plant tissues in liquid nitrogen by grinding with micro-pestles in 1.5 mL eppendorf tubes. Then 600 µL of protein extraction buffer were added to each tube. The homogenate was centrifuged twice at 16000 x g for 15 min and the clear supernatant was moved into a new 1.5 mL tubes. The GUS enzymatic assay was carried out by incubating 50 µL of protein extract with 350 µL of reaction buffer, containing the MUG, at 37°C. The released 4-methylumbelliferone (4-MU) was quantified with a table fluorometer (GE Helthcare) according to the manufacturer's instructions. The fluorescent values were converted to 4-MU pmoles, that were used to plot a line, whose slope represents how quickly the 4-MU (MU/min) was released. The GUS activity was expressed as nmol 4-MU released $min^{-1}µg^{-1}$protein (Jefferson et al. 1987). 50 µL of the reaction mix were withdrawn at serial time intervals (5 min, 20 min and 60 minutes) and the reaction was stopped in 950 µL 0.2 M Na_2CO_3. The data points were normalized on protein content, quantified according to the standard Bradford protocol [5 µL of protein extract mixed with 150 µL of Bradford solution (Bradford 1976)]. The combination between enzymatic activity values and protein quantification allows to calculate the final activity data as mol of 4-MU released min^{-1} $µg^{-1}$of protein.

2.24 Anthocyanin extraction and quantification

Total anthocyanins were extracted from plant material using methanol supplemented with 1% HCl. After grinding in liquid N_2, samples were incubated overnight at room temperature in extraction buffer. Then extracts were centrifuged at 16000 *g* for 15 min and absorption of the extracts was determined at 530 and 657 nm by means of a spectrophotometer. Finally, anthocyanins were quantified according to the formula proposed by Mehrtens et al. (2005) as follows: $Q_{Anthocyanins} = (A_{530}–0.25\times A_{657}) \times M^{-1}$; $Q_{Anthocyanins}$: amount of anthocyanins, A_{530}: absorption at 530 nm wavelength, A_{657}: absorption at 530 nm wavelength, and M: fresh weight (g) of the tissues.

2.25 Pollen viability test

The pollen viability of transgenic tobacco was assessed by staining grains in 1% 2,5-diphenyl monotetrazolium bromide (MTT) in 5% sucrose (Norton 1966, Khatun and Flowers 1995, Rodriguez-Riano and Dafni 2000). Tobacco flowers were collected when anthers started to burst and pollen grains were dispersed in 800 µL of MTT solution in eppendorf tubes for 10 min followed by centrifugation at 10000 x g for 1 min. Heat killed (at 80°C for 2 h) WT tobacco pollen grains were used as negative controls. 10 µL of suspension were placed in a BÜRKER chamber to count pollen grains under a microscope (LEICA DM5000B). Pollen grains were considered viable only when they turned deep pink (Wang et al. 2004).

2.26 Electronic microscopy

Fresh tobacco leaves at the same developmental stages were observed under low-pressure conditions by means of *Environmental Scanning Electron Microscopy* (*ESEM*). The experiment was performed using a FEI Quanta 200 instrument at the CUGAS facilities of University of Padova.

2.27 Light microscopy

The microscopic observations were performed by directly placing the specimens under the objective lens using a LEICA MZ16F stereo microscope.

2.28 Statistical analysis

Correlation analysis used in this study were performed by Pearson's correlation coefficient (r) and the p-values were calculated using online program 'p-value calculator for correlation coefficients' (http://www.danielsoper.com/statcalc3/).

Chapter III

3 Results

3.1 Regulation of anthocyanin biosynthesis in peach

3.1.1 Identification of peach genes coding for MYB10 and GL3/EGL3 bHLH transcription factors (TFs)

A phylogenetic study was carried out to select R2R3-MYB and bHLH TFs since they are the best candidates for a possible involvement in the regulation of anthocyanin synthesis in peach during fruit ripening. In *Arabidopsis*, there are 126 R2R3-MYB TFs (Yanhui et al. 2006) which are subdivided into 24 subgroups (Stracke et al. 2001). Among these, AtMYB75/PAP1 (AT1G56650), AtMYB90/PAP2 (AT1G66390), AtMYB113/PAP3 (AT1G66370) and AtMYB114/PAP4 (AT1G66380) belong to subgroup 6 and control anthocyanin biosynthesis in vegetative tissues (reviewed by Dubos et al. 2010). In addition, AtMYB123 (AT5G35550)/Transparent Testa 2 (TT2) (subgroup 5) regulates the synthesis of proanthocyanidins (PAs) in the seed coat (Lepiniec et al. 2006). A bootstrap circular tree was generated using MEGA5 (Tamura et al. 2011) to compare peach MYB TFs with those of the model plant species *Arabidiopsis* (sequences of both species were retrieved from PlantTFDB, http://planttfdb.cbi.edu.cn/). The analysis showed that only six peach MYB-like sequences cluster with *Arabidopsis* anthocyanin-related MYB TFs (Figure 3.1). An additional phylogenetic tree was constructed in order to understand the relationships of these six peach and four *Arabidopsis* MYB genes with similar genes known to be involved in the regulation of anthocyanin pigmentation in other dicot species (Figure 3.2). Three peach genes clustered together with those of fruit species belonging to Amygdaleae, while ppa024617m and ppa018744m clustered with strawberry and raspberry (both Rosoideae) sequences. Five peach genes appeared separated from the cluster containing those of both rosids (*Arabidopsis* and grape) and asterids (snapdragon, petunia and tomato) species. The ppa010069m did not cluster with any known anthocyanin promoting MYB. Interestingly, the

genes from pome fruits (apple, pear, medlar, quince, loquat, all of them members of Maloideae) formed a separate clade. Most of these genes were named according to apple *MdMYB10* (Espley et al. 2007).

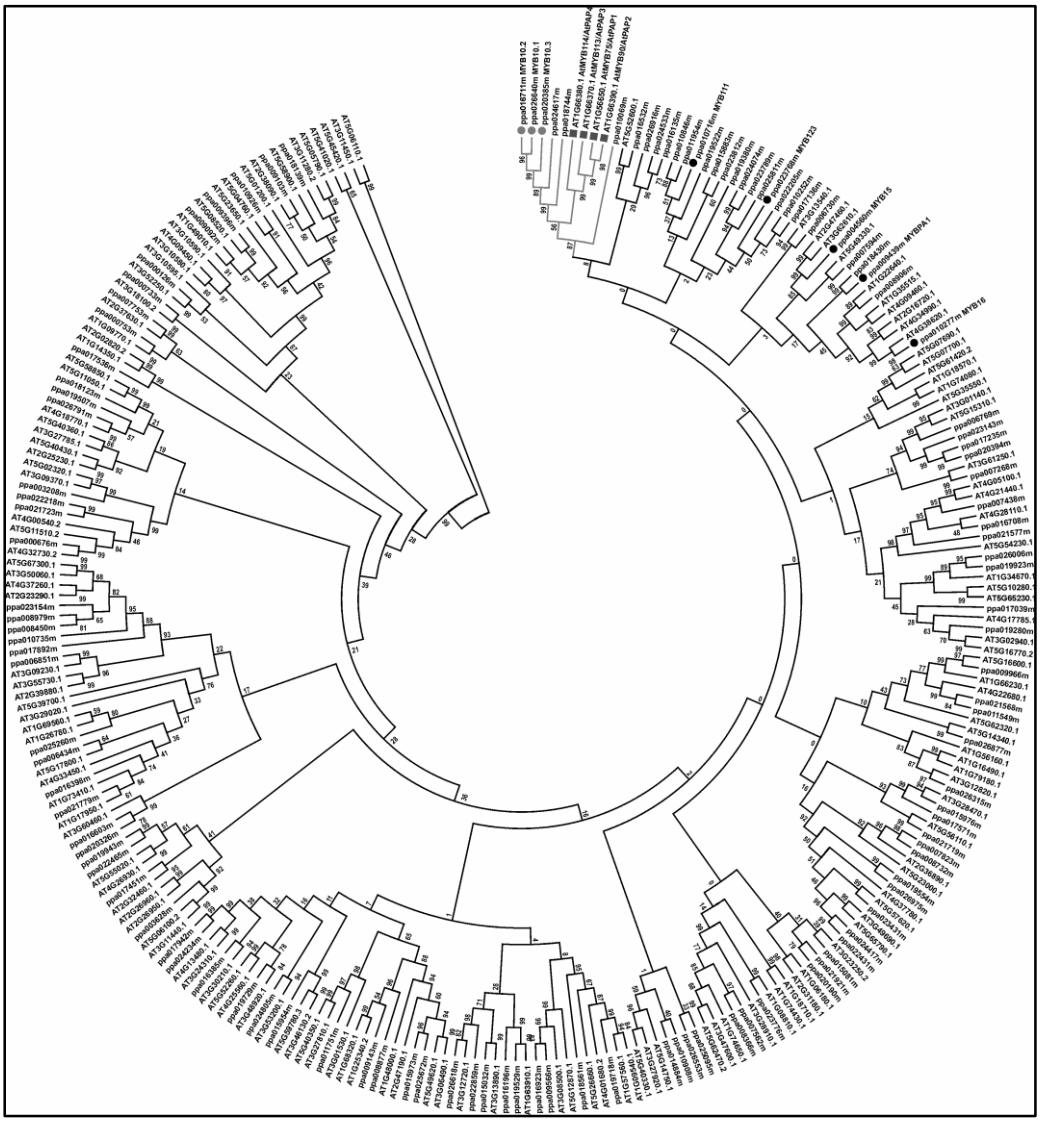

Figure 3.1 Phylogenetic tree showing relationships between *Arabidopsis* and peach MYB TFs. Grey solid squares: *Arabidopsis* anthocyanin promoting MYB TFs; grey solid circles: peach MYB10-like TFs; black solid circles: peach MYB TFs named according to Ravaglia et al. (2013). Phylogenetic and molecular evolutionary analyses were conducted using neighbor-joining method and 1000 bootstrap replicates using MEGA version 5 (Tamura et al. 2011).

Of the six peach sequences potentially regulating anthocyanin synthesis, three were found to be expressed in the fruit (Figure 3.3) by means of microarray analyses ((Trainotti et al. 2012) the microarray data available in Prof. Trainotti's Laboratory) and were named MYB10.1 (ppa026640m), MYB10.2 (ppa016711m) and MYB10.3 (ppa020385m), respectively. These were selected to be studied in

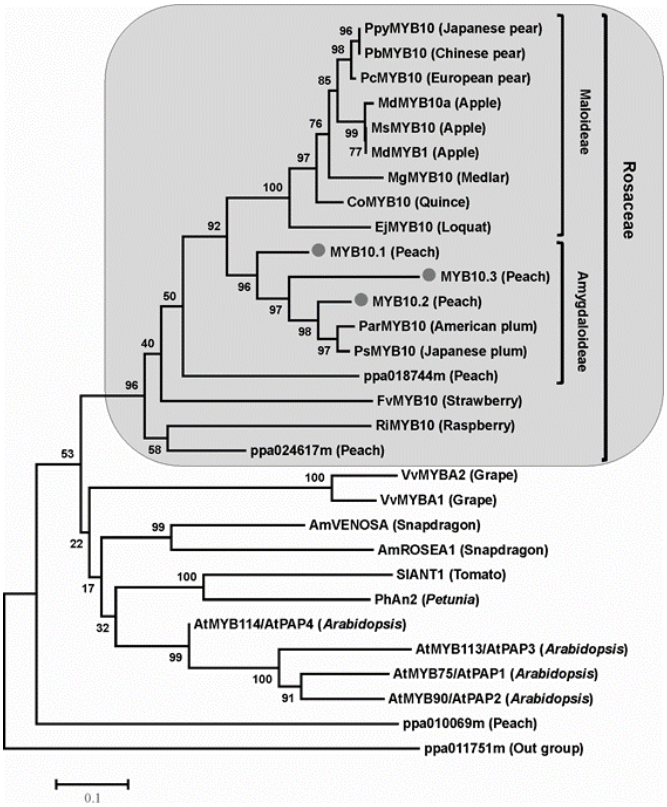

Figure 3.2 A phylogenetic tree of peach MYB10-like TFs with other known MYB10 TFs from other species involved in the regulation of anthocyanin biosynthesis. The protein sequences were aligned using ClustalW method and the tree was constructed by neighbor-joining method. Numbers close to the nodes indicates bootstrap values and scale shows 0.05 amino acid substitutions per site. The deduced amino acid sequences were retrieved from the www.phytozome.net and GeneBank database. Accessions are ABB83828 (AmVENOSA); ABX71488 (PpyMYB10); ABX71489 (PbMYB10); ABX71487 (PcMYB10); ABX71485 (MsMYB10); ABK58136 (MdMYB1); (MdMYB10a); ABX71486 (MgMYB10); ABX71483 (CoMYB10); ABX71484 (EjMYB10); ABX71490 (ParMYB10); ABX79946 (PsMYB10); ABX79948 (FvMYB10); ABX79950 (RiMYB10); AAQ55181 (SlANT1); ABB83826 (AmROSEA1); BAD18978 (VvMYBA2); XP_002262684 (VvMybA1); AT1G56650 (AtMYB75/PAP1); At1g66390 (AtMYB90/PAP2); At1g66370 (AtMYB113/PAP3); At1g66380 (AtMYB114/PAP4); ppa026640m (MYB10.1); ppa016711m (MYB10.2); ppa020385m (MYB10.3); PhAN2 (AAF66727) and ppa011751m (peach R2R2-MYB) used as an out group. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 5 (Tamura et al. 2011).

			Flower	Seed						Mesocarp					
GeneName	Description	val MAX	F_m	S1_m	S2_m	S3_m	S4_m	S5_m	S6_m	M1_m	M2_m	M3_m	M4_m	M5_m	M6_m
ppa026640m	MYB10.1	1680.11	100.00	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	3.82	0.06	35.84
ppa016711m	MYB10.2	6800.01	100.00	0.01	0.01	1.20	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	1.73
ppa020385m	MYB10.3	474.97	100.00	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	38.64
ppa024617m		31.84	3.14	3.14	3.14	100.00	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14
ppa010069m		428.46	100.00	0.23	0.23	4.10	16.65	0.23	4.48	0.23	0.23	0.23	0.23	0.23	0.23
ppa018744m		na	na	na	na	na	na	na	na	na	na	na	na	na	na

Figure 3.3 Microarray expression profiles of *MYB*s genes (using microarray data available in the laboratory of Prof. Livio Trainotti) in different stages of seed and mesocarp development in peach cv. 'Fantasia'. The highest value for each gene has been arbitrarily set to 100 (blue), and the others accordingly (0 = white). Val MAX: maximum expression value; F_m: relative expression value in flower; S1_m to S6_m and M1_m to M6_m represents relative expression values at various developmental stages of seed and mesocarp, respectively. M6_m is the relative expression value in the ripe mesocarp; 'na' indicates no expression values.

more details, while the remaining three (ppa024617m, ppa010069m and ppa018744m) were not considered further. According to the microarray expression profiles, the maximum expression was found in the flower for all three genes, while in the mesocarp only the expression of *MYB10.1* and *MYB10.3* was found to be relevant during ripening. Since the MYB TFs promoting anthocyanin biosynthesis have been shown to interact with bHLH TFs (Heim et al. 2003), the peach and *Arabidopsis* bHLHs were compared in order to find candidate bHLHs potentially participating in the regulation of anthocyanin synthesis. Only three peach proteins (bHLH3, ppa002884m; bHLH33, ppa002645m; and GL3, ppa002762m) clustered together with *Arabidopsis* bHLHs known to be involved in flavonoid synthesis regulation (Figure 3.4). Interestingly, three separate clades were formed and each one contained a single peach bHLH clustering with other known proteins from other species that are known to interact with flavonoid/anthocyanin promoting MYB TFs (Figure 3.5). Of the three identified peach proteins, bHLH3 and bHLH33 clustered with apple (MdbHLH3 and MdbHLH33), strawberry (FabHLH3 and FabHLH33), grape (VvMYC1 and VvMYCA1), and *Arabidopsis* (AtTT8 and AtMYC1) sequences, whereas the third member (i.e. GL3) formed a separate clade with strawberry FvGL3, grape VvGL3 and *Arabidopsis* AtEGL1 and AtGL3.

The expression profiles of these peach *bHLH* genes were analyzed using microarray expression data and only *bHLH3* was found to increase during ripening. Furthermore, *bHLH3*, with *bHLH33* and *GL3*, appeared to be expressed also during fruit growth, possibly participating in the control of the synthesis of other flavonoids (Figure 3.6).

Also a third regulator, a WD40-repeat protein, is known to be involved in the regulation of anthocyanin biosynthesis and, to this purpose, a peach WD40-repeat protein (ppa008187m) was identified through blasting with *Arabidopsis* (AtTTG1/At5G24520) and apple (MdTTG1/ADI58760) WD40-repeat proteins. According to a microarray expression profile, also *WD40* was highly expressed, although not exclusively, during ripening (Figure 3.7).

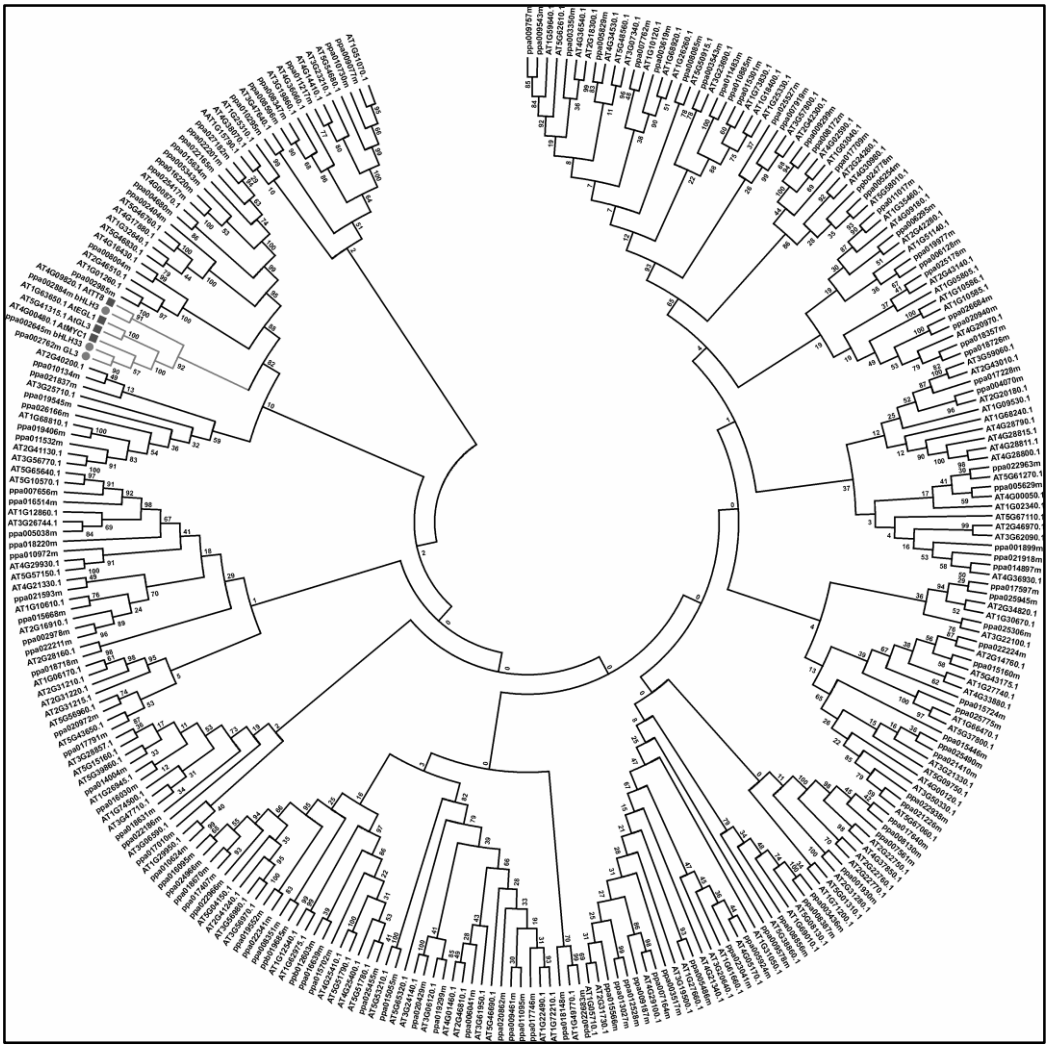

Figure 3.4 Phylogenetic tree showing relationships between *Arabidopsis* and peach bHLH TFs. Grey solid squares: *Arabidopsis* bHLH TFs which act as co-regulators of MYB TFs for the regulation of anthocyanin biosynthetic genes; grey solid circles: peach bHLH-like TFs belongs to the same clade with the *Arabidopsis* co-regulators. Phylogenetic and molecular evolutionary analyses were conducted using neighbor-joining method and 1000 bootstrap replicates using MEGA version 5 (Tamura et al. 2011).

Of all the peach *MYB10* and *bHLH* genes selected in this study, the CDSs were PCR amplified from either cDNA or gDNA, and their sequences were aligned with those of the reference peach genome sequence (Appendix I). The sequence analysis showed that only one nucleotide changed in *MYB10.1* (ppa026640m) at position 347 (A to G) thus leading to change a Lysine (K) for an Arginine (R) in the protein sequence. In *MYB10.2* there were four nucleotide changes from the reference sequence (ppa016711m) at positions 72 (C to T), 113 (A to G), 282 (C to A) and 351 (A to G) but only those at positions 113 and 351 led to an amino acid change (from tyrosine (Y) to cysteine (C) and isoleucine (I) to methionine

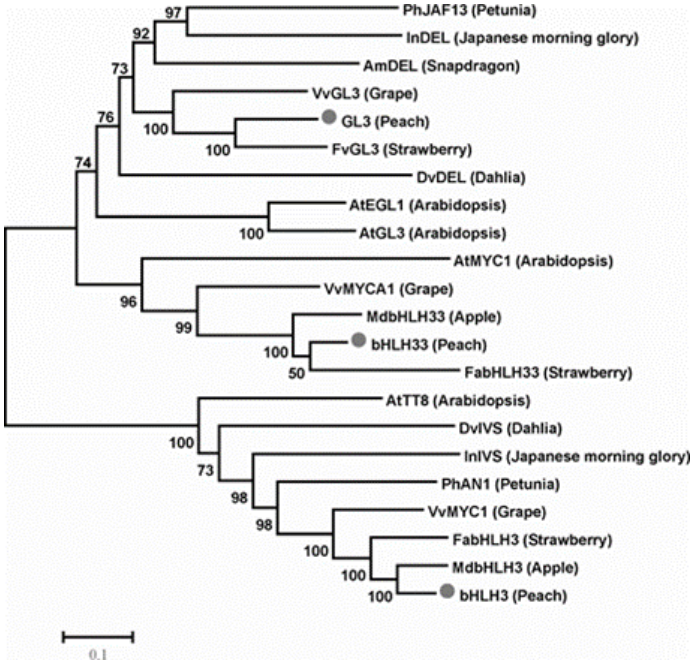

Figure 3.5 Phylogenetic tree for bHLH TFs known to be involved as a co-activator of R2R3-MYB TFs during anthocyanin biosynthesis. The deduced amino acid sequences were retrieved from www.phytozome.net and GeneBank databases. Accessions are ppa002884m (bHLH3); ppa002645m (bHLH33); ppa002762m (GL3); NP_176552 (AtEGL1); NP_680372 (AtGL3); CAC14865 (AtTT8), NP_191957 (AtMYC1); ADL36597 (MdbHLH3); ABB84474 (MdbHLH33); AAC39455 (PhJAF13); AAG25928 (PhAN1); AAA32663 (AmDEL); BAJ33515 (DvIVS); BAJ33516 (DvDEL); BAE94393 (InDEL); BAE94394 (InIVS); AFL02463 (FabHLH3); AFL02465 (FabHLH33); XP_004298802 (FvGL3); ACC6868 (VvMYC1); ABM92332 (VvMYCA1) and XP_002270239 (VvGL3). The protein sequences were aligned using ClustalW method and the tree was constructed by neighbor-joining method using MEGA version 5 (Tamura et al. 2011). Numbers close to the nodes indicates bootstrap values and scale shows 0.1 amino acid substitutions per site.

				Seed						Mesocarp					
GeneName	Description	val MAX	F_m	S1_m	S2_m	S3_m	S4_m	S5_m	S6_m	M1_m	M2_m	M3_m	M4_m	M5_m	M6_m
ppa002884m	bHLH3	5024.52	44.33	18.36	16.60	33.53	9.86	11.41	10.57	44.98	52.98	87.98	61.38	73.16	100.00
ppa002645m	bHLH33	2437.70	15.27	74.28	11.68	6.13	10.96	7.69	21.13	35.82	100.00	41.85	53.99	7.01	4.06
ppa002762m	GL3	1785.22	28.77	33.28	6.28	6.44	3.92	0.06	6.43	63.51	100.00	74.71	92.57	48.31	36.20

Figure 3.6 Microarray expression profiles of bHLH genes (using microarray data available in the laboratory of Prof. Livio Trainotti) in different stages of seed and mesocarp development in peach cv. 'Fantasia'. The highest value for each gene has been arbitrarily set to 100 (blue), and the others accordingly (0 = white). Val MAX: maximum expression value; F_m: relative expression value in flower; S1_m to S6_m and M1_m to M6_m represents relative expression values at various developmental stages of seed and mesocarp, respectively. M6_m is the relative expression value in the ripe mesocarp; 'na' indicates no expression values.

			Flower	Seed						Mesocarp					
GeneName	Description	val MAX	F_m	S1_m	S2_m	S3_m	S4_m	S5_m	S6_m	M1_m	M2_m	M3_m	M4_m	M5_m	M6_m
ppa008187m	PpWD40	3467.95	82.750822	58.15612	55.45376	62.187112	82.58263	67.44715	70.61454	80.46257	65.21541	95.93638	78.94481	88.04908	100

Figure 3.7 Microarray expression profile of *WD-40* (using microarray data available in the laboratory of Prof. Livio Trainotti) in different stages of seed and mesocarp development in peach cv. 'Fantasia'. The highest value for each gene has been arbitrarily set to 100 (blue), and the others accordingly (0 = white). Val MAX: maximum expression value; F_m: mean expression in peach flowers; S1_m to S6_m and M1_m to M6_m represents various developmental stages of seed and mesocarp, respectively. M6_m is the mesocarp at ripening stage.

(M), respectively). In *MYB10.3*, five nucleotides differed from the reference sequence (ppa020385m); changes were at positions 169 (G to T), 211 (G to A), 566 (A to G), 662 (A to G) and 679 (C to G) which change amino acids valine (V) to leucine (L), valine (V) to isoleucine (I), glutamic acid (E) to glycine (G), asparagine (N) to serine (S), leucine (L) to valine (V), respectively.

After cloning and sequencing of the *bHLH3*cDNA, it was clear that the length was 240 bp longer than expected. As in the reference genome sequence this part was inaccurate, the coding sequence (CDS) prediction was wrong, shorter than it should have been. The genomic region encompassing this 240 bp CDS fragment was further validated by amplifying it from both the cDNA and the genomic DNA (gDNA) of different peach cvs including 'Redhaven', 'Roza', 'Fantasia' and 'Stark Red Gold'. The cDNA templates always amplified (using 1.2_F: 5'-CGTCGTCGAACTAGGCACCA-3' and 1.5_R: 5'-GGGGCCGAAACCGCTGGAG-3' primers) a fragment with the same length (351 bp), while gDNA templates amplified a 457 bp (106 bp larger than expected) fragments for all the cvs analyzed (Figure 3.8a). Sequencing of the PCR product from the 'Stark Red Gold' genomic DNA clarified the nature of a 100 bp fragment, while a short stretch of 36 bp was absent. The new corrected genomic sequence of locus ppa002884m allowed to confirm that exon five is 240 bp longer than expected, as the first part of this exon could not be predicted with the inaccurate sequence. After the correction of the reference sequence, it was found that in the CDS of *bHLH3*, which codes for a protein that is 80 aa longer, only one nucleotide changed from the reference sequence (ppa002884m) at position 1124 (A to G) leading to change from aspartic acid (D) to glycine (G) (Figure 3.8b). There was only one nucleotide changed in *bHLH33* from the reference sequence (ppa002645m) at position 571 (A to G, leading from I to V). In the case of *GL3*, two nucleotides changed from the reference sequence (ppa002762m) at position 292 (C to T, leading from proline (P) to S) and 1086 (A to C, silent). The *GL3* had a 15 bp (5'-GGCGCCTCCATTCAG-3') insertion at position 478 leading to five (GASIQ) additional amino acids.

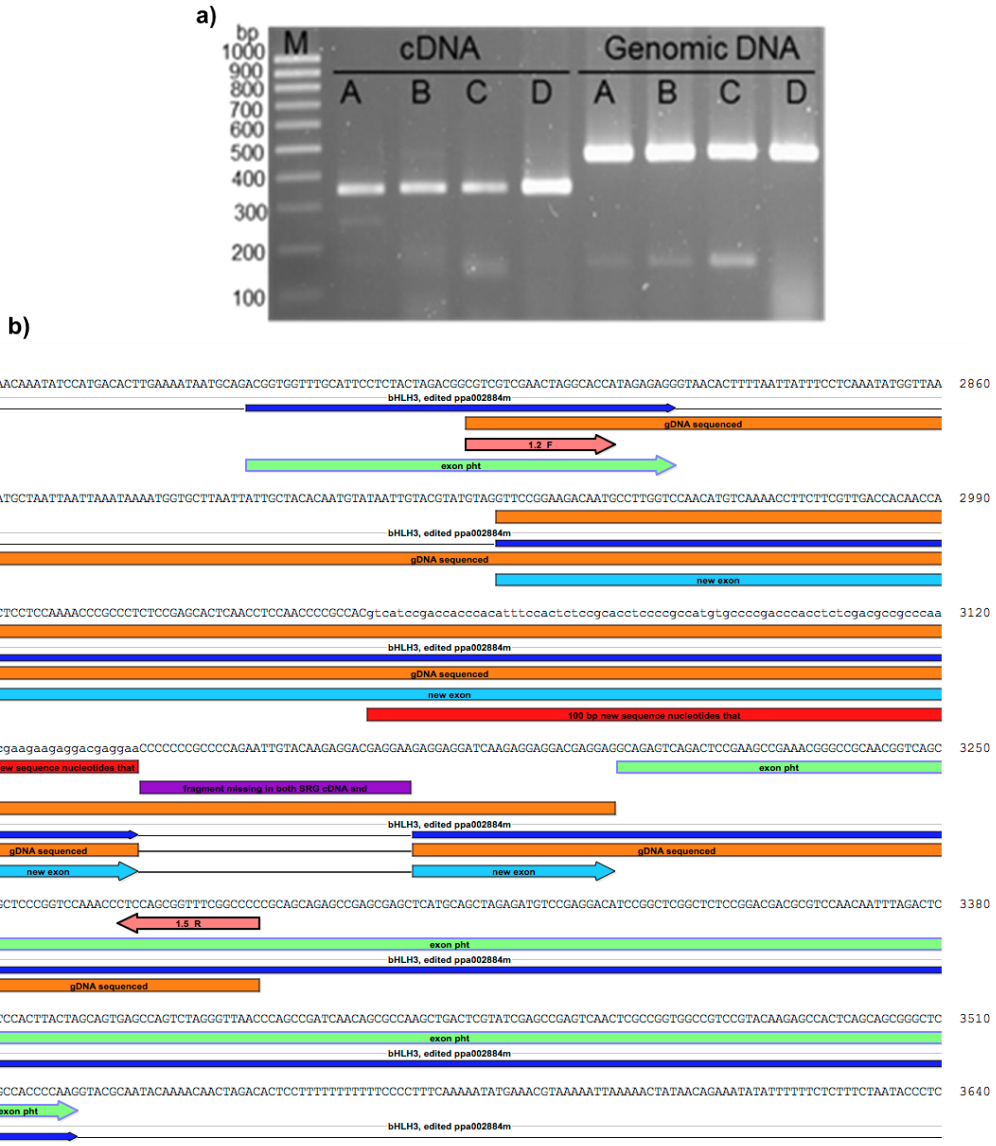

Figure 3.8 a) PCR amplification of the approximately sequenced fragment of bHLH3 from peach cvs using bHLH3.2_F and bHLH3.5_R primers. The letters A, B, C and D corresponding to peach cvs 'Redhaven', 'Roza', 'Fantasia' and 'Strak Red Gold', respectively while M indicates 100 bp DNA ladder. b) Analysis of the fragment of bHLH3 using gDNA of 'Stark Red Gold'. PCR primers are shown in the pink filled arrows, while the orange filled box represents the sequence of the PCR product. The green filled arrows indicate exons predicted in Phytozome, the light blue arrow the fragment of exon five that could be predicted and the dark blue arrow the correct exon predictions. The red box is the missing fragment in both the cDNA and gDNA of 'Stark Red Gold', while the purple one is a fragment present in the reference genome but absent in the sequenced cDNA and gDNA fragments (the removal of this fragment will join the two parts of exon five, here interrupted).

3.1.2 Accumulation of anthocyanins in peach at ripening

Total anthocyanins were measured spectrophotometrically in different parts of fruits of three peach cultivars ('Redhaven', 'Roza' and 'Fantasia') at ripening. These are melting flesh cultivars able to accumulate anthocyanin pigments in both

the peel (i.e. epicarp) and the mesocarp (Figure 3.9). In mesocarp the highest accumulation was found in the part around the stone (Figure 3.10). Among these cultivars, the highest quantity of anthocyanins was found in 'Fantasia' which accumulated more than twofold the amount of pigments present in 'Roza' and 'Redhaven' both in the peel and in the mesocarp around the stone.

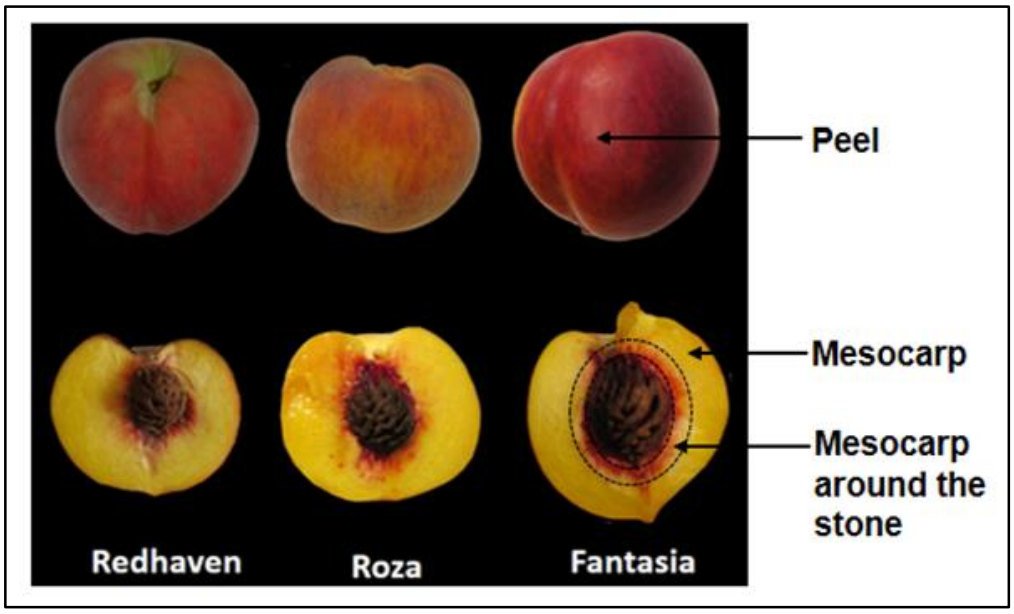

Figure 3.9 Peaches (S4 stage) of three cultivars showing the different parts from where sampling was done for quantification of total anthocyanin and gene expression analysis.

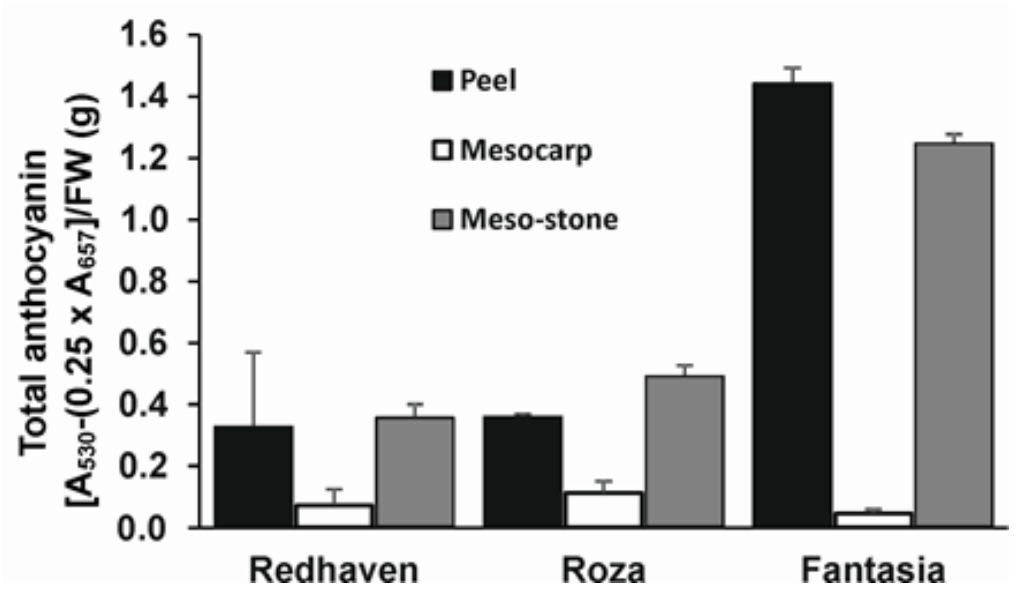

Figure 3.10 Total anthocyanin content in fruit parts of three different peach cultivars. Error bars are ± SD of the means of triplicates.

3.1.3 Expression profiles of *MYB10, bHLH* and anthocyanin biosynthetic genes in peach at ripening

The same fruit parts used to determine the anthocyanin amounts at ripening in the three cvs were also used to quantify, by means of quantitative real-time PCR (qRT-PCR), the expression of peach *MYB10* (*MYB10.1*, *MYB10.2* and *MYB10.3*), *bHLH* (*bHLH3*, *bHLH33* and *GL3*) and several structural (*CHS*, *CHI*, *F3H*, *DFR*, *UFGT*, *LAR* and *ANR*) genes. The expression of *MYB10.1* and *MYB10.3* was detectable in all tested samples and was highest in mesocarp around the stone, slightly lower in the peel and lowest in the mesocarp for all three cultivars (Figure 3.11). Albeit having a similar expression pattern that also matched the pattern of anthocyanin accumulation (Figure 3.10 and Figure 3.11), the transcript level of *MYB10.1* was over 10-fold that of *MYB10.3*. On the contrary, *MYB10.2* transcripts were hardly detectable in any sample but for the peel of 'Redhaven' and 'Fantasia' fruits where, nonetheless, the expression values were extremely low. Out of the three bHLH genes tested (Figure 3.10), only *bHLH3* was found to have an expression profile correlating with the anthocyanin accumulation, while

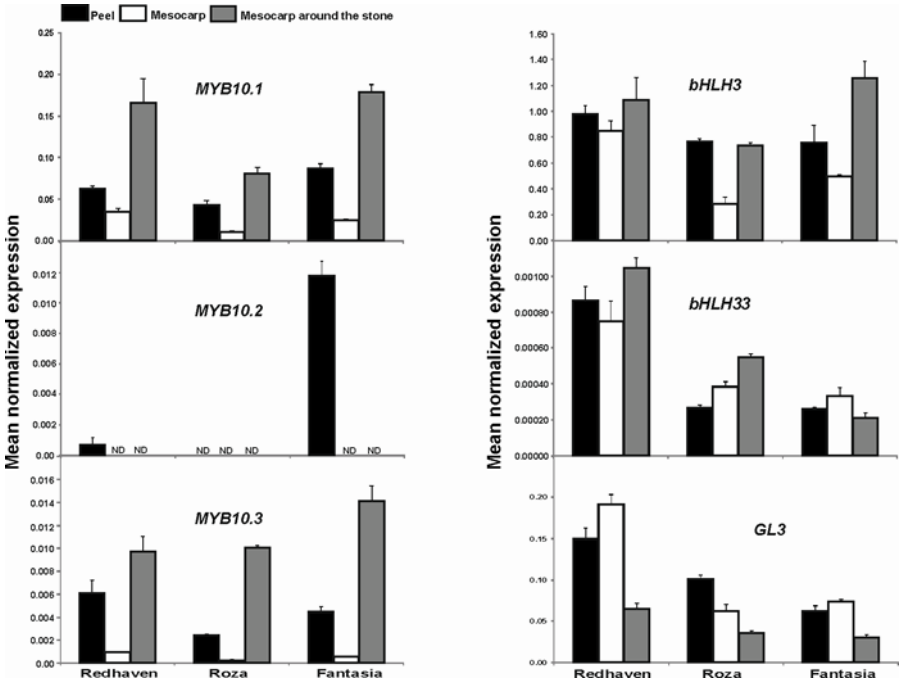

Figure 3.11 Expression profiles of *MYB10* (*MYB10.1*, *MYB10.2* and *MYB10.3*) and *bHLH* (*bHLH3*, *PpbHLH33* and *GL3*) genes in fruit parts of three peach cvs at ripening. Error bars are ± SE of the means of triplicates.

bHLH33 and *GL3*, besides being expressed at lower levels, showed profiles that differed according to the cv in which they were tested (Figure 3.11). The structural genes of the anthocyanin biosynthetic pathway have been well

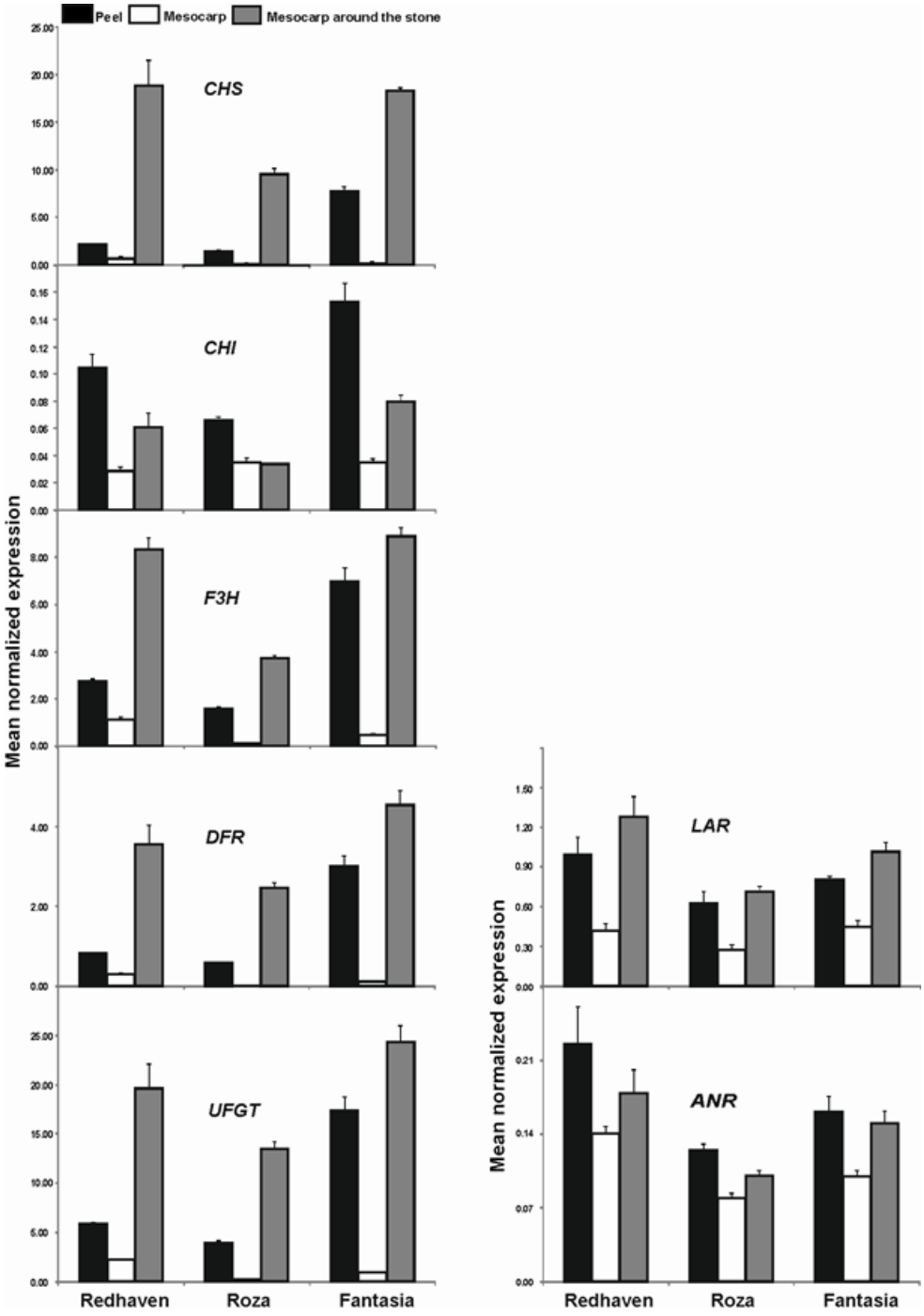

Figure 3.12 Expression profiles of peach anthocyanin and proanthocyanidin (PA) biosynthetic genes in three peach cvs during ripening. The analyzed anthocyanin and PA biosynthetic genes are as follows: *CHS*, Chalcone synthase; *CHI*, chalcone isomerase; *F3H*, flavanone-3β-hydroxylase; *DFR*, dihydroflavonol 4-reductase; *LAR*, leucoanthocyanidin reductase; *ANR*, anthocyanidin reductase; *UFGT*, UDP-glucose:flavonoid-3-*O*-glucosyltransferase. Error bars shown are ± SE from the means of triplicates.

characterized in several plants including maize, *Arabidopsis*, *Petunia*, grape, apple and pear (Kubasek et al. 1992, Lancaster and Dougall 1992, Sparvoli et al. 1994, Holton and Cornish 1995, Honda et al. 2002, Takos et al. 2006, Li et al. 2012). Besides the availability of the peach genome sequence (Verde et al, 2013), a recent study (Ravaglia et al. 2013) was the guide for selecting the peach structural genes to be used in this work. In particular, the expression profiles of five anthocyanin biosynthetic genes (*CHS*, *CHI*, *F3H*, *DFR* and *UFGT*) and two proanthocyanidin (PA) biosynthetic genes (*LAR* and *ANR*) at fruit ripening were tested in the same samples used above (Figure 3.12). In agreement with the pattern of anthocyanin accumulation in the different fruit parts during peach ripening, *CHS*, *F3H*, *DFR* and *UFGT*, but also *LAR* and *NAR*, were expressed at higher levels in peel and mesocarp around the stone and at lower, sometimes almost at background levels, in the mesocarp proper of the three cultivars (Figure 3.12). The transcript levels of genes encoding enzymes situated at the top (CHS) and end (UFGT), respectively, of the anthocyanin biosynthetic pathway were much higher than those of other biosynthetic genes. Correlation tests (Pearson correlation coefficient, r) between gene expression and anthocyanin content were found for *F3H*, *DFR* and *UFGT* (Table 3.1).

Table 3.1 Pearson's correlation (r) between anthocyanin content [(A_{530}-(0.25 x A_{657}))/ FW (g)] and regulatory as well as anthocyanin and PAs biosynthetic genes.

Gene name	**Anthocyanin content**	
	Correlation (r)	**p-value**
MYB10.1	0.610**	0.041
MYB10.3	0.570	0.055
bHLH3	0.390	0.150
bHLH33	-0.290	0.225
GL3	-0.280	0.233
CHS	0.570	0.055
CHI	0.770**	0.008
F3H	0.780**	0.007
DFR	0.780**	0.007
UFGT	0.790**	0.006
LAR	0.460	0.106
ANR	0.280	0.233

[*: $p \leq 0.05$; **: $p \leq 0.01$ and ***: $p \leq 0.001$]

Table 3.2 Pearson's correlation (r) of transcript levels of *MYB10s* (*MYB10.1* and *MYB10.3*) with anthocyanins and PAs biosynthetic genes during peach ripening.

Gene name		*MYB10.1*		*MYB10.3*	
		Correlation (r)	**p-value**	**Correlation (r)**	**p-value**
Anthocyanins	*CHS*	0.980***	0.000	0.900****	0.000
	CHI	0.320	0.201	0.230	0.276
	F3H	0.950***	0.000	0.780**	0.007
	DFR	0.950***	0.000	0.900***	0.000
	FGT	0.950***	0.000	0.890**	0.001
PAs	*LAR*	0.890**	0.001	0.800**	0.005
	ANR	0.460	0.106	0.370	0.164

[*: $p \leq 0.05$; **: $p \leq 0.01$ and ***: $p \leq 0.001$]

Table 3.3 Pearson's correlation (r) of transcript levels of *bHLHs* (*bHLH3*, *bHLH33* and *GL3*) with anthocyanins and PAs biosynthetic genes during peach ripening.

Genes name		*bHLH3*		*bHLH33*		*GL3*	
		Correlation (r)	**p-value**	**Correlation (r)**	**p-value**	**Correlation (r)**	**p-value**
Anthocyanins	*CHS*	0.980***	0.000	0.170	0.331	-0.570	0.055
	CHI	0.230	0.276	-0.140	0.360	-0.010	0.490
	F3H	0.890**	0.001	0.100	0.399	-0.430	0.124
	DFR	0.750**	0.010	0.190	0.312	-0.310	0.208
	UFGT	0.870**	0.001	-0.040	0.459	-0.520	0.076
PAs	*LAR*	0.700**	0.018	0.520	0.076	-0.020	0.480
	ANR	0.270	0.241	0.440	0.118	0.460	0.106

[*: $p \leq 0.05$; **: $p \leq 0.01$ and ***: $p \leq 0.001$]

3.1.5 Functional analysis of peach *MYB10* and *bHLH* genes by transient over-expression assays in heterologous and homologous systems

To gain more insights on the regulatory properties of peach *MYB10* and *bHLH* genes, their transactivating capabilities were tested in fully expanded tobacco leaves and ripening peach fruits by means of *Agrobacterium* mediated transient over-expression assays. Tobacco leaves were infiltrated with the nine possible combinations arising by mixing each one of the three *MYB10* (*MYB10.1*, *MYB10.2*

and *MYB10.3*) genes with each of the three *bHLH* (*bHLH3*, *bHLH33* and *GL3*) genes. When co-infiltrated with *bHLH3*, the *MYB10.1* gene and, to a lesser extent *MYB10.3*, induced anthocyanin formation 5 days after transformation, and this was visible as purple patches in the injected areas, (Figure 3.13, panel-a). The total anthocyanin content in the infiltrated areas of tobacco leaves was highest for the *MYB10.1*/*bHLH3* combination, being twice than that of the *MYB10.3*/*bHLH3* combination (Figure 3.14). With all the other combinations only basal levels of anthocyanin formation were detected.

To exclude the possibility that disparities in transformation efficiencies were the causes of the observed differences in pigment accumulation and gene transactivation, the expression of peach *MYB10s* and *bHLHs* were also carefully

Figure 3.13 Anthocyanin accumulation in agro-infiltrated sectors of tobacco leaves (panel-a) and peach fruits (panel-b) 5 days after transformation.

	MYB10.1	MYB10.2	MYB10.3
bHLH3	2.34	0.17	0.92
bHLH33	0.33	0.26	0.05
GL3	0.39	0.64	0.57
35S::Lhg4	0.81		

Figure 3.14 Heat map showing the total anthocyanin content [(A_{530}-(0.25 x A_{657}))/ FW (g)] in the agro-infiltrated sectors of tobacco leaves injected with different combinations of MYB (MYB10.1, MYB10.2 and MYB10.3) and bHLH (bHLH3, bHLH33 and GL3) TFs. The values are the means of three infiltrated leaves. The highest expression value has been arbitrarily set to 100 (red), and the others accordingly (0 = white). 35S::Lhg4 is the negative control (sector injected with empty plasmid).

quantified by qRT-PCR in the injected sectors of tobacco leaves and normalized on the expression of *GUS*, a transgene under the same regulatory system controlling also the expression of the peach MYB and bHLH TFs. As the expression levels of peach *MYB10s* and *bHLHs* were similar in tobacco leaves injected with different combinations of TFs, the possibility that the observed differences on the activation of *NtCHS*, *NtDFR* and *NtUFGT* were due to disparities in transgene expression was ruled out (Figure 3.15).

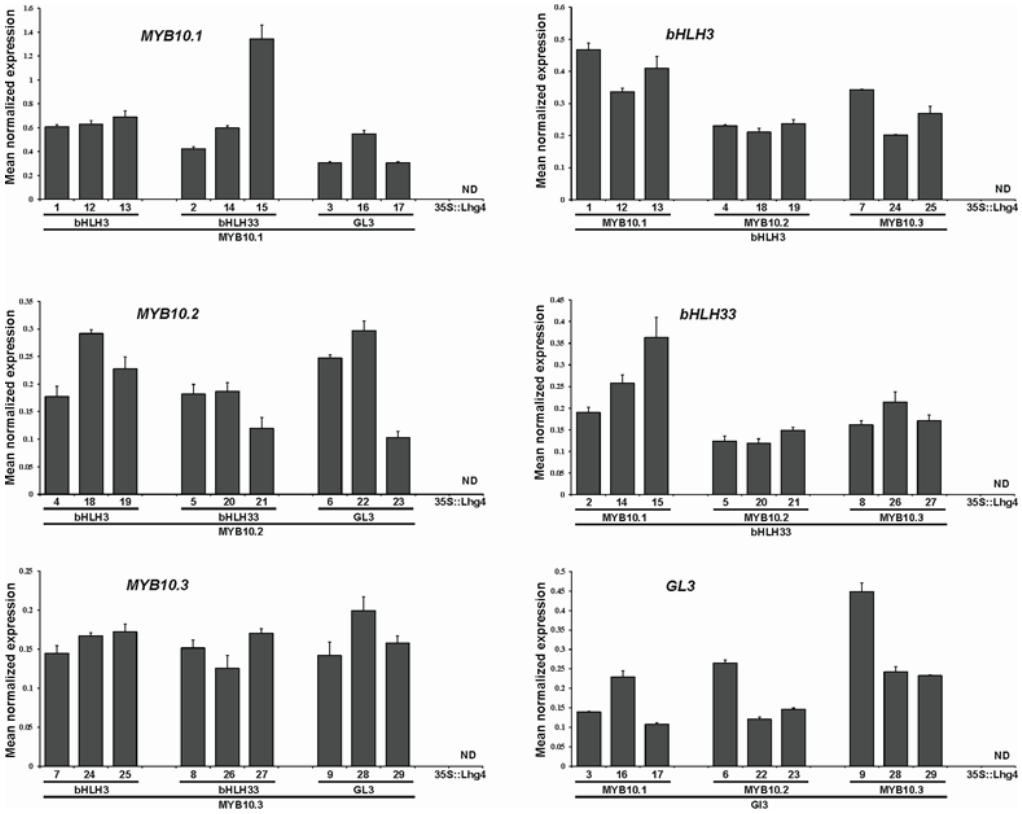

Figure 3.15 Transcript levels of peach *MYB10s* and *bHLHs* detected by qRT-PCR from agro-infiltrated tobacco leaves. The expression values of the target genes were normalized by the expression values of the *GUS* gene, used as internal standard. Error bars shown are ± SE from the means of triplicates.

The expression profiles of *NtCHS*, *NtDFR*, *NtUFGT*, *NtFLS* and *NtLAR* were determined in the infiltrated sectors of tobacco leaves by qRT-PCR. At the molecular level, a strong transactivation of *NtCHS* was detected only with *MYB10.1/bHLH3* and *MYB10.3/bHLH3* combinations. The transactivation of *NtDFR* was strong with the over-expression of *MYB10.1*/*bHLH3* and still significant, albeit weaker, with the *MYB10.3/bHLH3* combination (Figure 3.16).

NtUFGT, besides being better induced by the former and to a lower extent by the latter combination, was transactivated also by the *MYB10.3/bHLH33* combination (Figure 3.16). Transcript levels of *NtFLS* and *NtLAR* were low with any of the tested combinations (Figure 3.16), indicating their scarce ability to transactivate both the flavonol and PA pathways.

		MYB10.1	MYB10.2	MYB10.3
NtCHS	bHLH3	0.663	0.007	0.266
		0.338	0.072	0.131
		0.751	0.054	0.346
	bHLH33	0.013	0.017	0.004
		0.012	0.008	0.002
		0.032	0.002	0.000
	GL3	0.005	0.002	0.001
		0.022	0.007	0.002
		0.017	0.000	0.004
NtDFR	bHLH3	1.296	0.005	0.300
		0.660	0.056	0.113
		1.535	0.066	0.300
	bHLH33	0.095	0.003	0.017
		0.086	0.007	0.008
		0.294	0.004	0.006
	GL3	0.003	0.009	0.004
		0.086	0.008	0.009
		0.019	0.001	0.014
NtUFGT	bHLH3	0.253	0.002	0.072
		0.151	0.016	0.058
		0.230	0.027	0.083
	bHLH33	0.007	0.004	0.111
		0.007	0.007	0.125
		0.039	0.001	0.095
	GL3			0.006
		0.007	0.001	0.008
		0.006		0.019
NtFLS	bHLH3	0.008	0.001	0.006
		0.003	0.049	0.003
		0.008	0.011	0.000
	bHLH33	0.011	0.004	0.002
		0.021	0.009	0.001
		0.041	0.002	0.000
	GL3	0.002	0.008	0.000
		0.015	0.003	0.002
		0.013	0.001	0.002
NtLAR	bHLH3	0.001	0.000	0.001
		0.000	0.000	0.001
		0.001	0.000	0.002
	bHLH33			0.000
				0.000
				0.000
	GL3		0.000	
		0.000	0.000	
			0.000	

Figure 3.16 Heat map showing the expression of anthocyanin (*NtCHS*, *NtDFR*, *NtUFGT)*, flavonol (*NtFLS)* and PA (*NtLAR)* biosynthetic genes by qRT-PCR in agro-infiltrated tobacco leaves. The expression values of the target genes were normalized by the expression values of the *GUS* gene, used as internal standard. Each value represent the means of triplicates. The highest expression value has been arbitrarily set to 100 (blue), and the others accordingly (0 = white).

Thus, the heterologous tobacco transient over-expression assays showed that *MYB10.1* and *MYB10.3* were able to induce purple pigmentation when *bHLH3* was present as a partner. Since the experiments with tobacco had shown that the best MYB10s' partner was bHLH3, the latter gene was used in all the possible combinations with the three *MYB10* genes for transient over-expression assays in peach fruits. In particular, fruits of the 'Springcrest' cultivar, harvested at the transition between S3 to S4 stages of mesocarp development (i.e. at the pre-climacteric stage), were used for these transient over-expression assays. The agroinjection is technically more difficult in peach fruits than in tobacco leaves, thus leading to some variability among the samples injected with the same gene combination.

Anyway, even if it was not possible to determine which was the best inducer of the anthocyanin synthesis among the three *MYB10* genes (Figure 3.16), the agroinjections showed that all three *MYB10* genes were capable to induce pigment accumulation in combination with *bHLH3* (Figure 3.12, panel-b).

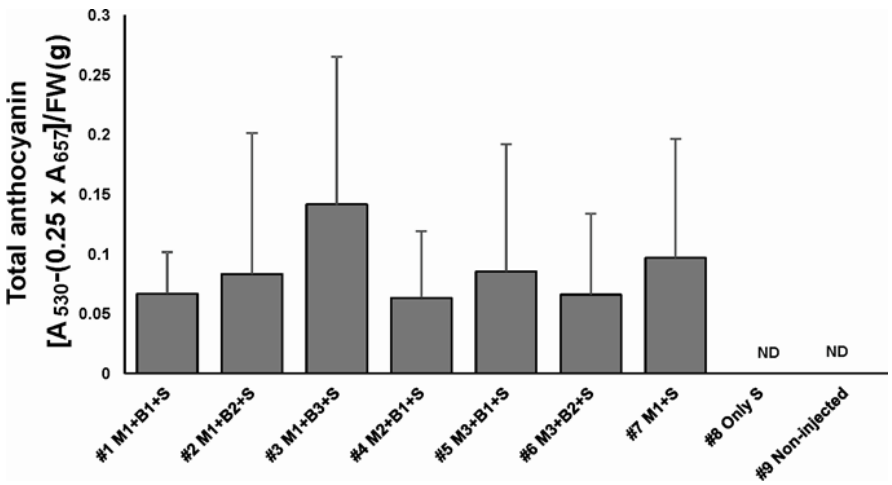

Figure 3.17 Total anthocyanin content in agro-infiltrated peach fruit mesocarp. Total anthocyanins were quantified spectrophotometrically as [A_{530}-(0.25 x A_{657}]/FW (g) and normalized by the GUS activity. The M1, M2, M3, B1, B2, B3 and S indicates *MYB10.1*, *MYB10.2*, *MYB10.3*, *bHLH3*, *bHLH33*, *GL3* and *35S::Lhg4*, respectively. Error bars ± SD of the means of three fruits.

3.2 Functional characterization of a peach R2R3-MYB TF gene by its over-expression in tobacco

3.2.1 Identification and cloning of the peach *MYB10.1* gene

The MYB10-like sequences have been described in section 3.1.1 of this chapter and *MYB10.1* was found as one of the best regulators of anthocyanin biosynthesis in peach during ripening. The *MYB10.1* cDNA was isolated from peach fruit cv 'Stark Red Gold' cDNA library. Analysis of the sequence showed that it encodes an R2R3-MYB TF with a CDS of 720 bp and a deduced peptide of 239 amino acid (aa) residues. It has nucleotide sequence homology with anthocyanin promoting *Arabidopsis* MYB TFs like AtPAP1, AtPAP2, AtPAP3 and AtPAP4 *(*58.6% and 63.3%, 52.6% and 57.7% similarity, respectively). Actually, a relatively high similarity (79.2%) was found with MdMYB10 TF (Table 3.4).

Table 3.4 Comparison of peach MYB10.1 (an R2R3-MYB) TF with anthocyanin promoting *Arabidopsis* and apple R2R3-MYB TFs.

Name of the TFs	GeneBank/TAIR/ Phytozome accessions	gDNA (bp)	CDS (bp)	Protein (aa)	MYB10.1 Similarity (%)*
MYB10.1	ppa026640m	1858	720	239	100
AtPAP1/AtMYB75	AT1G56650/AAG42001	1543	747	248	58.6
AtPAP2/AtMYB90	AT1G66390/NP_176813	1585	750	249	63.0
AtPAP3/AtMYB113	AT1G66370/NP_176811	971	741	246	52.6
AtPAP4/AtMYB114	AT1G66380/NP_176812	1078	420	139	57.7
MdMYB10	EU518249.2	4050	732	243	79.2

* CDS for each TFs were compared with MYB10.1 using "Lasergene" software package (DNASTAR).

3.2.2 Generation of transgenic tobacco plants over-expressing the peach *MYB10.1* gene

The functions of the peach *MYB10.1* gene were analyzed by over-expressing it in the tobacco heterologous system (see materials and methods). Several independent lines over-expressing the peach *MYB10.1* gene were obtained from the transformation events and the presence of the transgene was confirmed by PCR. For the phenotypic and molecular characterization, mainly T_0 plants were used but,

in some cases, also T_1 and F_1 (crossing between transgenic plants for *35S::Lhg4* and promoter-less *MYB10.1* plants) tobacco plants were used for further analysis. The transgenic plants were compared with wild type tobacco plants, also propagated by tissue culture, at the same developmental stages.

3.2.3 Phenotypes of transgenic tobacco plants over-expressing the peach *MYB10.1* gene

Transgenic plants showing strong defects in vegetative and reproductive characteristics are collectively described as "Type-I", whereas "Type-II" includes independent clones showing only minor changes in vegetative and reproductive development when compared with wild type (WT) tobacco plants (Figure 3.18). Six independent transgenic lines (named 1.2, 5.2, 9.2, 17.2, 19.2, and 20.2) could be assigned to type-I and seven (2.2, 8.2, 10.2, 16.2 18.2, 22.2, and 24.2) to type-II. Two independent lines (type-I: 1.2 and 17.2; type-II: 2.2 and 10.2) for each category of phenotype were analyzed further. Since the transgenic plants also contained a GUS reporter gene under the same transcriptional regulatory system (see materials and methods), GUS activity was assayed and result showed that at least 4-folds higher in type-I than in type-II transgenic plants, while no activity was detected in the WT (Figure 3.19). Thus, the highest GUS activity was detected in clones showing the strong phenotypes (type-I).

At the vegetative level, type-I transgenic plants showed reduced plant height (Figure 3.18) and leaf size; furthermore, also leaf shape was altered compared to type-II and WT tobacco plants (Figure 3.20a-e). An *ESEM* (*Environmental Scanning Electron Microscope*) analysis of leaf epidermis evidenced that the pavement cells of both the type-I and the type-II transgenic lines had significant reduction in their lengths compared to WT ones (Figure 3.20f-j and Figure 3.21a). Moreover, the breadth of the pavement cells was significantly increased only in type-I transgenic lines while there were not significant differences between type-II and WT (Figure 3.21b). In addition, the number of pavement cells per unit of area was also significantly reduced in type-I transgenic lines while in type-II lines it was similar to WT (Figure 3.21c).

Type-I transgenic plants showed abnormalities in floral development while type-II transgenic plant had flowers similar to WT (Figure 3.22). Type-I transgenic flowers had a huge reduction in the length of calyx, corolla, androecium and gynoecium compared to type-II and WT (Figure 3.23a-b and Table 3.5). The filament length of the anthers in type-I flowers were drastically reduced so that anthers did not reach the stigmas (Figure 3.24) thus preventing auto-pollination. Moreover, there was not nectary formation (Figure 3.23c). Whereas only petals were pigmented in type-II and WT flowers, in type-I transgenic flowers besides the strong pigmentation was observed in petals, some purple pigmentation was observed also in anthers and ovaries (Figure 3.23b-c). Moreover, in type-I transgenic plants usually the anthers did not dehisce; nonetheless, sometimes, at an extremely late stage when flowers were going to drop off, it could be observed that few anthers released some pollen at their tips (Figure 3.25a). At that stage stigmas had completely lost their ability to allow germination of pollen grains and thus all the flowers were fated to drop so that only remnants of inflorescences without capsules remained on the plant (Figure 3.25b). As a consequence of this altered development no seeds were produced by type-I transgenic plants. On the contrary, an increase in petal pigmentation and purple coats in developing seeds were the only phenotypic differences observed in type-II transgenic plants compared to WT (Figure 3.23d-e).

The total anthocyanin content was measured in petals, stamens, pistils and developing seeds of transgenic and WT tobacco flowers at anthesis (Figure 3.26). Leaves and stems were not pigmented in all transgenic T_0 plants during vegetative growth so they were not considered for this analysis. Type-I transgenic plants displayed enhanced anthocyanins accumulation in petals, stamens and pistils compared to type-II transgenic and WT flowers. The maximum accumulation of anthocyanins was detected in developing seeds of type-II transgenic plants while pigments were barely detectable in WT seeds. These results suggests that in a heterologous system peach *MYB10.1* is capable to induce anthocyanin pigmentation only in the reproductive tissues.

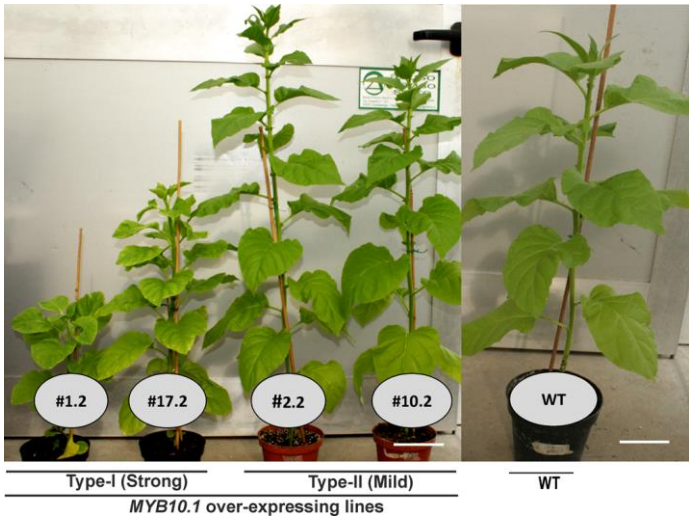

Figure 3.18 Phenotypes of peach *MYB10.1* over-expressing transgenic tobacco lines. Type-I, strong phenotype (1.2 and 17.2), Type-II, mild phenotype (2.2 and 10.2) and WT, wild type.

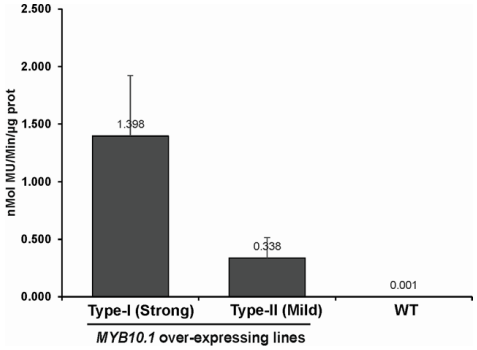

Figure 3.19 Enzymatic GUS activity in *MYB10.1* over-expressing transgenic tobacco lines and WT. Error bars are ± SD of the means of three independent plants.

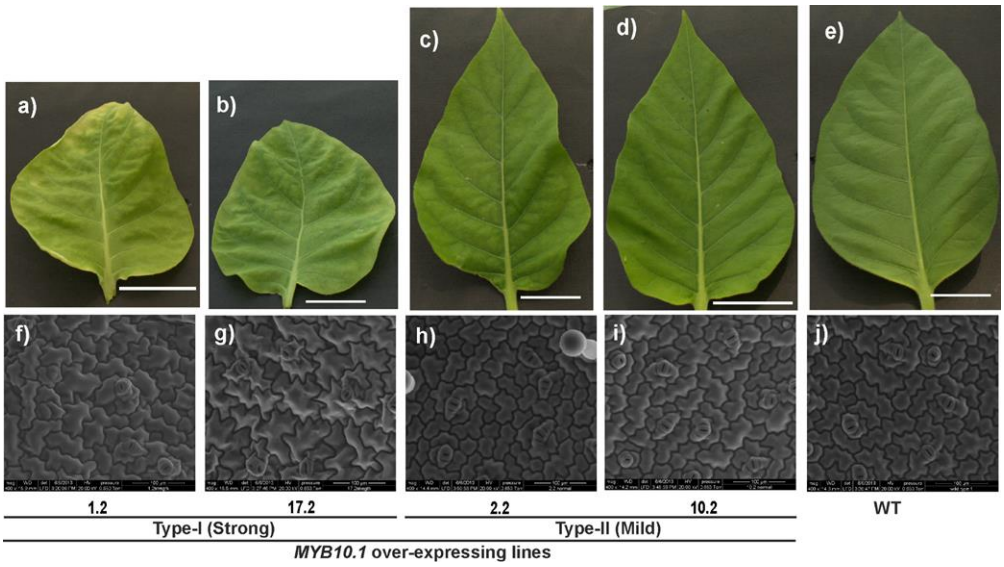

Figure 3.20 Phenotypic analysis of the upper leaf surface of the transgenic tobacco lines compare to WT. Digital photographs (a-e) of leaves with the same age and internode for both transgenic and WT plants. The images (f-j) of the upper surfaces of leaves were taken by means of an *Environmental Scanning Electron Microscope (ESEM*).

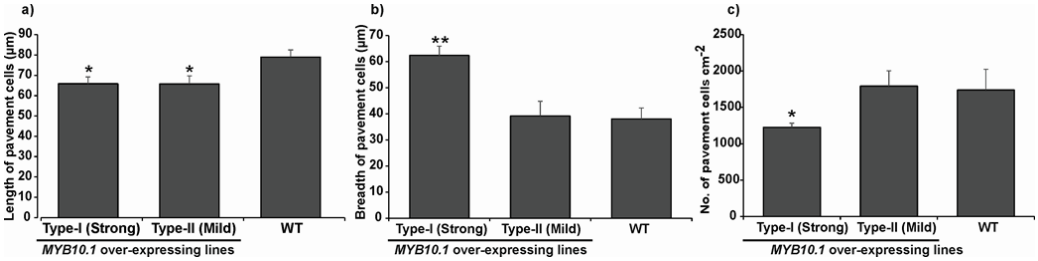

Figure 3.21 Analysis of the leaf epidermal (adaxial surface) cells of transgenic tobacco lines. Length of the pavement cells (a), breadth of the pavement cells (b) and number of the pavement cells cm^{-2} (c) of the tobacco leaf adaxial epidermis. Length and breadth were measured using ImageJ software program. At least 15 pavement cells of each independent transgenic line were used for length and breadth measurements. Cell density was calculated counting all the cells present in single photographs. All the photographs used in the measurement were taken by *ESEM*. Error bars are ± SD of the means of two independent lines. Single and double asterisks indicate $p \leq 0.05$ and **: $p \leq 0.01$, respectively.

Figure 3.22 Alteration of floral development in *MYB10.1* over-expressing transgenic tobacco (Type-I) compared to WT (stages were defined according to Koltunow et al. 1990).

Figure 3.23 Floral phenotypes of *MYB10.1* over-expressing transgenic tobacco lines compare to WT. Panel-a, complete flowers; panel-b, open flowers showing the anther-stigma distance; panel-c, nectary region of the ovary, panel-d, green capsules and panel-e, open capsules showing seed coat pigmentation. N/A indicates the absence of capsules and seeds in type-I transgenic plants.

Table 3.5 Length of different floral parts of *MYB10.1* over-expressing transgenic tobacco lines compared to WT. In brackets are indicated the percentage variations with the WT.

Phenotype	Genotype	Sepal (cm)	Petal (cm)	Androecium (cm)	Gynoecium (cm)
		Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
Strong	1.2	1.13 ± 0.06 (70%)	2.27 ± 0.12 (46%)	0.99 ± 0.13 (23%)	1.48 ± 0.03 (36%)
	17.2	1.2 ± 0.08 (74%)	2.30 ± 0.18 (47%)	1.04 ± 0.22 (24%)	1.70 ± 0.10 (100%)
Mild	2.2	1.68 ± 0.05 (103%)	4.55 ± 0.06 (92%)	4.13 ± 0.10 (95%)	4.10 ± 0.00 (100%)
	10.2	1.67 ± 0.04 (103%)	4.60 ± 0.07 (93%)	4.19 ± 0.05 (97%)	4.10 ± 0.08 (100%)
Control	WT	1.62 ± 0.13 (100%)	4.94 ± 0.09 (100%)	4.34 ± 0.05 (100%)	4.1 ± 0.08 (100%)

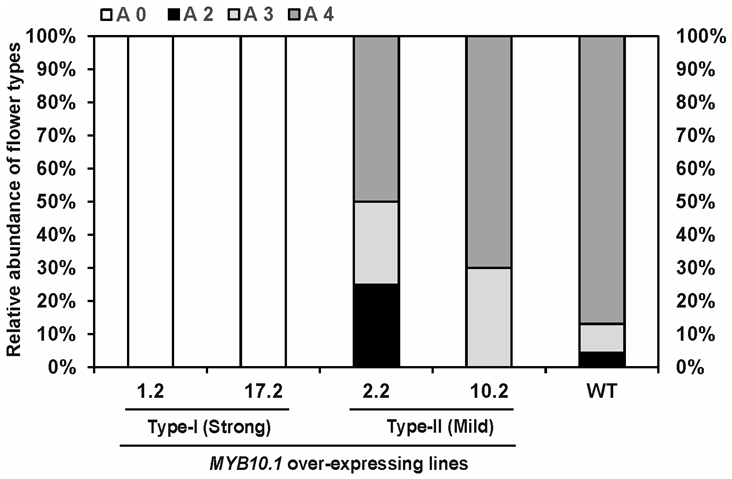

Figure 3.24 The anther filament elongation pattern of *MYB10.1* over-expressing transgenic tobacco flowers. A0, A2, A3 and A4 indicate flowers with 0, 2, 3 and 4 stamens touching the stigmas, respectively.

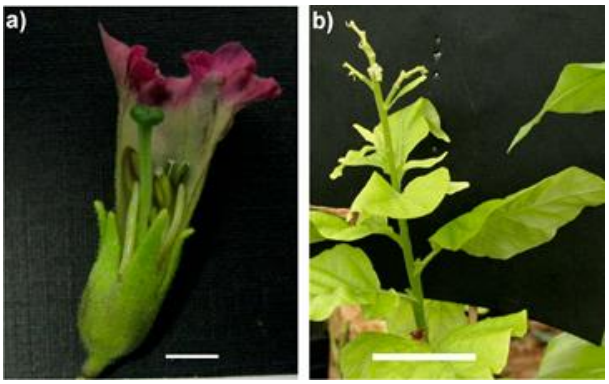

Figure 3.25 Shattering of anthers at extremely late stages of type-I flower development when they are going to abscise (a); only petioles of the floral inflorescence without capsules remain on type-I transgenic plants (b).

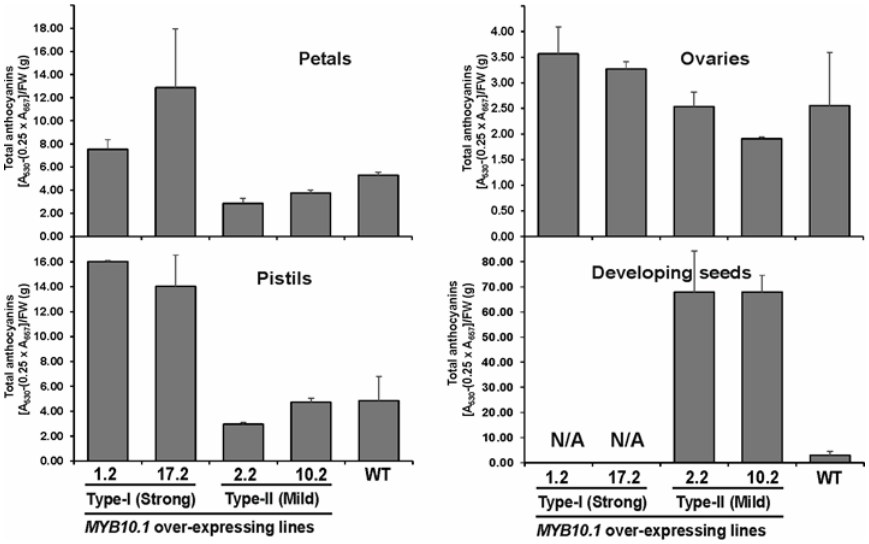

Figure 3.26 Total anthocyanin content in reproductive parts of transgenic and WT tobacco plants at anthesis. The quantification was done spectrophotometrically and presented as [A_{530}-(0.25 x A_{657}]/FW (g). N/A, seeds were absent in type-I transgenic plants. Error bars are ± SD of means.

3.2.4 Alteration of gene expression in transgenic tobacco flowers

The expression of genes related to anthocyanin biosynthesis and floral development was analyzed by RT-PCR in both the transgenic and the WT tobacco flowers at anthesis. As shown in Figure 3.27, transcripts of the peach *MYB10.1* gene were at high levels in type-I transgenic flowers, while were relatively low in type-II transgenic flowers. As expected, *MYB10.1* transcripts were not detected in WT flowers. The expression level of the endogenous *NtAN2* (an R2R3-MYB related to the regulation of anthocyanin biosynthesis in tobacco flower and the orthologue of *MYB10.1*) was low in type-I, moderate in type-II and high in WT tobacco flowers. On the other hand, the expression pattern of endogenous *NtAN1b* (encoding bHLH TF) was similar to WT in both types of transgenic flowers.

The transcript abundance of anthocyanin biosynthetic genes was found to be from slightly (*chalcone isomerase* (*NtCHI*), *flavanone-3-hydroxylase* (*NtF3H*), *dihydroflavonol 4-reductase* (*NtDFR*) *anthocyanidin synthase* (*NtANS*) and *UDP-glucose:flavonoid-3-O-glucosyltransferase* (*NtUFGT*)) to strongly (*phenylalanine ammonia-lyase* (*NtPAL*), *chalcone synthase* (*NtCHS*)) increased in transgenic plants of both types, with and induction that seems even stronger in type-II than in type-I clones (e.g. for *CHI*, *F3H* and *DFR* genes). The enhanced expression levels

of the abovementioned biosynthetic genes correlate with the purple seed coat produced by type-II transgenic tobacco plants. A similar expression pattern was also found for *leucoanthocyanidin reductase* (*NtLAR*), while *anthocyanidin reductase 1* (*NtANR1*) expression was low and not so different from that in WT, meaning that *MYB10.1* might also regulate proanthocyanidin biosynthetic genes in tobacco flowers.

Since type-I transgenic tobacco plants had shown defects in their reproductive organs, other genes encoding MYB TFs related to floral development were analyzed in this study. In *Arabidopsis* three MYBs (AtMYB21, AtMYB24 and AtMYB57), are predominantly expressed in flowers (Li et al. 2006; Cheng et al. 2009). As regards tobacco, NtMYB305, besides being the orthologous MYB of the abovementioned *Arabidopsis* genes, controls nectarin and flavonoid biosynthetic gene expression in flowers (Liu et al. 2009). Considering these MYBs, a phylogenetic analysis was carried out and NtMYB305 was confirmed to be orthologous to *Arabidopsis* AtMYB21, AtMYB24 and AtMYB57, while only a single peach gene (ppa011751, named MYB24 following the *Arabidopsis* nomenclature) was found in this clade (Figure 3.28). The expression of *NtMYB305*, known to be at maximum between stages 9 and 10 (Liu et al., 2009) was reduced in type-I and similar to WT in type-II transgenic flowers (Figure 3.27). As the expression of *NtNEC1*, a gene encoding the major nectarin protein, is controlled by *NtMYB305*, it was analyzed in *MYB10.1* over-expressing plants. Not surprisingly, *NtNEC1* was not expressed in type-I transgenic flowers, which also had extremely low amounts of *NtMYB305* transcripts, whereas it appeared expressed at similar levels in both type-II and WT flowers. These profiles correlate with the nectary glandless phenotype of type-I transgenic tobacco flowers (Figure 3.23c).

Jasmonic acid (JA) is required for normal androecium development in *Arabidopsis* acting also trough the induction of MYB genes (Mandaokar et al., 2006). As type-I plants had phenotypes resembling *Arabidopsis* mutants defective in JA synthesis or signaling (Mandaokar and Browse, 2009), *NtAOS* (*allene oxide synthase*, for JA biosynthesis) and *NtJAZd* (encodes a jasmonate ZIM-domain protein, for JA signaling) were tested for their expression. Both genes were up-regulated in type-I transgenic flowers but had levels similar to WT in type-II clones.

The type-I transgenic T0 plants were unable to set seeds either by self-pollination, as occurs in WT flowers, or by manual pollination, even with WT pollen. Ethylene plays a role in megasporogenesis (De Martinis and Mariani, 1999) and

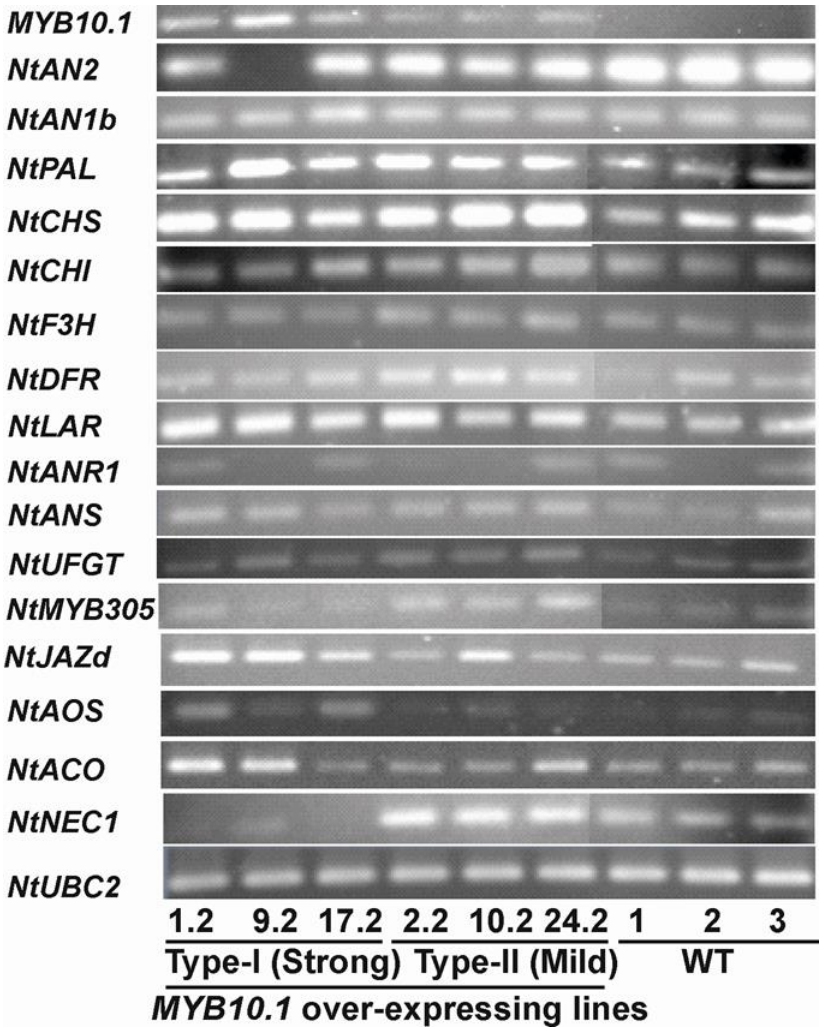

Figure 3.27 Expression patterns of anthocyanin biosynthetic and other floral development genes by RT-PCR in flowers of six independent transgenic and three WT tobacco lines at anthesis. The PCR cycle number was optimized for each genes and *NtUBC2* (Koyama et al. 2003) was used as control for equal loading. The genes included in the expression analysis are as follows: *MYB10.1*, peach MYB10.1 TF gene; *NtMYB305*, tobacco MYB305 TF gene orthologous of *Arabidopsis* stamen filament growth related genes *AtMYB21*, *AtMYB24* and *AtMYB57* (Cheng et al. 2009); *NtPAL*, *phenylalanine ammonia-lyase*; *NtCHS*, *chalcone synthase*; *NtCHI*, *chalcone isomerase*; *NtF3H*, *flavanone-3-hydroxylase*; *NtDFR*, *dihydroflavonol 4-reductase*; *NtLAR*, *Leucoanthocyanidin reductase*; *NtANR1*, *anthocyanindin reductase 1*; *NtANS*, *anthocyanidin synthase; UFGT*, *UDP-glycose:flavonoid-3-O-glycosyltransferase; NtJAZd*, encodes a jasmonate ZIM-domain protein; *NtAOS*, *allene oxide synthase*, *NtNEC1*, *nectarin 1*; and *NtACO*, *1-aminocyclopropane-1-carboxylate oxidase*.

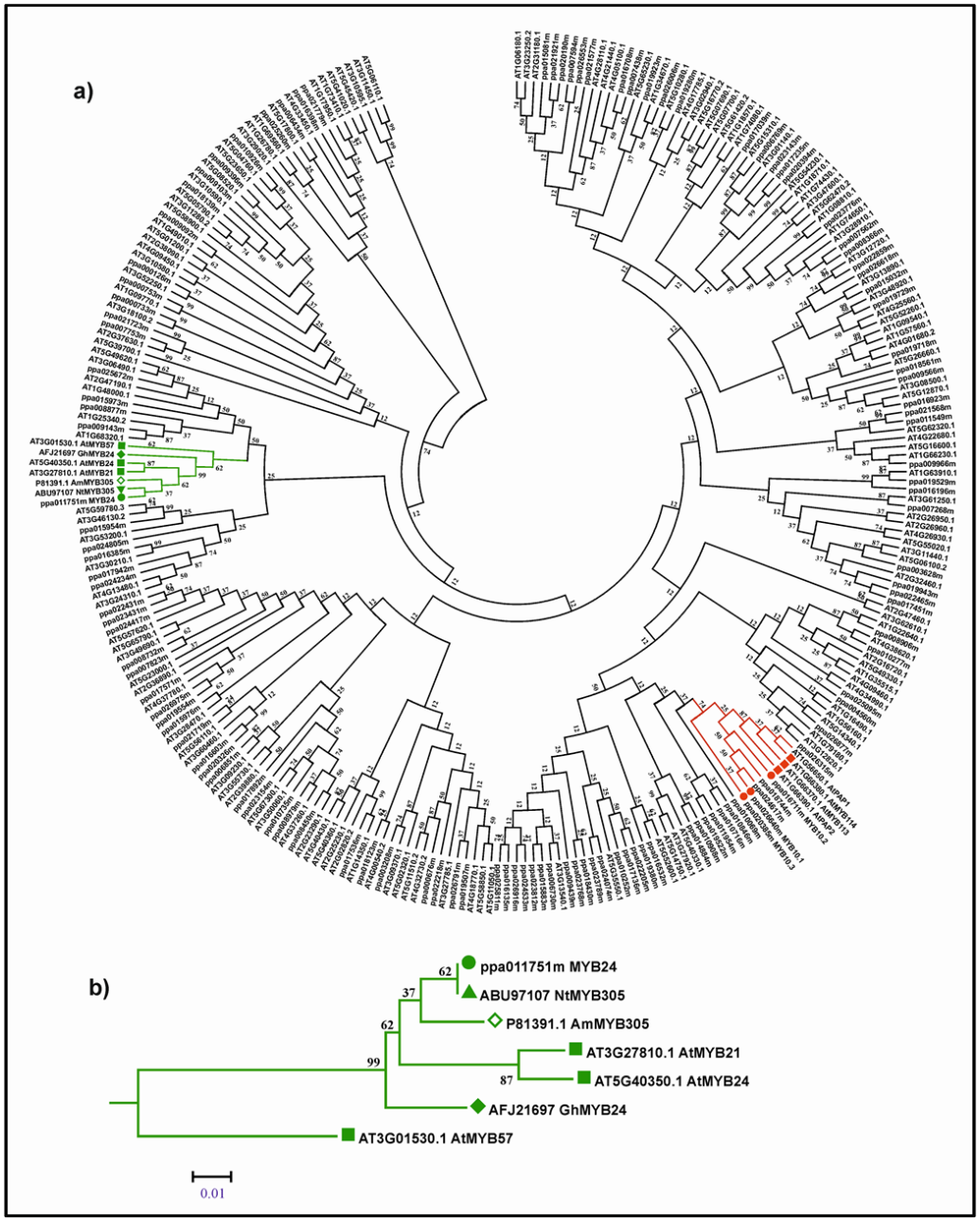

Figure 3.28 Phylogenetic tree showing relationships between *Arabidopsis*, tobacco and peach MYB TFs (a). Magnification of the clade where only MYB TFs related to floral development were clustered (b). Red solid boxes: *Arabidopsis* anthocyanin promoting MYB TFs; green solid boxes: *Arabidopsis* MYB TFs related to flower development (*AtMYB21*, *AtMYB24*, *AtMYB57*); red solid circles: peach MYB10-like TFs; green solid circle: flower specific peach MYB TF, green solid diamond: cotton GhMYB24 (ASJ21697), green unfilled diamond: snapdragon AmMYB305 (P81391) and green solid triangle: tobacco NlxNsMYB305 (ABU97107). The protein sequences of *Arabidopsis* and peach were downloaded from www.phytozome.net. Other protein sequences were retrieved from GeneBank. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 5 (Tamura et al. 2011).

thus to investigate whether the hormone had a role in female reproductive organ development of type-I plants, the expression of the *NtACO* gene encoding 1-aminocyclopropane-1-carboxylate oxidase (ACO) was analyzed. The result showed that the transcript level of *NtACO* was up-regulated in type-I transgenic flowers, while it was at WT levels in type-II plants.

3.2.5 Functional analysis of floral reproductive parts of transgenic lines

To get insights on the inability to set seeds in type-I transgenic plants, their male and female reproductive parts were analyzed in more detail. To this purpose, the pollen viability from non-dehiscent, dehiscent (only few anthers underwent dehiscence but at a very late stage, when flowers had already entered the senescence process), and partially dehiscent anthers was assessed. The pollen viability assay (Wang et al. 2004) showed that about 67.2% of pollen grains were viable in dehiscent anthers, 34.24% in the non-dehiscent and 34.38% in the partially dehiscent anthers while 90% was the viability for the WT pollen (Table 3.6). Therefore, it seems that transgenic pollens are less viable than WT but still able to fertilize ovules, even if the anthers do not open to release pollens. To discriminate whether the problem was due to the failure of anther opening or to the pollen maturation process reciprocal crosses were performed. To check the fertility of the type-I transgenic pistils, stigmas of T_0 flowers were manually pollinated with WT pollen. The fertilization was a total failure and all the flowers were shed from the plants. On the contrary, when WT stigmas were hand pollinated with type-I pollen from lately dehiscent anthers, there was 85% of successful fertilization (Table 3.7). It has to be noted that most of the seedlings coming from these crosses were GUS positive, meaning that they originated from transgenic pollen grains. This indicates that *MYB10.1* somehow affect more the fertility of the female reproductive part than the male one.

Table 3.6 Pollen viability of type-I (strong) transgenic tobacco flowers compared to WT.

	Type-I (strong)			WT	
	Dehiscent anthers	**Non-dehiscent anthers**	**Partially dehiscent anthers**	**Heat killed anthers**	**Not-heat killed anthers**
% viable pollen	67.20	34.24	34.38	0.00	90.00

Table 3.7 Pistil fertility test by manual pollination. Twenty flowers for each cross were pollinated and their fertility was assessed.

	♀**Type-I (Strong)** × ♂ **WT**	♀ **WT** × ♂ **Type-I (Strong)**
% fertilized flowers	0	85

3.2.6 Phenotype of T_1 generation of transgenic lines

The T_1 seeds of type-II and F_1 seeds from crosses between transgenic plant for *35S::Lhg4* and promoterless *MYB10.1* plants were sown in petri dishes on filter paper soaked with water containing kanamycin and hygromycin to allow the growth of only the transgenic plants expressing *MYB10.1*. Some seedlings had a stunted growth, with cotyledons that looked bleached and accumulating pigments and died afterwards (Figure 3.29a-b); on the contrary no pigments were found in WT cotyledons that grew normally (Fig 3.29d). Also the cotyledons of F_1 seedlings (♀ *35S::Lhg4* × ♂ promoterless *MYB10.*1) showed anthocyanin accumulation and chlorophyll bleaching (Figure 3.29c). After germination on petri dishes, some of these T_1 of seedlings were transferred to soil and grown in the greenhouse. The presence of the transgene was confirmed by histochemical GUS-staining and later on by PCR. Several individuals of the T_1 generation originated from the self-fertilization of type-II transgenic plants showed pigment accumulation also in the calyx, capsule and developing seed coats (Figure 3.30) while no pigment accumulation was observed in the capsule of their previous (T_0) generation (Figure 3.23d).

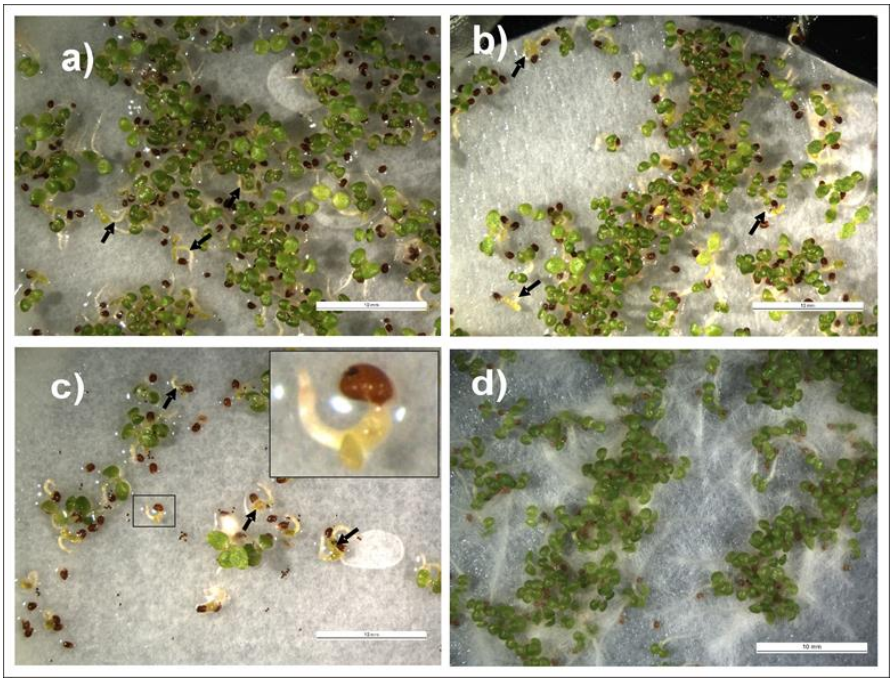

Figure 3.29 Phenotypes of transgenic (a, b and c, over-expressing *MYB10.1*)) and WT (d) tobacco seedlings; a) T_1 seedlings from the type-II 2.2 line (selfed); b) T_1 seedlings from the type-II 10.2 line (selfed); c) F_1 seedlings from the cross: ♀ *35S::Lhg4* x ♂ *35S::MYB10.1*. Arrows indicate pigmented cotyledons of transgenic tobacco seedlings.

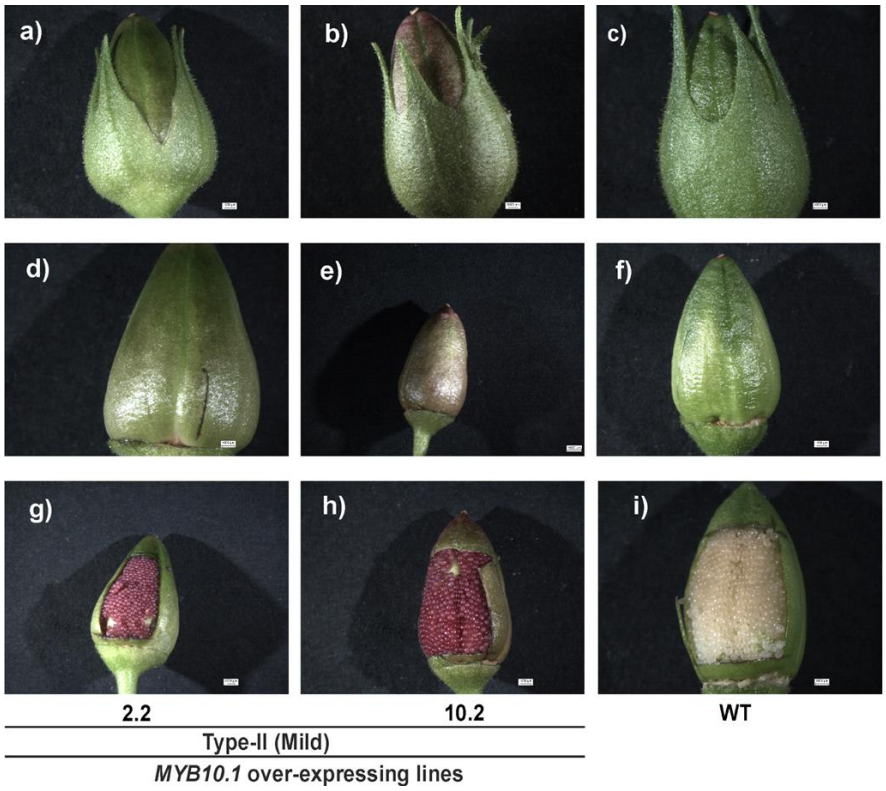

Figure 3.30 Phenotypes of the capsule and developing seed-coat produced by individuals of the T_1 generation from selfed Type-II plants, compared to WT. The images (a-c), (d-f) and (g-i) indicate capsules with calyx, capsules without calyx and open capsules showing developing seeds, respectively.

Chapter IV

4 Discussion

4.1 Regulation of anthocyanin biosynthesis in peach

Peach fruits can accumulate anthocyanins during ripening and this work has shown that there is a differential pigment accumulation in different fruit parts (peel, mesocarp and mesocarp around the stone), besides quantitative differences among different peach cultivars (e.g. 'Redhaven', 'Roza' and 'Fantasia'). Similar pigment amounts were always found in the peel and in the mesocarp around the stone, while negligible quantities were present in the un-pigmented part of the mesocarp (Figure 3.10). This result is in agreement with previous studies on peach (Andreotti et al. 2008, Ravaglia et al. 2013), as most of the cultivars that are on the market display a pigment accumulation pattern that is similar to the one of the genotypes here described.

In peach, the anthocyanin biosynthetic genes were differentially expressed during development and in the three fruit parts at ripening. The transcript levels of the biosynthetic genes (except *CHI*) peaked in the mesocarp around the stone followed by peel and were at lower, sometimes undetectable, levels in the mesocarp of each cultivar. Recently, Ravaglia et al. (2013) reported that the expression of the anthocyanin biosynthetic gene *UFGT* correlated with the anthocyanin concentration in the peel at early and late stages of fruit development whereas the expression of the flavonoid biosynthetic gene *FLS* correlated with the concentration of flavonols at the S1 stage; lastly, also the expression of *ARN* and *LAR* correlated with the accumulation of proanthocyanidins at mid-development in 'Stark Red Gold' nectarines. The present experimental findings also showed similar correlation between anthocyanin accumulation and transcript levels of anthocyanin biosynthetic genes in peach during mesocarp ripening (Table 3.1).

Correlation of the expression of key biosynthetic genes with flavonoid accumulation is well documented in many systems (Grotewold 2006), also in fruit, as Boss et al. (1996) showed that the expression of *UFGT* was high in the skin of

red grapes and undetectable in white grape and Ban et al. (2007) demonstrated that anthocyanin accumulation was positively correlated with the expression level of anthocyanin biosynthetic genes in apple skin.

Besides anthocyanin, also the PA biosynthetic genes *LAR* and *ANR* followed a similar expression pattern but their transcript level was extremely low. Although PA quantification data for the portion of the mesocarp that is around the stone are not available, these expression profiles are in agreement with the PA accumulation found in the skin by (Andreotti et al. 2008, Ravaglia et al. 2013).

In dicot species, anthocyanin early and late biosynthetic genes are differently regulated: the early biosynthetic genes C*HS, CHI* and *F3H* are regulated by R2R3-MYB TFs without co-regulators while the late biosynthetic genes *DFR, LDOX* and *UFGT* are regulated by a MBW-complex.

The MYB family of TFs is diverse in function such as regulation of primary and secondary metabolism, cellular morphogenesis and regulation of meristem formation, responses to biotic and abiotic stresses and the cell cycle (reviewed from Martin and Paz-Ares 1997, Dubos et al. 2010). In particular, the R2R3-MYB TFs regulate anthocyanin biosynthetic genes. In *Arabidopsis*, four genes, *PAP1* (AtMYB75), *PAP2* (AtMYB90), *PAP3* (AtMYB113) and *PAP4* (AtMYB114) belonging to this R2R3-MYB family of TFs, are involved in the regulation of anthocyanin pigment accumulation (Tohge et al. 2005, Gonzalez et al. 2008, Heppel et al. 2013). The *MdMYB1* regulates the anthocyanin pathway genes and the transcript level of this regulator was the genetic basis for apple skin color (Takos et al. 2006). In peach, the MYB-bHLH-WD40 (MBW) complex might regulate the anthocyanin biosynthesis (Figure 3.9). Being the whole peach genome sequence available (Verde et al. 2013), phylogenetic analyses allowed us to select the putative best candidates of the peach MBW complex. A similar approach was undertaken by Ravaglia et al. (2013). As regards the MYBs, these Authors focused their attention on MYB genes from several sub-clades, but the one they chose as MYB10, i.e. ppa016711m, corresponding to our MYB10.2, was not the best candidate among the possible six. On the contrary, we focused our attention only on true MYB10-like sequences. Among all the peach *MYB10*s studied, *MYB10.1* was the

highly expressed one, followed by *MYB10.3*, in the mesocarp around the stone and in the peel in ripe fruit. Being 'Fantasia' very similar to 'Stark Red Gold', the nectarine cultivar used by Ravaglia et al. (2013), it is unlikely that the low expression of *MYB10.2* is linked to the genotype selection. The remaining three genes (ppa024617m, ppa010069m and ppa018744m) were excluded from further analyses as they are not expressed in the fruit. These results together with expression profiles of biosynthetic genes suggest that higher expression of *UFGT* is mainly due to the higher expression of *MYB10.1* and *MYB10.3*, which ultimately channelize the flavonoids pathway to anthocyanin biosynthetic pathway. Besides changes in expression profiles, sub-functionalization among different members of the same gene family may arise due to amino acid changes in critical positions. The proteins of anthocyanin promoting MYB TFs have some critical amino acid residues like arginine (R) in the R2-MYB domain and valine (V) and alanine (A) in the R3-MYB domain which are not conserved in any other R2R3-MYB TFs (Heppel et al. 2013). In peach, *MYB10.1* and *MYB10.2* have the same conserved amino acid residues in conserved positions, like other known anthocyanin related MYB TFs, but *MYB10.3* changed the conserved V (nucleotide position, 211) to isoleucine (I) in the R3-MYB domain (Figure 4.1). Nonetheless, although one critical amino acid underwent a conservative change, still *MYB10.3* is capable to induce pigmentation in tobacco leaves when ectopically expressed, meaning that the protein is still able to transactivate its target(s). The dicot anthocyanins regulating MYB TFs have also a four amino acid conserved motif "[A/S/G]NDV" in the R3-MYB domain (Lin-Wang et al. 2010). This specific motif is perfectly conserved only in *MYB10.1*, while the third position (D) is changed to N and G in MYB10.2 and MYB10.3, respectively. Nonetheless, the two changes might have different effects on protein activity as the first one (D to N in MYB10.2) converts to negative a positive charge, while the latter (D to G), by only removing the positive charge, could be less critical. Only side directed mutagenesis of MYB10.1 at this position could rule out whether these changes are critical for MYB10.2 and MYB10.3 activities, as both proteins could induce anthocyanin accumulation in peach (both the first and the latter) and tobacco (the latter).

The identification of the proper bHLH was easier, as only three peach genes fall into the clade of *TT8, GL3* and *MYC1*. Although the expression profiles of peach *bHLHs* did not show any striking tissue specificity, it was *bHLH3*, which expression level was highest, to show a better correlation with the pigment accumulation and the expression of biosynthetic genes. Thus suggest that, as seen in apple (Ban et al. 2007) and other systems, MYB10 TFs might be the main regulators of the anthocyanin biosynthesis in peach.

Figure 4.1 Protein sequence alignment of peach MYB10 TFs with other MYB TFs that are known to be involved in anthocyanin and PA biosynthesis from other species. Two highly conserved MYB domains were shown above the alignment. The amino acids in the circle (R in the R2-MYB domain, V and A in the R3-MYB domain) are the three key amino acid residues which mediate the specificity of the pathway either anthocyanin or PA (Heppel et al. 2013). The amino acid motif "[DE]Lx_2[RK]$x_3Lx_6Lx_3$R" below the alignment indicates interacting site for bHLH TFs (Zimmermann et al. 2004). The protein sequences of AtMYB123 (AT5G35550); VvMYBPA1 (CAJ90831); peach MYB123 and MYBPA protein sequence were as described by Ravaglia et al. (2013). The remaining protein sequences were same accessions as described in Figure 3.2.

The transient transformation of tobacco leaves demonstrated that both *MYB10.1* and *MYB10.3*, but not *MYB10.*2, in combination with *bHLH3*, were capable to induce the accumulation of anthocyanin pigments. Some of the anthocyanin biosynthetic genes were selected as markers to verify if a preferential transactivation of the target genes could be ascribed to any of the MYB and bHLH TFs tested. The activation of *NtCHS*, *NtDFR* and *NtUFGT* confirmed that *MYB10.1*, with *bHLH3* as partner, was the best gene for anthocyanin biosynthesis. Nonetheless, while *NtCHS* was activated only by MYB10.1 and MYB10.3 in combination with bHLH3, *NtDFR* was activated by MYB10.1 also in combination with bHLH33. Moreover, the second best activation of *NtUFGT* was obtained with the MYB10.3/bHLH33 couple, showing roughly half the transactivation strength of MYB10.1/bHLH3. None of the combinations with MYB10.2 was effective on

tobacco genes, even though it was shown to be able to transactivate the expression of a reporter gene driven by the promoter of the peach *UGFT* gene, in combination with bHLH3, in a heterologous expression system (Ravaglia et al. 2013). Neither the pathway leading to flavonols nor the one for the synthesis of PAs are activated by *MYB10-like* genes. This is in agreement to what has been found in other systems and with recent findings also in peach, as Ravaglia et al. (2013) have shown that the peach LAR promoter is transactivated by the MYBPA1/bHLH3 couple but not by the MYB10.2/bHLH3 one.

The ectopic expression of *MYB10.1* without *bHLH* partner did not ever induce any pigment accumulation in the tobacco leaves (data not shown), thus confirming that in tobacco leaves, MYB10-like proteins work on late genes only when a bHLH co-regulator is present. This is because the anthocyanin biosynthetic pathway is generally not active in tobacco leaves, thus providing a background-free system to study this pathway. On the contrary, the peach mesocarp system is not background-free. The transient transformation of peach fruits showed that all three peach *MYB10* genes were able to induce anthocyanin in the fruit mesocarp, even when *bHLH3* was not present as partner, as for the case of *MYB10.1* (Figure 3.13, panel-b). The anthocyanin accumulation data (Figure 3.17) did not allow to get a clear picture of the activation system in peach. The efficiency of the transient transformation technique in peach is not as good as it is in tobacco, complicated by the presence of large cells (Masia et al. 1992), rich in water and sugars (Vizzotto et al. 1996), that contrast the uniform spreading of agrobacteria within the tissues and that, compared to tobacco leaves, shorten the fruit shelf-life. Thus, despite the fact that, contrary to tobacco, all three peach *MYB10*s work nicely in the peach system, results were too coarse to distinguish the best MYB10/bHLH partners by ectopic expression (Figure 3.17). Nonetheless, this experiment shows that also MYB10.2, despites the abovementioned difference in a critical position, can induce anthocyanin synthesis.

Several traits related to leaf, flower and fruit color have already been mapped on the *Prunus* reference map. For example, “flesh color around the stone (*Cs*)”, “flower color (*Fc*)” and “anther color (*Ag*)” have been mapped on linkage group (LG) 3, and “fruit skin color (*Sc*)” and “leaf color (*Gr*)” have been mapped on LG

6 (Dirlewanger et al. 2004b, Yamamoto et al. 2005). As it is well known that some R2R3-MYB TFs are the key genetic determinant for the regulation of red anthocyanin pigmentation in different species (Kobayashi et al. 2004, Ban et al. 2007, Butelli et al. 2012), it would be interesting to see if also the genes here characterized are linked to the color related traits. This has already been shown, as an example, in apple, where *MdMYB10*, *MdMYBA* and *MdMYB1*, which are allelic, are key determinant for skin, flesh and leaf color (Ban et al. 2007, Lin-Wang et al. 2010, Espley et al. 2013). Gillen and Bliss (2005) mapped blood flesh trait (*bf* locus) to the top of LG 4 from an F_2 population of single F_1 plant from a cross between peach cvs 'Harrow Blood (HB)' × 'Okinawa (Oki)'. Recently, Shen et al. (2013) provided genetic evidence of blood-flesh phenotype of 'Wu Yue Xian' that is controlled by a single dominant locus, called dominant blood flesh (*DBF*) located on the top of LG5. Chagné et al. (2007) identified a locus named *Rni* on LG 9 of the apple genome for the red color of leaf and flesh of apple fruit and showed that *MdMYB10* co-segregated with this locus.

In the available peach genome sequence the three *MYB10s* (*MYB10.1, MYB10.2 and MYB10.3*) here described are all located within 80 kb on linkage group (LG3) (Figure 4.2), within the two markers that were used to map three color related traits (*Cs*, *Fc* and *Ag*) (Dirlewanger et al. 2004b). Similarly, the other three peach MYB10-like TFs (ppa018744m, ppa024617m and ppa010069m) that were also grouped with *Arabidopsis* anthocyanin promoting MYB TFs (AtPAP1/AtMYB75, AtPAP2/AtMYB90, AtPAP3/AtMYB113 and AtPAP4/AtMYB114, see Figure 3.1) are located in LG 6 closed to *fruit skin color* (*Sc*) and *leaf color* (*Gr*). Whereas for the latter the gene positions are weakly associated with the mapped traits, the former practically overlap, strongly suggesting that peach at least one among *MYB10.1*, *MYB10.2* and *MYB10.3* might be the genetic determinant of *flesh color around the stone*, *anther color* and *flower color*, but genetic evidences are needed to support this hypothesis.

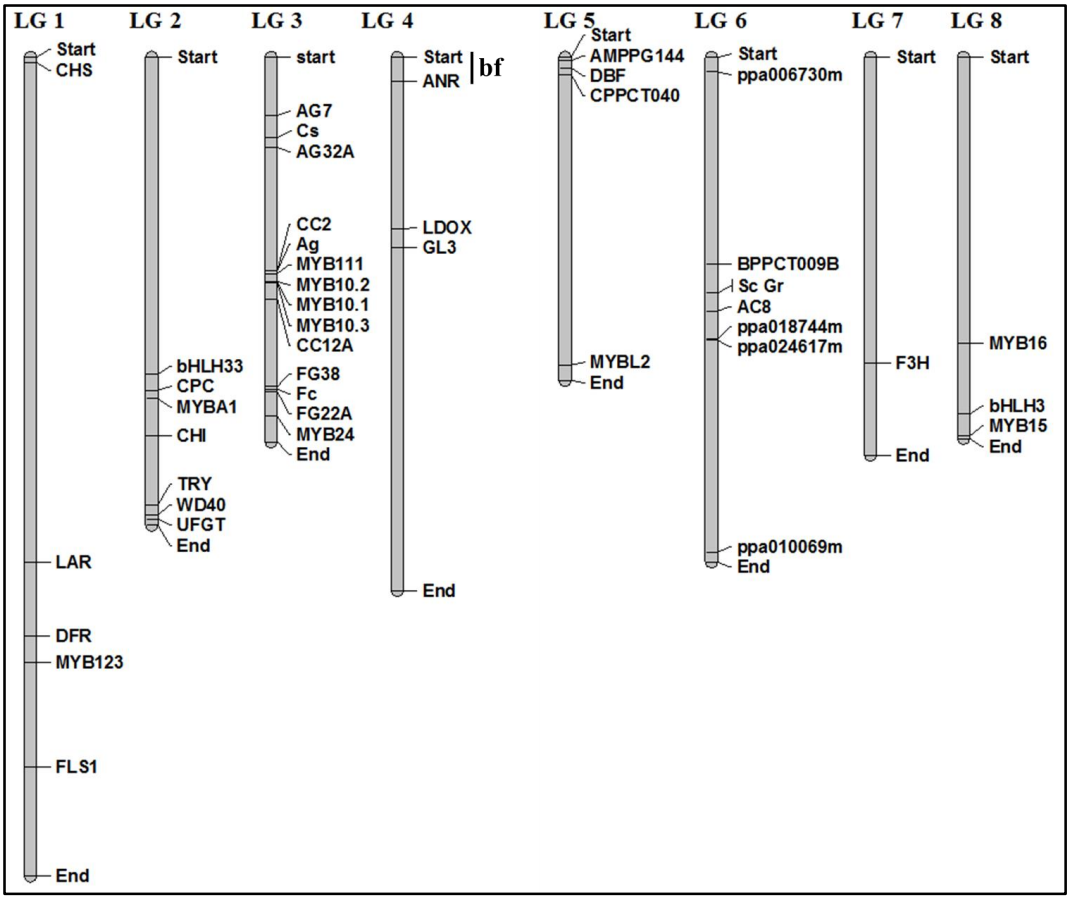

Figure 4.2 Position of color related traits, the regulatory and the biosynthetic genes of anthocyanins on the peach genome. *Cs*, flesh color around the stone; *Ag*, anther color; *Fc*, flower color; *bf*, blood flesh (top of the LG4 according to Gillen and Bliss 2005), *DBF*, dominant blood flesh; *Sc*, fruit skin color; and Gr, leaf color. Each traits were shown in between two molecular markers like AG7-***Cs***-AG32A; CC2-***Ag***-CC12A; FG38-***Fc***-FG22A; AMPPG144-***DBF***-CPPCT040; and BPPCT009B-***Sc/Gr***-AC8. The map was drawn using MapChart version 2.2 (Voorrips 2002).

4.2 Functional characterization of a peach R2R3-MYB transcription factor gene by its over-expression in tobacco

There several reports on both homologous and heterologous over-expression of *MYB* genes like *MdMYBA* in tobacco (Ban et al. 2007), *MdMYB1* in *Arabidopsis* (Takos et al. 2006), *MdMYB10* in apple (Espley et al. 2007), *PyMYB10* in *Arabidopsis* (Feng et al. 2010), *VvMYBA1* in tobacco and grape (Li et al. 2011), *VvMYB5a* in tobacco and grape (Deluc et al. 2006), *LeANT1* in tobacco and tomato (Mathews et al. 2003), *IbMYB1a* in *Arabidopsis* (Chu et al. 2013), *IbMYB1* induced in *Arabidopsis* and sweet potato (Mano et al. 2007), *GMYB10* in tobacco (Elomaa et al. 2003), *AtPAP1* in tobacco and *Arabidopsis* (Borevitz et al. 2000). Ban et al (2007) reported that the ectopic expression of *MdMYBA* in tobacco induces the accumulation of anthocyanins in the reproductive tissues. Feng et al. (2010) also demonstrated similar results when *PyMYB10* was overexpressed in *Arabidopsis*. The ectopic expression of *MYB10.1* showed similar anthocyanin accumulation patterns, always limited to reproductive parts particularly in petals and young seeds. Since tobacco flowers are naturally able to make anthocyanins, the pathway is already operating although only active in the flower. On the contrary, the anthocyanin biosynthetic pathway is not active in tobacco vegetative parts, as leaves, and thus the over-expression of *MYB10.1* could enhance the anthocyanin pigmentation only in the reproductive parts (flowers and young seeds).

Chu et al. (2013) showed that over-expression of *IbMYB1a* gene (coding for sweet potato R2R3-MYB TF) up-regulates structural genes, like *CHI, F3H,* and *DFR*, in the anthocyanin biosynthetic pathway in transgenic *Arabidopsis*. Huang et al. (2013a) also demonstrated that over-expression of *EsMYB1* (encoding a R2R3-MYB TF from *Epimedium sagittatum*) up-regulates important flavonoid related genes in both transgenic tobacco and *Arabidopsis*. Huang et al. (2013b) also found significant up-regulation of anthocyanin biosynthetic genes such as *NtCHS, NtCHI*, *NtDFR* and *NtANS* in transgenic tobacco lines over-expressing *MrMYB1* gene (encoding R2R3-MYB) from Chinese bayberry. Also the heterologous expression of *MYB10.1* modulates the transcription of most of the structural genes in the anthocyanin biosynthetic pathway, thus leading to purple anthocyanin pigmentation in the reproductive tissues of transgenic tobacco plants. These findings suggest that

the over-expression of peach *MYB10.1* up-regulates anthocyanin production in tobacco flower by inducing the transcription of the biosynthetic genes through the interaction with endogenous WD40 and bHLH co-activators.

The *AN2* is a key anthocyanin regulatory MYB gene specifically expressed in reproductive tissues of tobacco and its over-expression up-regulates the flavonoid pathway genes both in transgenic tobacco and *Arabidopsis* while its by RNAi caused the formation of white flowers (Pattanaik et al. 2010). Recently, Huang et al. (2013b) showed that over-expression of *MrMYB1* produced enhanced level of anthocyanin in reproductive parts of transgenic tobacco except leaves by up-regulating anthocyanin biosynthetic genes as well as endogenous *bHLH* (*NtAN1a* and *NtAN1b*) genes. On the contrary, the over-expression of *MYB10.1* slightly decreases *NtAN2* levels, while living *NtAN1b* almost unchanged.

In addition to regulation of anthocyanin biosynthesis, over-expressing of *MYB10.1* caused phenotypic variation in transgenic tobacco lines. After analyzing expression pattern of *MYB10.1* in transgenic tobacco plants, it was found that tobacco plants (type-I) with higher expression of transgenes were defective with their vegetative and reproductive development. On the contrary, tobacco plants (type-II) with moderate expression of transgenes were normal in their vegetative and floral development excluding seed coat pigmentation. The type-I transgenic lines exhibit shorter plants and reduced leaf length but similar leaf breadth as compared to type-II and WT. This is because of the irregular size and shape of epidermal cells such as reduced length and increased breadth as well as less number of epidermal cells per unit area. This suggests the involvement of *MYB10.1* on the plant cell growth and developments, as a similar regulatory system composed of TTG1(WD40)-EG3/EGL3(bHLH)-GL1(R2R3-MYB) has been widely studied in *Arabidopsis* epidermal cell differentiation (Grebe 2012). Although there are reports that sustain that "the trichomes of *Arabidopsis* and *Nicotiana* are merely analogous structures, and that the *MYB* genes regulating their differentiation are specific and separate" (Payne et al. 1999), and that the *bHLH* genes have no effect on epidermal cell development in plants belonging to the Asterid division (Serna and Martin 2006), we cannot exclude that the high expression levels achieved by means of the pOp/LhG4 expression system (Rutherford et al. 2005) of *MYB10.1* had an impact

on cell fate determination, reducing their sizes and thus decreasing plant and leaf growth.

Li et al. (2013) reported that *GhMYB24* (R2R3-MYB) is preferentially expressed in anthers/pollen. Its over-expression in *Arabidopsis* leads to sterile plants while lower to moderate expression level of transgene produce fertile plants. A similar phenotype was observed when peach *MYB10.1* was over-expressed in tobacco plants, even though the two MYBs belong to different clades. Thus the out-titration of MYB proteins, similar in sequence but differing in function, might affect endogenous tobacco MYBs, even those that are involved in other functions. It is known that *AtMYB21*, *AtMYB24*, and *AtMYB57* are mainly expressed in the flowers and play a critical role for the reproductive organ development (Cheng et al. 2009, Dubos et al. 2010, Li et al. 2006, Shin et al. 2002). Yang et al. (2007) showed that *AtMYB24* was expressed in flowers specifically in microspores and ovules and plays important role in anther development. According to them, over-expression of *AtMYB24* cause pleiotropic phenotypes such as reduced plant height as well as defective anther development. After double mutant (*myb21, myb24*) analysis in *Arabidopsis*, Mandaokar et al. (2006) showed that *AtMYB21* and *AtMYB24* are involved in floral organ development such as flower opening, petal expansion, filament elongation, anther dehiscence, inhibition of lateral vascular development in unfertilized carpels, and abscission of sepals, petals and stamens. In tobacco, NtMYB305 (encoding a R2R3-MYB TF most similar to AtMYB21, AtMYB24, and AtMYB57) is involved in floral organ developmental processes related to nectary gland formation (Liu et al. 2009). In transgenic tobacco plants over-expressing peach *MYB10.1*, *NtMYB305* showed extremely low levels of expression in type-I transgenic plants with defective flowers. Thus the nectary phenotype of *MYB10.1* over-expressing plant might be explained by the suppression of *NtMYB305*. More complicated is the explanation of the anther phenotype, as *NtMYB305* has been reported to be not expressed in this flower part (Liu et al. 2009), as, on the contrary is for its strict homologues from *Arabidopsis* (MYB21/MYB24, Mandaokar et al. 2006), cotton (Li et al. 2013), but also peach (Figure 4.3).

Jasmonate (JA) controls several plant processes including plant growth, fertility, development, anthocyanin accumulation and defense and within the signaling cascades activated by JA, the JASMONATE-ZIM DOMAIN (JAZ) repressor proteins are the major components (reviewed from Pauwels and Laurens 2011, Browse 2009). Thines et al. (2007) demonstrated that AtJAZ1 protein repressed the transcription of JA responsive genes and application of exogenous JA initiate JAZ1 degradation. Qi et al (2011) reported that JAZ proteins interact with bHLH TFs like Transparent Testa8 (TT8), Glabra3 (GL3) and Enhancer of Glabra3 (EGL3) and R2R3-MYB TFs like MYB75/PAP1 and Glabra1 (GL1) to repress JA-mediated anthocyanin production and trichome formation. In the absence of JA, JAZ proteins bind to the downstream TFs and limits their transcriptional activity while availability of JA leads to degradation of the JAZ proteins to free the downstream TFs for the transcriptional regulation of target genes (Chini 2007, Qi et al. 2011). JA is involved in stamen development and pollen maturation process in plants. It has been demonstrated by Mandaokar et al. (2006) that *AtMYB21* and *AtMYB24* are induced by JA, while Song et al. (2011) reported that JAZ proteins interact with *AtMYB21* and *AtMYB24* to decrease their transcriptional function; upon perception of JA signal, *COI1* recruits JAZs to the SCF^{COI1} complex for ubiquitination and degradation through the 26S proteasome to release *AtMYB21* and *AtMYB24* thus triggering transcription of various important genes for JA-mediated stamen development. JA deficient mutants (*opr3*) triggered male-sterility but fertility was restored by exogenous JA application (Cheng et al. 2009).

Li et al. (2013) stated that the over-expression of *GhMYB24* (code for a R2R3-MYB TF from cotton) might modulate the expression of *AtJAZ1* and other genes in the JA signaling pathway. The over-expression of *MYB10.1* also up-regulates the JA biosynthetic and signaling pathway genes *NtAOS* and *NtJAZd* (orthologous to *AtJAZ1* gene), in type-I transgenic tobacco flowers. Thus suggests an imbalance in JA action in these transgenic plants, which show a phenotype similar to JA deficient mutants but increased levels in transcripts coding for JA synthesis and action. This imbalance occurs only when high levels of *MYB10.1* accumulate in transgenic plants, probably as a feed-back mechanism that is trying to counteract an induction usually mediated by JA. On the contrary, type-II and WT showed normal

expression pattern of JA-related genes and flowers has normal development, probably because these plants could react to the higher *MYB10.1* levels by decreasing *NtAN2* and *NtMYB305* levels, thus keeping the MYB-bHLH-WD40 (MBW) complex below threshold level. Li et al. (2004) characterized a tomato mutant, *jasmonic acid–insensitive1* (*jai1*) with defective in JA signaling resulting in female reproductive part sterility. They also reported that the sterility was due to the defect in the maternal control of seed maturation linked with the loss of accumulation of JA regulated proteinase inhibitor proteins in reproductive tissues. Thus, an impaired JA metabolisms might be the cause of altered female fertility also in tobacco plants over-expressing *MYB10.1*.

Also ethylene plays an important role in early stages of female sporogenesis and ovule fertilization in tobacco (De Martinis and Mariani 1999). They showed that suppression of a pistil-specific *NtACO* gene caused female sterility due to an arrest in ovule development in transgenic tobacco plants. Type-I transgenic tobacco flowers showed opposite result with higher expression level of *NtACO* gene, possibly to balance the enhanced JA level.

The nectary is rich in carbohydrates and secretes nectar to attract pollinators as well as defend floral reproductive tissues against microorganisms (Carter et al. 2007). The tobacco nectar contains five different types of nectarin proteins (NEC1 to NEC5) but NEC1 is most abundant and its expression is restricted to nectary (reviewed from Carter and Thornburg 2004)). The expression of *NtNEC1* is regulated by tobacco *NtMYB305* (Liu et al. 2009). Liu et al. (2009) also demonstrated that the expression level of *NtMYB305*, *NtNEC1*, *NtNEC5* and anthocyanin biosynthetic genes such as *NtPAL* and *NtCHI* were reduced in tobacco by *NtMYB305* knockdown experiment. The type-I transgenic flowers also showed similar results. These transgenic flowers are defective in floral development and have no nectary gland revealed that with extremely low expression, *NtMYB305* could not activate transcription of *NtNEC1* in type-I transgenic tobacco flowers. Therefore, these transgenic flowers failed to develop nectary gland at the base of ovary.

To better elucidate the role of tobacco NtMYB305 TF, the expression levels of *MYB24* (orthologous of *NtMYB305*, *AmMYB305*, and *AtMYB21*, *AtMYB24* and *AtMYB57*) and of the tree previously identified *MYB10s* (*MYB10.1*, *MYB10.*2 and *MYB10.1*) genes were analyzed in different parts of peach flower by qRT-PCR. The expression profiles revealed that in peach, *MYB24* expression was high in sepal, petal, and androecium but low in gynoecium and was undetected in the fruit mesocarp (Figure 4.3) suggesting that *MYB24* is flower specific. In addition, among the three anthocyanin promoting *MYB10* genes, the expression level of *MYB10.1* was the highest, being higher in androecium followed by petal whereas almost undetectable in the gynoecium (Figure 4.3). The high transcript level of *MYB10.1* in androecium is in agreement with the anthocyanin content accumulation in purple stamens, highest than in any other part of the peach flower (Figure 4.4). This expression profile is in agreement with the stamen specificity observed for *Arabidopsis* AtMYB24, but differs from the nectary specific expression of *NtMYB305*. Considering these findings, it is hypothesized that the over-expression of *MYB10.1*, by altering JA/ethylene levels, suppress *NtMYB305* thus causing defective flowers/missing nectary.

Based on the present results obtained through transgenic plants analysis, it can be concluded that over-expression of *MYB10.1* in tobacco regulates anthocyanin biosynthesis in the reproductive parts. Furthermore, they might play direct or indirect role in other process like vegetative and reproductive development. These findings will promote further investigation to elucidate the role of R2R3-MYB TFs.

Description	Mean normalized expression			
	MYB10.1	*M2YB10.2*	*MYB10.3*	*MYB24*
Complete flower	0.290	0.143	0.038	1.876
Sepal	0.152	0.242	0.028	1.387
Petal	0.842	0.327	0.129	2.516
Androecuim	1.354	0.195	0.160	1.783
Gynoecium	0.006	0.022	ND	0.123
Mesocarp around the stone	0.593	ND	0.062	ND

Figure 4.3 Heat map showing the expression levels of *MYB10.1*, *MYB10.2*, *MYB10.3* and *MYB24* by qRT-PCR in peach flower (Complete flower, sepals, petals, androecium, and gynoecium) and, as a comparison with previous experiments, in fruit mesocarp around the stone. The expression values of the target genes were normalized by the expression values of the *PpN1* gene, used as internal standard. The highest expression value has been arbitrarily set to 100 (blue), and the others accordingly (0 = white).

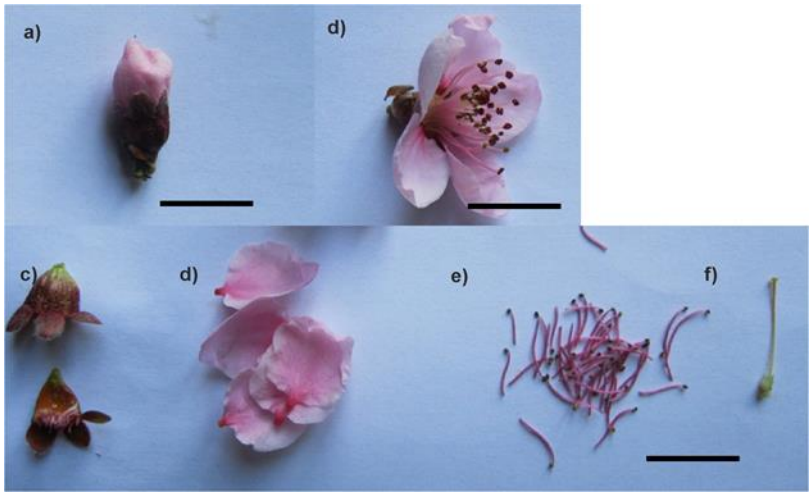

Figure 4.4 Different parts of peach flowers. A flower bud (a), a fully bloomed flower (b), sepals (c), petals (d), androecium (e), and gynoecium (f).

Chapter V

5 Conclusions

5.1 Regulation of anthocyanin biosynthesis in peach

Peaches are fleshy fruits and they change color during ripening due to the accumulation of phenolic compound including anthocyanins. Anthocyanins are beneficial to human health due to their antioxidant potentiality. Three MYB10-like and three bHLH-like encoding genes were identified as best candidates to be the regulators of the anthocyanin accumulation in peach. In particular, *MYB10.1* is the most highly expressed gene in ripe peach fruits. Anthocyanins accumulate differentially in the ripe fruit with the highest content in the peel, followed by the mesocarp around the stone and the mesocarp proper with the lowest levels. The expression of *MYB10.*1 and *MYB10.3* genes correlates well with the anthocyanin levels in the different fruit parts. They have also positive correlation with the expression of key structural genes of the anthocyanin pathway. Interestingly, the transient transformation of tobacco leaves revealed that the over-expression of peach *MYB10.1*/*bHLH3* and *MYB10.3*/*bHLH3* can induce anthocyanin production by up-regulating the *NtDFR* and *NtUFGT* genes. These results suggest that *MYB10.1* is the best regulator followed by *MYB10.3* and can induce anthocyanin synthesis with *bHLH3* as a co-regulator. These findings will aid fruit breeders to obtain new peach cultivars with higher antioxidant level as well as colored fruit.

5.2 Functional characterization of a peach R2R3-MYB transcription factor gene by its over-expression in tobacco

The functional characterization of peach *MYB10.1* (which is the highly expressed genes among three *MYB10*s in ripe fruit) was carried out by tobacco stable transformation. The over-expression of peach *MYB10.1* induces anthocyanin pigmentation in the reproductive parts and up-regulates the expression level of many important genes of the flavonoid biosynthetic pathway in tobacco flower, more specifically of the anthocyanin pathway like *NtCHS*, *NtCHI*, *NtF3H*, *NtDFR*, *NtANS* and *NtUFGT* as well as of the proanthocyanindin biosynthetic pathway like *NtLAR* in both types of transgenic tobacco flowers. Besides, *MYB10.1* over-expression also causes phenotypic alteration in transgenic tobacco lines. The strong phenotypes (type-I) exhibited irregular leaf shape and size, reduced plant height, reduced filament length of anthers, non-dehiscent anthers, reduced pistil length, no nectary gland formation and no capsule development but reproductive parts including androecium, gynoecium and petals were pigmented. On the other hand, mild phenotypes (type-II) showed regular plant growth and development but the developing seed coats were pigmented. Surprisingly, over-expression of peach *MYB10.1* leads to suppression of *NtMYB305*, which is required for floral development, in type-I transgenic tobacco plants. The *NECTARIN 1* (*NtNCE1*) that is known to be involved in nectary gland formation was also suppressed since this gene is regulated by *NtMYB305*. Moreover, peach *MYB10.1* somehow up-regulates JA biosynthetic (*NtAOS)* and signaling (*NtJAZd)* genes, and *1-aminocyclopropane-1-carboxylate oxidase* (*NtACO)*. The alteration of these hormonal pathways might be among the causes of the observed floral abnormalities. On the contrary, the type-II transgenic plants showed opposite results for *NtMYB305*, *NtNEC1* and *NtJAZd* thus showing that the mild over-expression of *MYB10.1* in these plants might mimic that of *NtMYB305*. Finally, it can be concluded that the over-expression of peach *MYB10.1* in tobacco not only regulates flavonoid biosynthesis (anthocyanin and proanthocyanidin) in the reproductive parts but also plays a role in other processes like vegetative and reproductive development.

Chapter VI

6 References

Aharoni A, De Vos C, Wein M, Sun Z, Greco R, Kroon A, Mol JN, O'Connell AP (2001) The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. The Plant Journal 28:319-332

Alba R, Payton P, Fei Z, McQuinn R, Debbie P, Martin GB, Tanksley SD, Giovannoni JJ (2005) Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. The Plant Cell 17:2954-2965

Alkema J, Seager SL (1982) The chemical pigments of plants. J Chem Educ 59:183

Allan AC, Hellens RP, Laing WA (2008) MYB transcription factors that colour our fruit. Trends Plant Sci 13:99-102

An XH, Tian Y, Chen KQ, Wang XF, Hao YJ (2012) The apple WD40 protein MdTTG1 interacts with bHLH but not MYB proteins to regulate anthocyanin accumulation. J Plant Physiol 169:710-717

Andreotti C, Ravaglia D, Ragaini A, Costa G (2008) Phenolic compounds in peach (*Prunus persica*) cultivars at harvest and during fruit maturation. Ann Appl Biol 153:11-23

Atchley WR, Terhalle W, Dress A (1999) Positional dependence, cliques, and predictive motifs in the bHLH protein domain. J Mol Evol 48:501-516

Azeredo H (2009) Betalains: properties, sources, applications, and stability- a review. Int J Food Sci Tech 44:2365-2376

Bailey PC, Martin C, Toledo-Ortiz G, Quail PH, Huq E, Heim MA, Jakoby M, Werber M, Weisshaar B (2003) Update on the basic helix-loop-helix transcription factor gene family in *Arabidopsis* thaliana. The Plant Cell 15:2497-2502

Ban Y, Honda C, Hatsuyama Y, Igarashi M, Bessho H, Moriguchi T (2007) Isolation and functional analysis of a MYB transcription factor gene that is a key regulator for the development of red coloration in apple skin. Plant and cell physiology 48:958-970

Bapat VA, Trivedi PK, Ghosh A, Sane VA, Ganapathi TR, Nath P (2010) Ripening of fleshy fruit: molecular insight and the role of ethylene. Biotechnol Adv 28:94-107

Bassi D, Monet R (2008) Botany and Taxonomy. In: Bassi D, Layne D (eds) The Peach: Botany, Production and Uses. CABI international, Oxfordshire OX10 8DE, UK, pp 1-36

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

Berger F, Linstead P, Dolan L, Haseloff J (1998) Stomata patterning on the hypocotyl of *Arabidopsis thaliana* is controlled by genes involved in the control of root epidermis patterning. Dev Biol 194:226-234

Bomal C, Duval I, Giguère I, Fortin É, Caron S, Stewart D, Boyle B, Séguin A, MacKay JJ (2013) Opposite action of R2R3-MYBs from different subgroups on key genes of the shikimate and monolignol pathways in spruce. J Exp Bot:ert398 [doi: 10.1093/jxb/ert398]

Bonghi C, Trainotti L, Botton A, Tadiello A, Rasori A, Ziliotto F, Zaffalon V, Casadoro G, Ramina A (2011) A microarray approach to identify genes involved in seed-pericarp cross-talk and development in peach. BMC plant biology 11:107

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

Borovsky Y, Oren-Shamir M, Ovadia R, De Jong W, Paran I (2004) The *A* locus that controls anthocyanin accumulation in pepper encodes a *MYB* transcription factor homologous to *Anthocyanin2* of *Petunia*. Theor Appl Genet 109:23-29

Boss PK, Davies C, Robinson SP (1996) Expression of anthocyanin biosynthesis pathway genes in red and white grapes. Plant Mol Biol 32:565-569

Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254

Bramley PM (2002) Regulation of carotenoid formation during tomato fruit ripening and development. J Exp Bot 53:2107-2113

Brivanlou AH, Darnell JE (2002) Signal transduction and the control of gene expression. Science 295:813-818

Brooks SJ, Moore JN, Murphy JB (1993) Quantitative and qualitative changes in sugar content of peach genotypes [*Prunus persica* (L.) Batsch.]. J Am Soc Hort Sci 118:97-100

Brouillard R (1982) Chemical structure of anthocyanins. Academic Press: New York

Browse J (2009) Jasmonate passes muster: a receptor and targets for the defense hormone. Annual review of plant biology 60:183-205

Busatto N (2012) Functional characterization of a ripening induced RGF-like peptide hormone in peach. PhD Thesis. Università degli studi di Padova:97

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato-Recupero G, Martin C (2012) Retrotransposons Control Fruit-Specific, Cold-Dependent Accumulation of Anthocyanins in Blood Oranges. Plant Cell 24: 1242–1255

Butelli E, Titta L, Giorgio M, Mock HP, Matros A, Peterek S, Schijlen EGWM, Hall RD, Bovy AG, Luo J (2008) Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nat Biotechnol 26:1301-1308

Carey CC, Strahle JT, Selinger DA, Chandler VL (2004) Mutations in the pale aleurone color1 regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar TRANSPARENT TESTA GLABRA1 gene in *Arabidopsis thaliana*. The Plant Cell 16:450-464

Carr MD, Wollborn U, McIntosh PB, Frenkiel TA, McCormick JE, Bauer CJ, Klempnauer K, Feeney J (1996) Structure of the B‐Myb DNA‐binding Domain in solution and evidence for multiple conformations in the region of repeat‐2 involved in DNA binding. European Journal of Biochemistry 235:721-735

Carter C, Healy R, Nicole M, Naqvi SS, Ren G, Park S, Beattie GA, Horner HT, Thornburg RW (2007) Tobacco nectaries express a novel NADPH oxidase implicated in the defense of floral reproductive tissues against microorganisms. Plant Physiol 143:389-399

Carter C, Thornburg RW (2004) Is the nectar redox cycle a floral defense against microbial attack? Trends Plant Sci 9:320-324

Chagné D, Carlisle C, Blond C, Volz R, Whitworth C, Oraguzie N, Crowhurst R, Allan A, Espley R, Hellens R (2007) Mapping a candidate gene (MdMYB10) for red flesh and foliage colour in apple. BMC Genomics 8:212

Chalmers D, Ende B (1975) A reappraisal of the growth and development of peach fruit. Functional Plant Biology 2:623-634

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Rep 11:113-116

Chang S, Tan C, Frankel EN, Barrett DM (2000) Low-density lipoprotein antioxidant activity of phenolic compounds and polyphenol oxidase activity in selected clingstone peach cultivars. J Agric Food Chem 48:147-151

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL, Veatch O, Herdes RE, Diebold L, Saw J (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. Genetics 177:2481-2491

Cheng H, Song S, Xiao L, Soo HM, Cheng Z, Xie D, Peng J (2009) Gibberellin acts through jasmonate to control the expression of *MYB21*, *MYB24*, and *MYB57* to promote stamen filament growth in *Arabidopsis*. PLoS genetics 5:e1000440

Chini A, Fonseca S, Fernandez G, Adie B, Chico J, Lorenzo O, Garcia-Casado G, Lopez-Vidriero I, Lozano F, Ponce M (2007) The JAZ family of repressors is the missing link in jasmonate signalling. Nature 448:666-671

Chiu LW, Zhou X, Burke S, Wu X, Prior RL, Li L (2010) The purple cauliflower arises from activation of a MYB transcription factor. Plant Physiol 154:1470-1480

Chu H, Jeong JC, Kim W, Chung DM, Jeon HK, Ahn YO, Kim SH, Lee H, Kwak S, Kim CY (2013) Expression of the sweetpotato R2R3-type IbMYB1a gene induces anthocyanin accumulation in *Arabidopsis*. Physiol Plantarum 148:189-99

Chung MY, Vrebalov J, Alba R, Lee JM, McQuinn R, Chung JD, Klein P, Giovannoni J (2010) A tomato (*Solanum lycopersicum*) *APETALA2/ERF* gene, *SlAP2a*, is a negative regulator of fruit ripening. The Plant Journal 64:936-947

Craft J, Samalova M, Baroux C, Townley H, Martinez A, Jepson I, Tsiantis M, Moore I (2005) New pOp/LhG4 vectors for stringent glucocorticoid‐dependent transgene expression in *Arabidopsis*. The Plant Journal 41:899-918

Craig W, Beck L (1999) Phytochemicals: Health Protective Effects. Canadian journal of dietetic practice and research: a publication of Dietitians of Canada Revue canadienne de la pratique et de la recherche en dietetique: une publication des Dietetistes du Canada 60:78

De Martinis D, Mariani C (1999) Silencing gene expression of the ethylene-forming enzyme results in a reversible inhibition of ovule development in transgenic tobacco plants. The Plant Cell 11:1061-1071

de Vetten N, Quattrocchio F, Mol J, 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

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

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

Dirlewanger E, Cosson P, Boudehri K, Renaud C, Capdeville G, Tauzin Y, Laigret F, Moing A (2006) Development of a second-generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit. Tree genetics & genomes 3:1-13

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O (2004a) Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid—location of root-knot nematode resistance genes. TAG Theoretical and Applied Genetics 109:827-838

Dirlewanger E, Graziano E, Joobeur T, Garriga-Calderé F, Cosson P, Howad W, Arús P (2004b) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proc Natl Acad Sci USA 101:9891

Dirlewanger E, Pronier V, Parvery C, Rothan C, Guye A, Monet R (1998) Genetic linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular markers. TAG Theoretical and Applied Genetics 97:888-895

Dixon RA, Steele CL (1999) Flavonoids and isoflavonoids- a gold mine for metabolic engineering. Trends Plant Sci 4:394-400

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

EFS (1898) Ripening of Fleshy Fruits. Am Nat :208-210

Eliassen AH, Hendrickson SJ, Brinton LA, Buring JE, Campos H, Dai Q, Dorgan JF, Franke AA, Gao Y, Goodman MT (2012) Circulating carotenoids and risk of breast cancer: pooled analysis of eight prospective studies. J Natl Cancer Inst 104:1905-1916

Elomaa P, Uimari A, Mehto M, Albert VA, Laitinen RA, Teeri TH (2003) Activation of anthocyanin biosynthesis in *Gerbera hybrida* (Asteraceae) suggests conserved protein-protein and protein-promoter interactions between the anciently diverged monocots and eudicots. Plant Physiol 133:1831-1842

Escribano-Bailón MT, Santos-Buelga C, Rivas-Gonzalo JC (2004) Anthocyanins in cereals. Journal of Chromatography A 1054:129-141

Espley RV, Bovy A, Bava C, Jaeger SR, Tomes S, Norling C, Crawford J, Rowan D, McGhie TK, Brendolise C (2013) Analysis of genetically modified red-fleshed apples reveals effects on growth and consumer attributes. Plant biotechnology journal 11:408-419

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC (2007) Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. The Plant Journal 49:414-427

Faust M, Timon B (1995) Origin and dissemination of peach. Hort.Rev 17:331-379

Feller A, Machemer K, Braun EL, Grotewold E (2011) Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. The Plant Journal 66:94-116

Feng S, Wang Y, Yang S, Xu Y, Chen X (2010) Anthocyanin biosynthesis in pears is regulated by a R2R3-MYB transcription factor PyMYB10. Planta 232:245-255

Ferré-D'Amaré AR, Prendergast GC, Ziff EB, Burley SK (1993) Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. Nature 363:38-45

Fisher DK, Guiltinan MJ (1995) Rapid, efficient production of homozygous transgenic tobacco plants with *Agrobacterium tumefaciens*: a seed-to-seed protocol. Plant Mol Biol Rep 13:278-289

Fraser LG, Seal AG, Montefiori M, McGhie TK, Tsang GK, Datson PM, Hilario E, Marsh HE, Dunn JK, Hellens RP (2013) An R2R3 MYB transcription factor determines red petal colour in an *Actinidia* (kiwifruit) hybrid population. BMC Genomics 14:28

Fraser LG, Tsang GK, Datson PM, De Silva HN, Harvey CF, Gill GP, Crowhurst RN, McNeilage MA (2009) A gene-rich linkage map in the dioecious species *Actinidia chinensis* (kiwifruit) reveals putative X/Y sex-determining chromosomes. BMC Genomics 10:102

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13:207-209

Gan L, Xia K, Chen J, Wang S (2011) Functional characterization of TRICHOMELESS2, a new single-repeat R3 MYB transcription factor in the regulation of trichome patterning in *Arabidopsis*. BMC plant biology 11:176

Gaziano JM, Manson JE, Branch LG, Colditz GA, Willett WC, Buring JE (1995) A prospective study of consumption of carotenoids in fruits and vegetables

and decreased cardiovascular mortality in the elderly. Ann Epidemiol 5:255-260

Gillen AM, Bliss FA (2005) Identification and mapping of markers linked to the Mi gene for root-knot nematode resistance in peach. J Am Soc Hort Sci 130:24-33

Giovannoni J (2001) Molecular biology of fruit maturation and ripening. Annual review of plant biology 52:725-749

Giovannoni JJ (2004) Genetic regulation of fruit development and ripening. The Plant Cell 16:S170-S180

Goff SA, Cone KC, Chandler VL (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

Gonzalez A, Zhao M, Leavitt JM, Lloyd AM (2008) Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. The Plant Journal 53:814-827

Goodman CD, Casati P, Walbot V (2004) A multidrug resistance–associated protein involved in anthocyanin transport in *Zea mays*. The Plant Cell 16:1812-1826

Goodrich J, Carpenter R, Coen ES (1992) A common gene regulates pigmentation pattern in diverse plant species. Cell 68:955-964

Grebe M (2012) The patterning of epidermal hairs in *Arabidopsis*-updated. Curr Opin Plant Biol 15:31-37

Grotewold E (2006) The genetics and biochemistry of floral pigments. Annu.Rev.Plant Biol. 57:761-780

Grotewold E, Athma P, Peterson T (1991) Alternatively spliced products of the maize P gene encode proteins with homology to the DNA-binding domain of myb-like transcription factors. Proceedings of the National Academy of Sciences 88:4587-4591

Grotewold E, Sainz MB, Tagliani L, Hernandez JM, Bowen B, Chandler VL (2000) Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R. Proceedings of the National Academy of Sciences 97:13579-13584

Han K, Matsumoto A, Shimada K, Sekikawa M, Fukushima M (2007) Effects of anthocyanin-rich purple potato flakes on antioxidant status in F344 rats fed a cholesterol-rich diet. Br J Nutr 98:914-921

Hayama H, Shimada T, Fujii H, Ito A, Kashimura Y (2006) Ethylene-regulation of fruit softening and softening-related genes in peach. J Exp Bot 57:4071-4077

Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC (2003) The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Mol Biol Evol 20:735-747

Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM (2000) pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol Biol 42: 819–832

Hemmat M, Weedon N, Manganaris A, Lawson D (1994) Molecular marker linkage map for apple. J Hered 85:4-11

Heppel SC, Jaffé FW, Takos AM, Schellmann S, Rausch T, Walker AR, Bogs J (2013) Identification of key amino acids for the evolution of promoter target specificity of anthocyanin and proanthocyanidin regulating MYB factors. Plant Mol Biol:1-15

Hippolyte I, Bakry F, Seguin M, Gardes L, Rivallan R, Risterucci AM, Jenny C, Perrier X, Carreel F, Argout X (2010) A saturated SSR/DArT linkage map of Musa acuminata addressing genome rearrangements among bananas. BMC plant biology 10:65

Hobert O (2008) Gene regulation by transcription factors and microRNAs. Science Signaling 319:1785

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

Honda C, Kotoda N, Wada M, Kondo S, Kobayashi S, Soejima J, Zhang Z, Tsuda T, Moriguchi T (2002) Anthocyanin biosynthetic genes are coordinately expressed during red coloration in apple skin. Plant Physiology and Biochemistry 40:955-962

Huang W, Sun W, Lv H, Luo M, Zeng S, Pattanaik S, Yuan L, Wang Y (2013a) A R2R3-MYB transcription factor from *Epimedium sagittatum* regulates the flavonoid biosynthetic pathway. PloS one 8:e70778

Huang Y, Song S, Allan AC, Liu X, Yin X, Xu C, Chen K (2013b) Differential activation of anthocyanin biosynthesis in *Arabidopsis* and tobacco over-expressing an R2R3 MYB from Chinese bayberry. Plant Cell, Tissue and Organ Culture (PCTOC):1-9

Jefferson RA, FAU KT, Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. The EMBO journal 6:3901-3907

Jia L, Clegg MT, Jiang T (2004) Evolutionary dynamics of the DNA-binding domains in putative R2R3-MYB genes identified from rice subspecies *indica* and *japonica* genomes. Plant Physiol 134:575-585

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

Jung S, Wu K, Giovannucci E, Spiegelman D, Willett WC, Smith-Warner SA (2013) Carotenoid intake and risk of colorectal adenomas in a cohort of male health professionals. Cancer Causes & Control 24:705-717

Kapila J, De Rycke R, Van Montagu M, Angenon G (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. Plant science 122:101-108

Karlova R, Rosin FM, Busscher-Lange J, Parapunova V, Do PT, Fernie AR, Fraser PD, Baxter C, Angenent GC, de Maagd RA (2011) Transcriptome and metabolite profiling show that APETALA2a is a major regulator of tomato fruit ripening. The Plant Cell 23:923-941

Kataoka I, Beppu K (2004) UV irradiance increases development of red skin color and anthocyanins in 'Hakuho' peach. HortScience 39:1234-1237

Kenis K, Keulemans J (2005) Genetic linkage maps of two apple cultivars (*Malus* × *domestica* Borkh.) based on AFLP and microsatellite markers. Mol Breed 15:205-219

Khatun S, Flowers T (1995) The estimation of pollen viability in rice. J Exp Bot 46:151-154

Kim CY, Ahn YO, Kim SH, Kim Y, Lee H, Catanach AS, Jacobs JM, Conner AJ, Kwak S (2010) The sweet potato IbMYB1 gene as a potential visible marker for sweet potato intragenic vector system. Physiol Plantarum 139:229-240

Kitamura S (2006) Transport of flavonoids: from cytosolic synthesis to vacuolar accumulation. In: The science of flavonoids. Springer, pp. 123-146

Klee H, Giovannoni J (2011) Genetics and control of tomato fruit ripening and quality attributes. Annu Rev Genet 45: 41-59

Klempnauer K, Bonifer C, Sippel AE (1986) Identification and characterization of the protein encoded by the human *c-myb* proto-oncogene. EMBO J 5:1903-1911

Klempnauer K, Gonda TJ, Michael Bishop J (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

Kobayashi S, Goto-Yamamoto N, Hirochika H (2004) Retrotransposon-Induced Mutations in Grape Skin Color. Science 304: 982–982

Kobayashi S, Ishimaru M, Hiraoka K, Honda C (2002) Myb-related genes of the Kyoho grape (*Vitis* l*abruscana*) regulate anthocyanin biosynthesis. Planta 215:924-933

Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci 10:236-242

Koltunow AM, Truettner J, Cox KH, Wallroth M, Goldberg RB (1990) Different temporal and spatial gene expression patterns occur during anther development. Plant Cell 2:1201-1224

Kong J, Chia L, Goh N, Chia T, Brouillard R (2003) Analysis and biological activities of anthocyanins. Phytochemistry 64:923-933

Koyama T, Okada T, Kitajima S, Ohme-akagi M, Shinshi H, Sato F (2003) Isolation of tobacco ubiquitin-conjugating enzyme cDNA in a yeast two-hybrid system with tobacco ERF3 as bait and its characterization of specific interaction. J Exp Bot 54:1175-1181

Kubasek WL, Shirley BW, McKillop A, Goodman HM, Briggs W, Ausubel FM (1992) Regulation of flavonoid biosynthetic genes in germinating *Arabidopsis* seedlings. The Plant Cell 4:1229-1236

Lambert P, Hagen L, Arus P, Audergon J (2004) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas× peach Earlygold reference map for *Prunus*. TAG Theoretical and Applied Genetics 108:1120-1130

Lancaster J, Dougall DK (1992) Regulation of skin color in apples. Crit Rev Plant Sci 10:487-502

Larkin JC, Walker JD, Bolognesi-Winfield AC, Gray JC, Walker AR (1999) Allele-specific interactions between *ttg* and *gl1* during trichome development in *Arabidopsis thaliana*. Genetics 151:1591-1604

Lee TI, Young RA (2000) Transcription of eukaryotic protein-coding genes. Annu Rev Genet 34:77-137

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

Lesley J (1940) A genetic study of saucer fruit shape and other characters in the peach 37:218-222

Li D, Roberts R (2001) Human Genome and Diseases: WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. Cellular and Molecular Life Sciences CMLS 58:2085-2097

Li J, Yang X, Wang Y, Li X, Gao Z, Pei M, Chen Z, Qu L, Gu H (2006) Two groups of MYB transcription factors share a motif which enhances trans-activation activity. Biochem Biophys Res Commun 341:1155-1163

Li L, Ban Z, Li X, Wu M, Wang A, Jiang Y, Jiang Y (2012) Differential Expression of Anthocyanin Biosynthetic Genes and Transcription Factor PcMYB10 in Pears (*Pyrus communis* L.). PloS one 7:e46070

Li L, Zhao Y, McCaig BC, Wingerd BA, Wang J, Whalon ME, Pichersky E, Howe GA (2004) The tomato homolog of CORONATINE-INSENSITIVE1 is required for the maternal control of seed maturation, jasmonate-signaled defense responses, and glandular trichome development. The Plant Cell 16:126-143

Li Y, Jiang J, Du M, Li L, Wang X, Li X (2013) A cotton gene encoding MYB-like transcription factor is specifically expressed in pollen and is involved in regulation of late anther/pollen development. Plant and Cell Physiology 54:893-906

Li ZT, Dhekney SA, Gray DJ (2011) Use of the *VvMybA1* gene for non-destructive quantification of promoter activity via color histogram analysis in grapevine (*Vitis vinifera*) and tobacco. Transgenic Res 20:1087-1097

Liebhard R, Koller B, Gianfranceschi L, Gessler C (2003) Creating a saturated reference map for the apple (*Malus* × *domestica* Borkh.) genome. TAG Theoretical and Applied Genetics 106:1497-1508

Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie T, Espley R, Hellens R, Allan A (2010) An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC plant biology 10:50

Lipsick JS (1996) One billion years of Myb. Oncogene 13:223-235

Liu G, Ren G, Guirgis A, Thornburg RW (2009) The MYB305 transcription factor regulates expression of nectarin genes in the ornamental tobacco floral nectary. The Plant Cell 21:2672-2687

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

Mandaokar A, Thines B, Shin B, Markus Lange B, Choi G, Koo YJ, Yoo YJ, Choi YD, Choi G (2006) Transcriptional regulators of stamen development in *Arabidopsis* identified by transcriptional profiling. The Plant Journal 46:984-1008

Mandaokar A, Browse J (2009) MYB108 Acts Together with MYB24 to Regulate Jasmonate-Mediated Stamen Maturation in *Arabidopsis*. Plant Physiol 149:851-862

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet 38:948-952

Mano H, Ogasawara F, Sato K, Higo H, Minobe Y (2007) Isolation of a regulatory gene of anthocyanin biosynthesis in tuberous roots of purple-fleshed sweet potato. Plant Physiol 143:1252-1268

Marrs KA, Alfenito MR, Lloyd AM, Walbot V (1995) A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. Nature 375:397-400

Martin C, Paz-Ares J (1997) MYB transcription factors in plants. Trends in Genetics 13:67-73

Masia A, Zanchin A, Rascio N, Ramina A (1992) Some biochemical and ultrastructural aspects of peach fruit development. J Am Soc Hort Sci 117:808-815

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

Matsui K, Umemura Y, Ohme‐Takagi M (2008) AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. The Plant Journal 55:954-967

Matsumoto H, Nakamura Y, Tachibanaki S, Kawamura S, Hirayama M (2003) Stimulatory effect of cyanidin 3-glycosides on the regeneration of rhodopsin. J Agric Food Chem 51:3560-3563

Matus J, Poupin M, Cañón P, Bordeu E, Alcalde J, Arce-Johnson P (2010) Isolation of WDR and bHLH genes related to flavonoid synthesis in grapevine (*Vitis vinifera* L.). Plant Mol Biol 72:607-620

Mazza G, Kay CD, Cottrell T, Holub BJ (2002) Absorption of anthocyanins from blueberries and serum antioxidant status in human subjects. J Agric Food Chem 50:7731-7737

Mehrtens F, Kranz H, Bednarek P, Weisshaar B (2005) The *Arabidopsis* transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. Plant Physiol 138:1083-1096

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

Montiel S, Montaña C (2000) Vertebrate frugivory and seed dispersal of a Chihuahuan Desert cactus. Plant Ecol 146:219-227

Moore I, Samalova M, Kurup S (2006) Transactivated and chemically inducible gene expression in plants. The Plant Journal 45:651-683

Moore I, Gälweiler L, Grosskopf D, Schell J, Palme K (1998) A transcription activation system for regulated gene expression in transgenic plants. Proceedings of the National Academy of Sciences 95:376-381

Morita Y, Saitoh M, Hoshino A, Nitasaka E, Iida S (2006) 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 47:457-470

Mosco A (2012) Tissue localization of betacyanins in cactus stems. Revista Mexicana de Biodiversidad 83: 413-420

Mueller LA, Goodman CD, Silady RA, Walbot V (2000) AN9, a *Petunia* glutathione S-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. Plant Physiol 123:1561-1570

Muller PY, Janovjak H, Miserez AR, Dobbie Z (2002) Processing of gene expression data generated by quantitative real-time RT-PCR. BioTechniques 32:1372

Murre C, Bain G, van Dijk MA, Engel I, Furnari BA, Massari ME, Matthews JR, Quong MW, Rivera RR, Stuiver MH (1994) Structure and function of helix-loop-helix proteins. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression 1218:129-135

Murre C, McCaw PS, Baltimore D (1989) A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, *MyoD*, and *myc* proteins. Cell 56:777-783

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

Norton JD (1966) Testing of plum pollen viability with tetrazolium salts 89:132-134

Ogata K, Morikawa S, Nakamura H, Hojo H, Yoshimura S, Zhang R, Aimoto S, Ametani Y, Hirata Z, Sarai A (1995) Comparison of the free and DNA-

complexed forms of the DMA-binding domain from c-Myb. Nature Structural & Molecular Biology 2:309-320

Ogundiwin E, Peace C, Gradziel T, Parfitt D, Bliss F, Crisosto C (2009) A fruit quality gene map of *Prunus*. BMC Genomics 10:587

Ohno S, Hosokawa M, Hoshino A, Kitamura Y, Morita Y, Park K, Nakashima A, Deguchi A, Tatsuzawa F, Doi M (2011) A bHLH transcription factor, DvIVS, is involved in regulation of anthocyanin synthesis in dahlia (*Dahlia variabilis*). J Exp Bot 62:5105-5116

Parker D, Zilberman D, Moulton K (1991) How quality relates to price in California fresh peaches. Calif Agric 45:14-16

Pattanaik S, Kong Q, Zaitlin D, Werkman JR, Xie CH, Patra B, Yuan L (2010) Isolation and functional characterization of a floral tissue-specific R2R3 MYB regulator from tobacco. Planta 231:1061-1076

Pauwels L, Goossens A (2011) The JAZ proteins: a crucial interface in the jasmonate signaling cascade. The Plant Cell 23:3089-3100

Payne CT, Zhang F, Lloyd AM (2000) GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. Genetics 156:1349-1362

Payne T, Clement J, Arnold D, Lloyd A (1999) Heterologous *myb* genes distinct from *GL1* enhance trichome production when overexpressed in *Nicotiana tabacum*. Development 126:671-682

Paz-Ares J, Ghosal D, Wienand U, Peterson P, 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

Pedreno M, Escribano J (2000) Studying the oxidation and the antiradical activity of betalain from beetroot. J Biol Educ 35:49-51

Perin C, Hagen L, De Conto V, Katzir N, Danin-Poleg Y, Portnoy V, Baudracco-Arnas S, Chadoeuf J, Dogimont C, Pitrat M (2002) A reference map of *Cucumis melo* based on two recombinant inbred line populations. TAG Theoretical and Applied Genetics 104:1017-1034

Petroni K, Tonelli C (2011) Recent advances on the regulation of anthocyanin synthesis in reproductive organs. Plant Science 181:219-229

Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 424:85-88

Qi T, Song S, Ren Q, Wu D, Huang H, Chen Y, Fan M, Peng W, Ren C, Xie D (2011) The Jasmonate-ZIM-domain proteins interact with the WD-Repeat/bHLH/MYB complexes to regulate jasmonate-mediated anthocyanin accumulation and trichome initiation in *Arabidopsis thaliana*. The Plant Cell 23:1795-1814

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

Quilot B, Wu B, Kervella J, Genard M, Foulongne M, Moreau K (2004) QTL analysis of quality traits in an advanced backcross between *Prunus persica* cultivars and the wild relative species *P. davidiana*. TAG Theoretical and Applied Genetics 109:884-897

Ramsay NA, Glover BJ (2005) MYB–bHLH–WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci 10:63-70

Rao A, Rao LG (2007) Carotenoids and human health. Pharmacological Research 55:207-216

Ravaglia D, Espley RV, Henry-Kirk RA, Andreotti C, Ziosi V, Hellens RP, Costa G, Allan AC (2013) Transcriptional regulation of flavonoid biosynthesis in nectarine (*Prunus persica*) by a set of R2R3 MYB transcription factors. BMC Plant Biology 13:68

Rehder A (1949) Manual of Cultivated Trees and Shrubs Hardy in North America: Exclusive of the Subtropical and Warmer Temperate Regions. Macmillan

Riechmann JL, Ratcliffe OJ (2000) A genomic perspective on plant transcription factors. Curr Opin Plant Biol 3:423-434

Rodriguez-Riano T, Dafni A (2000) A new procedure to asses pollen viability. Sexual Plant Reproduction 12:241-244

Rutherford S, Brandizzi F, Townley H, Craft J, Wang Y, Jepson I, Martinez A, Moore I (2005) Improved transcriptional activators and their use in mis-expression traps in *Arabidopsis*. The Plant Journal 43:769-788

Sablowski R, Moyano E, Culianez-Macia FA, Schuch W, Martin C, Bevan M (1994) A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. EMBO J 13:128

Salvatierra A, Pimentel P, Moya-León MA, Herrera R (2013) Increased accumulation of anthocyanins in *Fragaria chiloensis* fruits by transient suppression of *FcMYB1* gene. Phytochemistry 90:25-36

Sargent DJ, Clarke J, Simpson D, Tobutt K, Arus P, Monfort A, Vilanova S, Denoyes-Rothan B, Rousseau M, Folta K (2006) An enhanced microsatellite

map of diploid *Fragaria*. TAG Theoretical and Applied Genetics 112:1349-1359

Schaart JG, Dubos C, Romero De La Fuente, I., Houwelingen AMML, Vos RCH, Jonker HH, Xu W, Routaboul JM, Lepiniec L, Bovy AG (2013) Identification and characterization of MYB‐bHLH‐WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria* × *ananassa*) fruits. New Phytol 197:454-467

Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C (2006) A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. The Plant Cell 18:831-851

Serna L, Martin C (2006) Trichomes: different regulatory networks lead to convergent structures. Trends Plant Sci 11:274-280

Sharma A, Zhang L, Ashrafi H, Foolad MR (2009) A *Solanum lycopersicum*× *Solanum pimpinellifolium* linkage map of tomato displaying genomic locations of R-genes, RGAs, and candidate resistance/defense-response ESTs. International journal of plant genomics 2008

Shen Z, Confolent C, Lambert P, Poëssel J, Quilot-Turion B, Yu M, Ma R, Pascal T (2013) Characterization and genetic mapping of a new blood-flesh trait controlled by the single dominant locus *DBF* in peach. Tree Genetics & Genomes 9:1435-1446

Shin B, Choi G, Yi H, Yang S, Cho I, Kim J, Lee S, Paek N, Kim J, Song P (2002) AtMYB21, a gene encoding a flower‐specific transcription factor, is regulated by COP1. The Plant Journal 30:23-32

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arus P, Dandekar AM (2008) Multiple models for Rosaceae genomics. Plant Physiol 147:985-1003

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP (2010) The genome of woodland strawberry (*Fragaria vesca*). Nat Genet 43:109-116

Singh M, Arseneault M, Sanderson T, Murthy V, Ramassamy C (2008) Challenges for research on polyphenols from foods in Alzheimer's disease: bioavailability, metabolism, and cellular and molecular mechanisms. J Agric Food Chem 56:4855-4873

Smith TF, Gaitatzes C, Saxena K, Neer EJ (1999) The WD repeat: a common architecture for diverse functions. Trends Biochem Sci 24:181-185

Sompornpailin K, Makita Y, Yamazaki M, Saito K (2002) A WD-repeat-containing putative regulatory protein in anthocyanin biosynthesis in *Perilla frutescens*. Plant Mol Biol 50:485-495

Song S, Qi T, Huang H, Ren Q, Wu D, Chang C, Peng W, Liu Y, Peng J, Xie D (2011) The jasmonate-ZIM domain proteins interact with the R2R3-MYB transcription factors MYB21 and MYB24 to affect jasmonate-regulated stamen development in *Arabidopsis*. The Plant Cell 23:1000-1013

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C (1994) Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol Biol 24:743-755

Spelt C, Quattrocchio F, Mol JN, Koes R (2000) anthocyanin1 of *Petunia* encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. The Plant Cell 12:1619-1631

Spolaore S, Trainotti L, Casadoro G (2001) A simple protocol for transient gene expression in ripe fleshy fruit mediated by Agrobacterium. J Exp Bot 52:845-850

Springob K, Nakajima J, Yamazaki M, Saito K (2003) Recent advances in the biosynthesis and accumulation of anthocyanins. Nat Prod Rep 20:288-303

Stafford HA (1994) Anthocyanins and betalains: evolution of the mutually exclusive pathways. Plant Science 101:91-98

Stevens PF (2001) Angiosperm Phylogeny Website. Consultado el día, 8 (URL: http://www.mobot.org/MOBOT/research/APweb)

Stirnimann CU, Petsalaki E, Russell RB, Müller CW (2010) WD40 proteins propel cellular networks. Trends Biochem Sci 35:565-574

Stockinger E, Mulinix C, Long C, Brettin T (1996) A linkage map of sweet cherry based on RAPD analysis of a microspore-derived callus culture population. J Hered 87:214-218

Strack D, Vogt T, Schliemann W (2003) Recent advances in betalain research. Phytochemistry 62:247-269

Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, Weisshaar B (2007) Differential regulation of closely related R2R3‐MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. The Plant Journal 50:660-677

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

Sun Y, Li H, Huang J (2012) *Arabidopsis* TT19 functions as a carrier to transport anthocyanin from the cytosol to tonoplasts. Molecular plant 5:387-400

Takhtadzhian AL (1997) Diversity and the Classification of Flowering Plants. Columbia University Press

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

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731-2739

Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. The Plant Journal 54:733-749

Taylor LP, Grotewold E (2005) Flavonoids as developmental regulators. Curr Opin Plant Biol 8:317-323

Tedesco I, Luigi Russo G, Nazzaro F, Russo M, Palumbo R (2001) Antioxidant effect of red wine anthocyanins in normal and catalase-inactive human erythrocytes. J Nutr Biochem 12:505-511

Thines B, Katsir L, Melotto M, Niu Y, Mandaokar A, Liu G, Nomura K, He SY, Howe GA (2007) JAZ repressor proteins are targets of the SCFCOI1 complex during jasmonate signalling. Nature 448:661-665

Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M (2005) Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. The Plant Journal 42:218-235

Toledo-Ortiz G, Huq E, Quail PH (2003) The *Arabidopsis* basic/helix-loop-helix transcription factor family. The Plant Cell 15:1749-1770

Tomás-Barberán FA, Gil MI, Cremin P, Waterhouse AL, Hess-Pierce B, Kader AA (2001) HPLC-DAD-ESIMS analysis of phenolic compounds in nectarines, peaches, and plums. J Agric Food Chem 49:4748-4760

Tombuloglu H, Kekec G, Sakcali MS, Unver T (2013) Transcriptome-wide identification of R2R3-MYB transcription factors in barley with their boron responsive expression analysis. Molecular Genetics and Genomics :1-15

Trainotti L, Bonghi C, Ziliotto F, Zanin D, Rasori A, Casadoro G, Ramina A, Tonutti P (2006) The use of microarray μPEACH1.0 to investigate transcriptome changes during transition from pre-climacteric to climacteric phase in peach fruit. Plant Science 170:606-613

Trainotti L, Cagnin S, Forcato C, Bonghi C, Dhingra A, Koepke T, Prat L, Maldonado J, Silva H (2012) Functional Genomics: Transcriptomics. Genet.

Genomics Breed. Stone Fruits - CRC Press Book, 1st ed. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487-2742, pp 292–322

Trainotti L, Tadiello A, Casadoro G (2007) The involvement of auxin in the ripening of climacteric fruits comes of age: the hormone plays a role of its own and has an intense interplay with ethylene in ripening peaches. J Exp Bot 58:3299

Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, Stefanini M, Mank R, Moroldo M, Morgante M (2007) A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics 176:2637-2650

Tsuda T, Horio F, Uchida K, Aoki H, Osawa T (2003) Dietary cyanidin 3-O-β-D-glucoside-rich purple corn color prevents obesity and ameliorates hyperglycemia in mice. J Nutr 133:2125-2130

Tsuda T, Yamaguchi M, Honda C, Moriguchi T (2004) Expression of anthocyanin biosynthesis genes in the skin of peach and nectarine fruit. J Am Soc Hort Sci 129:857-862

van Nocker S, Ludwig P (2003) The WD-repeat protein superfamily in *Arabidopsis*: conservation and divergence in structure and function. BMC Genomics 4:50

Vancanneyt G, Schmidt R, O'Connor-Sanchez A, Willmitzer L, Rocha-Sosa M (1990) Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in *Agrobacterium*-mediated plant transformation. Molecular and General Genetics MGG 220:245-250

Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet 45:487-494

Vizzotto G, Pinton R, Varanini Z, Costa G (1996) Sucrose accumulation in developing peach fruit. Physiol Plant 96: 225–230.

Voorrips RE, (2002) MapChart version 2.2: software for the graphical presentation of linkage maps and QTLs. J Hered 93: 77–78

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J (2002) A MADS-box gene necessary for fruit ripening at the tomato *ripening-inhibitor* (*rin*) locus. Science 296:343-346

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC (1999) The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. The Plant Cell 11:1337-1349

Wang S, Barron C, Schiefelbein J, Chen J (2010) Distinct relationships between GLABRA2 and single‐repeat R3 MYB transcription factors in the regulation of trichome and root hair patterning in *Arabidopsis*. New Phytol 185:387-400

Wang Z, Ge Y, Scott M, Spangenberg G (2004) Viability and longevity of pollen from transgenic and nontransgenic tall fescue (*Festuca arundinacea*) (Poaceae) plants. Am J Bot 91:523-530

Wargovich MJ (2000) Anticancer properties of fruits and vegetables: The Role of Oxidative Stress and Antioxidants in Plant and Human Health. HortScience 35:573-575

Welch CR, Wu Q, Simon JE (2008) Recent advances in anthocyanin analysis and characterization. Current analytical chemistry 4:75

Werner DJ, Creller MA, Chaparro JX (1998) Inheritance of the blood-flesh trait in peach. HortScience 33:1243-1246

Williams CE, Grotewold E (1997) Differences between plant and animal Myb domains are fundamental for DNA binding activity, and chimeric Myb domains have novel DNA binding specificities. J Biol Chem 272:563-571

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

Xia M, Ling W, Zhu H, Wang Q, Ma J, Hou M, Tang Z, Li L, Ye Q (2007) Anthocyanin prevents CD40-activated proinflammatory signaling in endothelial cells by regulating cholesterol distribution. Arterioscler Thromb Vasc Biol 27:519-524

Yamamoto T, Kimura T, Shoda M, Imai T, Saito T, Sawamura Y, Kotobuki K, Hayashi T, Matsuta N (2002) Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. TAG Theoretical and Applied Genetics 106:9-18

Yamamoto T, Yamaguchi M, Hayashi T (2005) An integrated genetic linkage map of peach by SSR, STS, AFLP and RAPD. J Jpn Soc Hort Sci 74:204-213

Yang X, Li J, Pei M, Gu H, Chen Z, Qu L (2007) Over-expression of a flower-specific transcription factor gene AtMYB24 causes aberrant anther development. Plant Cell Rep 26:219-228

Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q (2006) The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol Biol 60:107-124

Zanchin A, Bonghi C, Casadoro G, Ramina A, Rascio N (1994) Cell enlargement and cell separation during peach fruit development. Int J Plant Sci :49-56

Zawel L, Reinberg D (1995) Common themes in assembly and function of eukaryotic transcription complexes. Annu Rev Biochem 64:533-561

Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. Development 130:4859-4869

Zhang H, Jin J, Tang L, Zhao Y, Gu X, Gao G, Luo J (2011) PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. Nucleic Acids Res 39:D1114-D1117

Zhao L, Gao L, Wang H, Chen X, Wang Y, Yang H, Wei C, Wan X, Xia T (2013) The R2R3-MYB, bHLH, WD40, and related transcription factors in flavonoid biosynthesis. Funct Integr Genomics 13:75-98

Zhu H, Fitzsimmons K, Khandelwal A, Kranz RG (2009) CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in Arabidopsis. Molecular Plant 2:790-802

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

Zuo J, Zhu B, Fu D, Zhu Y, Ma Y, Chi L, Ju Z, Wang Y, Zhai B, Luo Y (2012) Sculpting the maturation, softening and ethylene pathway: The influences of microRNAs on tomato fruits. BMC Genomics 13:7

7 Appendices

Appendix I: Sequence alignment for *MYB10s* and *bHLHs* with the reference sequence of the peach genome (www.phytozome.net).

Coding sequence of MYB10.1

```
ppa026640m      ATGGAGGGCTATAACTTGGGTGTGAGAAAAGGAGCTTGGACTAGAGAGGAAGATGATCTT 60
MYB10.1         ATGGAGGGCTATAACTTGGGTGTGAGAAAAGGAGCTTGGACTAGAGAGGAAGATGATCTT
                ************************************************************

ppa026640m      TTGAGGCAGTGCATTGAGAATCATGGAGAAGGAAAGTGGCACCAAGTTCCTAACAAAGCA 120
MYB10.1         TTGAGGCAGTGCATTGAGAATCATGGAGAAGGAAAGTGGCACCAAGTTCCTAACAAAGCA
                ************************************************************

ppa026640m      GGGTTGAACAGGTGCAGGAAGAGCTGTAGACTAAGGTGGATGAACTATTTGAAGCCAAAT 180
MYB10.1         GGGTTGAACAGGTGCAGGAAGAGCTGTAGACTAAGGTGGATGAACTATTTGAAGCCAAAT
                ************************************************************

ppa026640m      ATCAAGAGAGGAGAGTTTGCAGAGGATGAAGTAGATCTAATCATTAGGCTTCACAAGCTT 240
MYB10.1         ATCAAGAGAGGAGAGTTTGCAGAGGATGAAGTAGATCTAATCATTAGGCTTCACAAGCTT
                ************************************************************

ppa026640m      TTAGGAAACAGGTGGTCATTGATTGCTGGAAGGCTTCCAGGAAGGACAGCGAATGATGTG 300
MYB10.1         TTAGGAAACAGGTGGTCATTGATTGCTGGAAGGCTTCCAGGAAGGACAGCGAATGATGTG
                ************************************************************

ppa026640m      AAAAATTATTGGAACACTCGACTGCGGACGGATTCTCGCCTGAAAAAGGTGAAAGATAAA 360
MYB10.1         AAAAATTATTGGAACACTCGACTGCGGACGGATTCTCGCCTGAAAAGGGTGAAAGATAAA
                ************************************************** **********

ppa026640m      CCCCAAGAAACAATAAAGACCATCGTAATAAGACCTCAACCCCGAAGCTTCATCAAGAGT 420
MYB10.1         CCCCAAGAAACAATAAAGACCATCGTAATAAGACCTCAACCCCGAAGCTTCATCAAGAGT
                ************************************************************

ppa026640m      TCAAATTGTTTGAGCAGTAAAGAACCAATTTTGGATCATATTCAAACAGTCGAGAATTTT 480
MYB10.1         TCAAATTGTTTGAGCAGTAAAGAACCAATTTTGGATCATATTCAAACAGTCGAGAATTTT
                ************************************************************

ppa026640m      AGTACGCCGTCACAAACATCACCATCAACAAAGAATGGAAATGATTGGTGGGAAACCTTT 540
MYB10.1         AGTACGCCGTCACAAACATCACCATCAACAAAGAATGGAAATGATTGGTGGGAAACCTTT
                ************************************************************

ppa026640m      TTAGATGACGAGGATGTTTTTGAAAGAGCTACATGCTATGGTCTAGCATTAGAGGAAGAA 600
MYB10.1         TTAGATGACGAGGATGTTTTTGAAAGAGCTACATGCTATGGTCTAGCATTAGAGGAAGAA
                ************************************************************

ppa026640m      GAGTTCACAAGTTTTTGGGTTGATGATATGCCACAATCGAAAAGACAGTGTACCAATGTT 660
MYB10.1         GAGTTCACAAGTTTTTGGGTTGATGATATGCCACAATCGAAAAGACAGTGTACCAATGTT
                ************************************************************

ppa026640m      TCAGAAGAAGGACTAGGTAGAGGTGATTTCTCTTTTAACGTGGACTTTTGGAATCATTAA 720
MYB10.1         TCAGAAGAAGGACTAGGTAGAGGTGATTTCTCTTTTAACGTGGACTTTTGGAATCATTAA
                ************************************************************
```

Amino acid sequence of MYB10.1

```
ppa026640m      MEGYNLGVRKGAWTREEDDLLRQCIENHGEGKWHQVPNKAGLNRCRKSCRLRWMNYLKPN
MYB10.1         MEGYNLGVRKGAWTREEDDLLRQCIENHGEGKWHQVPNKAGLNRCRKSCRLRWMNYLKPN
                ************************************************************

ppa026640m      IKRGEFAEDEVDLIIRLHKLLGNRWSLIAGRLPGRTANDVKNYWNTRLRTDSRLKKVKDK
MYB10.1         IKRGEFAEDEVDLIIRLHKLLGNRWSLIAGRLPGRTANDVKNYWNTRLRTDSRLKRVKDK
                ************************************************************:***

ppa026640m      PQETIKTIVIRPQPRSFIKSSNCLSSKEPILDHIQTVENFSTPSQTSPSTKNGNDWWETF
MYB10.1         PQETIKTIVIRPQPRSFIKSSNCLSSKEPILDHIQTVENFSTPSQTSPSTKNGNDWWETF
                ************************************************************

ppa026640m      LDDEDVFERATCYGLALEEEEFTSFWVDDMPQSKRQCTNVSEEGLGRGDFSFNVDFWNH
MYB10.1         LDDEDVFERATCYGLALEEEEFTSFWVDDMPQSKRQCTNVSEEGLGRGDFSFNVDFWNH
                ************************************************************
```

Coding sequence of MYB10.2

```
ppa016711m      ATGGAGGGTTATGACTTGAGTGTGAGAAAAGGAGCTTGGACTAGAGAGGAAGATGATCTT 60
MYB10.2         ATGGAGGGTTATGACTTGAGTGTGAGAAAAGGAGCTTGGACTAGAGAGGAAGATGATCTT
                ************************************************************

ppa016711m      CTGAGGCAGTGCATTGAGAATCAAGGTGAAGGAAAGTGGCACCAGGTTCCTTACAAAGCA 120
MYB10.2         CTGAGGCAGTGTATTGAGAATCAAGGTGAAGGAAAGTGGCACCAGGTTCCTTGCAAAGCA
                *********** ********************************************* *******

ppa016711m      GGGTTGAACAGGTGCAGGAAGAGCTGTAGACTAAGGTGGGTGAACTATTTGAAGCCAAAT 180
MYB10.2         GGGTTGAACAGGTGCAGGAAGAGCTGTAGACTAAGGTGGGTGAACTATTTGAAGCCAAAT
                ************************************************************

ppa016711m      ATCAAGAAAGGAGAGTTTGCAGAGGATGAAGTAGATCTAATAATTAGGCTTCACAAGCTT 240
MYB10.2         ATCAAGAAAGGAGAGTTTGCAGAGGATGAAGTAGATCTAATAATTAGGCTTCACAAGCTT
                ************************************************************

ppa016711m      CTAGGAAACAGGTGGTCATTGATTGCTCGAAGACTTCCAGGCAGGACTGCGAATAATGTG 300
MYB10.2         CTAGGAAACAGGTGGTCATTGATTGCTCGAAGACTTCCAGGAAGGACTGCGAATAATGTG
                ******************************************* ******************

ppa016711m      AAAAATTACTGGAACACCCGATCGCGGACGGATTATTGCATGAAAAAGATAAAAGACAAA 360
MYB10.2         AAAAATTACTGGAACACCCGATCGCGGACGGATTATTGCATGAAAAAGATGAAAGACAAA
                ****************************************************** **********

ppa016711m      TCCCAAGAAACAATAAAGACCATAATAAGGCCACAACCAAGAAGATTCACCAAAAGTTCA 420
MYB10.2         TCCCAAGAAACAATAAAGACCATAATAAGGCCACAACCAAGAAGATTCACCAAAAGTTCA
                ************************************************************

ppa016711m      AATTGTTTGAGTTTTAAAGAACCAATTTTGGACCATACTCAACGTGATTGGTGGGAGACC 480
MYB10.2         AATTGTTTGAGTTTTAAAGAACCAATTTTGGACCATACTCAACGTGATTGGTGGGAGACC
                ************************************************************

ppa016711m      TTTTTAGATGACAAGGATGCTACTGAAAGAGCTACAGGTTCTGGTCTTGGGTTAGATGAA 540
MYB10.2         TTTTTAGATGACAAGGATGCTACTGAAAGAGCTACAGGTTCTGGTCTTGGGTTAGATGAA
                ************************************************************

ppa016711m      GAACTGCTCGCAAGTTTTTGGGTTGATGATGATATGCCACAATCGACAAGAAAATGCATC 600
MYB10.2         GAACTGCTCGCAAGTTTTTGGGTTGATGATGATATGCCACAATCGACAAGAAAATGCATC
                ************************************************************

ppa016711m      AATTTTTCTGAAGGACTAATTAGAGGTGATTTCTCTTTTAGCGTGGACCCTTGGAATCAT 660
MYB10.2         AATTTTTCTGAAGGACTAATTAGAGGTGATTTCTCTTTTAGCGTGGACCCTTGGAATCAT
                ************************************************************

ppa016711m      TCAAATGAAGAATATAGAAAGTAA 684
MYB10.2         TCAAATGAAGAATATAGAAAGTAA
                ************************
```

Amino acid sequence of MYB10.2

```
ppa016711m      MEGYDLSVRKGAWTREEDDLLRQCIENQGEGKWHQVPYKAGLNRCRKSCRLRWVNYLKPN
MYB10.2         MEGYDLSVRKGAWTREEDDLLRQCIENQGEGKWHQVPCKAGLNRCRKSCRLRWVNYLKPN
                ************************************ ***********************

ppa016711m      IKKGEFAEDEVDLIIRLHKLLGNRWSLIARRLPGRTANNVKNYWNTRSRTDYCMKKIKDK
MYB10.2         IKKGEFAEDEVDLIIRLHKLLGNRWSLIARRLPGRTANNVKNYWNTRSRTDYCMKKMKDK
                ************************************************************:***

ppa016711m      SQETIKTIIRPQPRRFTKSSNCLSFKEPILDHTQRDWWETFLDDKDATERATGSGLGLDE
MYB10.2         SQETIKTIIRPQPRRFTKSSNCLSFKEPILDHTQRDWWETFLDDKDATERATGSGLGLDE
                ************************************************************

ppa016711m      ELLASFWVDDDMPQSTRKCINFSEGLIRGDFSFSVDPWNHSNEEYRK
MYB10.2         ELLASFWVDDDMPQSTRKCINFSEGLIRGDFSFSVDPWNHSNEEYRK
                ************************************************
```

Coding sequence of MYB10.3

```
ppa020385m      ATGGGGGGAAATAACTTGGATGTGAAAAAAGGAGCTTGGACTAAAGAGGAAGATGCTCTT 60
MYB10.3         ATGGGGGGAAATAACTTGGATGTGAAAAAAGGAGCTTGGACTAAAGAGGAAGATGCTCTT
                ************************************************************

ppa020385m      CTGAGCAAGTGCATGGAGAATCATGGAGAAGGAAAGTGGCACGAGGTTCCTTACAAAGCA 120
MYB10.3         CTGAGCAAGTGCATGGAGAATCATGGAGAAGGAAAGTGGCACGAGGTTCCTTACAAAGCA
                ************************************************************

ppa020385m      GGCTTAAACAGATGCAGGAAGAGCTGTAGACTAAGGTGGTTGAACTATGTGAAGCCAAAT 180
MYB10.3         GGCTTAAACAGATGCAGGAAGAGCTGTAGACTAAGGTGGTTGAACTATTTGAAGCCAAAT
                ************************************************** ************

ppa020385m      ATCAAGAGAGGAGAGTTTACAGAGGATGAAGTAGATCTAATAATTAGGCTTCACAAGCTT 240
MYB10.3         ATCAAGAGAGGAGAGTTTACAGAGGATGAAATAGATCTAATAATTAGGCTTCACAAGCTT
                ****************************** ******************************

ppa020385m      TTAGGAAACAGGTGGGCACTGATTGCTGGAAGACTTCCAGGAAGGACATCGAACGGTGTG 300
MYB10.3         TTAGGAAACAGGTGGGCACTGATTGCTGGAAGACTTCCAGGAAGGACATCGAACGGTGTG
                ************************************************************

ppa020385m      AAAAATTATTGGAACACCCGACGGCGGACGAATTCTCTCCTGAAAACGACTACGAAAGAA 360
MYB10.3         AAAAATTATTGGAACACCCGACGGCGGACGAATTCTCTCCTGAAAACGACTACGAAAGAA
                ************************************************************

ppa020385m      AAATTCCAAGAAACAATAAAGCCCATCGTCACAAGGCCTCAACCGCGAAGTTTCACCAAA 420
MYB10.3         AAATTCCAAGAAACAATAAAGCCCATCGTCACAAGGCCTCAACCGCGAAGTTTCACCAAA
                ************************************************************

ppa020385m      AGTTCAAATTGTTCGAGTTTTGAAGAACCAGTTTTGGACCATACTCAACTAGAAGAAAAT 480
MYB10.3         AGTTCAAATTGTTCGAGTTTTGAAGAACCAGTTTTGGACCATACTCAACTAGAAGAAAAT
                ************************************************************

ppa020385m      TTTAGTACGCCATCACAAACATCACCATCAACAAGGATTGGAAATGATTGGTGGGATACC 540
MYB10.3         TTTAGTACGCCATCACAAACATCACCATCAACAAGGATTGGAAATGATTGGTGGGATACC
                ************************************************************

ppa020385m      TTTTTAGATGACAAGGATGCTACTGAAACAGCTACAGGTTCTGGTCCTGGGTTTGATGAA 600
MYB10.3         TTTTTAGATGACAAGGATGCTACTGGAACAGCTACAGGTTCTGGTCCTGGGTTTGATGAA
                ************************* ************************************

ppa020385m      GAACTGCTCACGAGTTTTTGGGTTGATGATGATATGCCACAATCGGCAAGAACATGCATC 660
MYB10.3         GAACTGCTCACGAGTTTTTGGGTTGATGATGATATGCCACAATCGGCAAGAACATGCATC
                ************************************************************

ppa020385m      AATTTTTCTGAAGAAGAACTGAGTATAAGTGATTTCTCTTTTAACTTGGACCTTTGGAAT 720
MYB10.3         AGTTTTTCTGAAGAAGAAGTGAGTATAAGTGATTTCTCTTTTAACTTGGACCTTTGGAAT
                * ***************** ****************************************

ppa020385m      CATTCAAAAGAAGAATAA 738
MYB10.3         CATTCAAAAGAAGAATAA
                ******************
```

Amino acid sequence of MYB10.3

```
ppa020385m      MGGNNLDVKKGAWTKEEDALLSKCMENHGEGKWHEVPYKAGLNRCRKSCRLRWLNYVKPN
MYB10.3         MGGNNLDVKKGAWTKEEDALLSKCMENHGEGKWHEVPYKAGLNRCRKSCRLRWLNYLKPN
                ************************************************************:***

ppa020385m      IKRGEFTEDEVDLIIRLHKLLGNRWALIAGRLPGRTSNGVKNYWNTRRRTNSLLKTTTKE
MYB10.3         IKRGEFTEDEIDLIIRLHKLLGNRWALIAGRLPGRTSNGVKNYWNTRRRTNSLLKTTTKE
                **********:*************************************************

ppa020385m      KFQETIKPIVTRPQPRSFTKSSNCSSFEEPVLDHTQLEENFSTPSQTSPSTRIGNDWWDT
MYB10.3         KFQETIKPIVTRPQPRSFTKSSNCSSFEEPVLDHTQLEENFSTPSQTSPSTRIGNDWWDT
                ************************************************************

ppa020385m      FLDDKDATETATGSGPGFDEELLTSFWVDDDMPQSARTCINFSEEELSISDFSFNLDLWN
MYB10.3         FLDDKDATGTATGSGPGFDEELLTSFWVDDDMPQSARTCISFSEEEVSISDFSFNLDLWN
                ******** ********************************.*****:**************

ppa020385m      HSKEE
MYB10.3         HSKEE
                *****
```

Coding sequence of bHLH3

```
ppa002884m      ATGGCTGCACCGCCAAGTAGCAGCAGGCTCCGCAGCATGTTGCAGGCGTCAGTCCAATCT 60
bHLH3           ATGGCTGCACCGCCAAGTAGCAGCAGGCTCCGCAGCATGTTGCAGGCGTCAGTCCAATCT
                ************************************************************

ppa002884m      GTCCAATGGACTTACAGTCTCTTCTGGCAAATCTGTCCACAACAAGGGATCTTAGTATGG 120
bHLH3           GTCCAATGGACTTACAGTCTCTTCTGGCAAATCTGTCCACAACAAGGGATCTTAGTATGG
                ************************************************************

ppa002884m      TCAGATGGGTACTACAATGGAGCCATCAAGACGAGGAAGACGGTGCAACCCATGGAAGTC 180
bHLH3           TCAGATGGGTACTACAATGGAGCCATCAAGACGAGGAAGACGGTGCAACCCATGGAAGTC
                ************************************************************

ppa002884m      AGTGCAGAAGAGGCCTCTCTTCAGCGGAGCCAGCAGCTTAGAGAACTGTATGACTCATTG 240
bHLH3           AGTGCAGAAGAGGCCTCTCTTCAGCGGAGCCAGCAGCTTAGAGAACTGTATGACTCATTG
                ************************************************************

ppa002884m      TCTGCCGGAGAAACTAACCAGCCCCCGGCACGCCGCCCTTGCGCCTCCTTATCCCCCGAG 300
bHLH3           TCTGCCGGAGAAACTAACCAGCCCCCGGCACGCCGCCCTTGCGCCTCCTTATCCCCCGAG
                ************************************************************

ppa002884m      GACTTAACCGAATCCGAATGGTTCTACTTGATGTGTGTCTCCTTCTCTTTCCCCCCTGGG 360
bHLH3           GACTTAACCGAATCCGAATGGTTCTACTTGATGTGTGTCTCCTTCTCTTTCCCCCCTGGG
                ************************************************************

ppa002884m      GTGGGGTTGCCAGGAAAAGCATATGCGAGGAGGCAGCATGTATGGCTCACCGGTGCAAAC 420
bHLH3           GTGGGGTTGCCAGGAAAAGCATATGCGAGGAGGCAGCATGTATGGCTCACCGGTGCAAAC
                ************************************************************

ppa002884m      GAAGTCGATAGCAAAACCTTTTCCAGAGCTATCTTAGCCAAGAGTGCTCGTATACAGACG 480
bHLH3           GAAGTCGATAGCAAAACCTTTTCCAGAGCTATCTTAGCCAAGAGTGCTCGTATACAGACG
                ************************************************************

ppa002884m      GTGGTTTGCATTCCTCTACTAGACGGCGTCGTCGAACTAGGCACCATAGAGAGGG----- 540
bHLH3           GTGGTTTGCATTCCTCTACTAGACGGCGTCGTCGAACTAGGCACCATAGAGAGGGTTCCG
                ************************************************************

ppa002884m      ------------------------------------------------------------ 600
bHLH3           GAAGACAATGCCTTGGTCCAACATGTCAAAACCTCCTTCGTTGACCACAACCACCATCCT

ppa002884m      ------------------------------------------------------------ 660
bHLH3           CCTCCAAAACCCGCCCTCTCCGAGCACTCAACCTCCAACCCCGCCACGTCATCCGACCAC

ppa002884m      ------------------------------------------------------------ 720
bHLH3           CCACATTTCCACTCTCCGCACCTCCCCGCCATGTGCCCCGACCCACCTCTCGACGCCGCC

ppa002884m      ----------------------------------------------------------CAGAG 780
bHLH3           CAAGAAGACGAAGAAGAGGACGAGGAAGAGGAGGATCAAGAGGAGGACGAGGAGGCAGAG
                                                                      *****

ppa002884m      TCAGACTCCGAAGCCGAAACGGGCCGCAACGGTCAGCCCGTTGCTCCCGGTCCAAACCCT 840
bHLH3           TCAGACTCCGAAGCCGAAACGGGCCGCAACGGTCAGCCCGTTGCTCCCGGTCCAAACCCT
                ************************************************************

ppa002884m      CCAGCGGTTTCGGCCCCCGCAGCAGAGCCGAGCGAGCTCATGCAGCTAGAGATGTCCGAG 900
bHLH3           CCAGCGGTTTCGGCCCCCGCAGCAGAGCCGAGCGAGCTCATGCAGCTAGAGATGTCCGAG
                ************************************************************

ppa002884m      GACATCCGGCTCGGCTCTCCGGACGACGCGTCCAACAATTTAGACTCGGATTTCCACTTA 960
bHLH3           GACATCCGGCTCGGCTCTCCGGACGACGCGTCCAACAATTTAGACTCGGATTTCCACTTA
                ************************************************************

ppa002884m      CTAGCAGTGAGCCAGTCTAGGGTTAACCCAGCCGATCAACAGCGCCAAGCTGACTCGTAT 1020
bHLH3           CTAGCAGTGAGCCAGTCTAGGGTTAACCCAGCCGATCAACAGCGCCAAGCTGACTCGTAT
                ************************************************************

ppa002884m      CGAGCCGAGTCAACTCGCCGGTGGCCGTCCGTACAAGAGCCACTCAGCAGCGGGCTCCAA 1080
bHLH3           CGAGCCGAGTCAACTCGCCGGTGGCCGTCCGTACAAGAGCCACTCAGCAGCGGGCTCCAA
                ************************************************************

ppa002884m      CCGCCACCCCAAGGACCCCTGGCATTGGAGGAGTTGACACATGATGACACTCACTATTCT 1140
bHLH3           CCGCCACCCCAAGGACCCCTGGCATTGGAGGAGTTGACACATGGTGACACTCACTATTCT
                ******************************************* ****************

ppa002884m      GAAACAGTCTCCACCATACTTCAAAGGCAGGCGACTAGGTGGACGGATTCCTCGTCCACC 1200
bHLH3           GAAACAGTCTCCACCATACTTCAAAGGCAGGCGACTAGGTGGACGGATTCCTCGTCCACC
                ************************************************************

ppa002884m      GATCAGGTCGCCTACTCAGCCCAATCAGCATTCATCAAGTGGACGACTCGCGTTGAACAT 1260
bHLH3           GATCAGGTCGCCTACTCAGCCCAATCAGCATTCATCAAGTGGACGACTCGCGTTGAACAT
                ************************************************************

ppa002884m      CACATGCTGGTGCCAATTGAGGGCACGTCCCAATGGCTCCTTAAATACATCTTGTTCAGC 1320
bHLH3           CACATGCTGGTGCCAATTGAGGGCACGTCCCAATGGCTCCTTAAATACATCTTGTTCAGC
                ************************************************************

ppa002884m      GTTCCGTTCCTCCACACAAAATACCGCGACGAGAACTCGCCAAAATCTCATGAGGGCGAT 1380
bHLH3           GTTCCGTTCCTCCACACAAAATACCGCGACGAGAACTCGCCAAAATCTCATGAGGGCGAT
                ************************************************************

ppa002884m      GCTTCCACTCGGTTGAGGAAAGGGACCCCACAAGATGAGCTCAGCGCCAATCATGTCATG 1440
bHLH3           GCTTCCACTCGGTTGAGGAAAGGGACCCCACAAGATGAGCTCAGCGCCAATCATGTCATG
                ************************************************************
```

```
ppa002884m      GCGGAACGGCGTCGTCGCGAGAAGCTTAATGAGAGGTTCATTATACTAAGGTCGCTAGTG 1500
bHLH3           GCGGAACGGCGTCGTCGCGAGAAGCTTAATGAGAGGTTCATTATACTAAGGTCGCTAGTG
                ************************************************************

ppa002884m      CCCTTTGTGACAAAAATGGACAAGGCCTCGATATTAGGGGACACGATCGAGTACGTGAAG 1560
bHLH3           CCCTTTGTGACAAAAATGGACAAGGCCTCGATATTAGGGGACACGATCGAGTACGTGAAG
                ************************************************************

ppa002884m      CAATTGCGTAAGAAGATTCAGGATCTTGAGGCACGTAACGTGCAGATGGAGGATGATCAA 1620
bHLH3           CAATTGCGTAAGAAGATTCAGGATCTTGAGGCACGTAACGTGCAGATGGAGGATGATCAA
                ************************************************************

ppa002884m      CGGTCAAGATCATCCGGGGAAATACATAGGTCAAATAGTATGAAAGAGTTGCGGAGCGGG 1680
bHLH3           CGGTCAAGATCATCCGGGGAAATACATAGGTCAAATAGTATGAAAGAGTTGCGGAGCGGG
                ************************************************************

ppa002884m      CTCACGGTAGTGGAGCGGACCCGGGTTGGTCCACCCGGGTCGGATAAAAGGAAGTTAAGG 1740
bHLH3           CTCACGGTAGTGGAGCGGACCCGGGTTGGTCCACCCGGGTCGGATAAAAGGAAGTTAAGG
                ************************************************************

ppa002884m      ATTGTAGAGGGAAGTGGTGGTGCGGCCGTTGCCAAGCCTAAAATGATGGAGGAGTCACCA 1800
bHLH3           ATTGTAGAGGGAAGTGGTGGTGCGGCCGTTGCCAAGCCTAAAATGATGGAGGAGTCACCA
                ************************************************************

ppa002884m      CCTTCACCACCGCCACCACCACCACAATCATCACCGACACCTATGGTGACGGGTACCTCT 1860
bHLH3           CCTTCACCACCGCCACCACCACCACAATCATCACCGACACCTATGGTGACGGGTACCTCT
                ************************************************************

ppa002884m      CTAGAGGTGTCGATAATCGAGAGTGACGGGTTGTTGGAGCTCCAATGCCCGTATAGAGAA 1920
bHLH3           CTAGAGGTGTCGATAATCGAGAGTGACGGGTTGTTGGAGCTCCAATGCCCGTATAGAGAA
                ************************************************************

ppa002884m      GGGTTGTTGCTTGATATCATGCAAACTCTAAGAGAGCTAAGAATTGAGACGACAGTTGTC 1980
bHLH3           GGGTTGTTGCTTGATATCATGCAAACTCTAAGAGAGCTAAGAATTGAGACGACAGTTGTC
                ************************************************************

ppa002884m      CAGTCCTCATTGAATAATGGGTTCTTCGTAGCTGAGCTGAGGGCCAAGGTGAAGGAGAAT 2040
bHLH3           CAGTCCTCATTGAATAATGGGTTCTTCGTAGCTGAGCTGAGGGCCAAGGTGAAGGAGAAT
                ************************************************************

ppa002884m      GTGAATGGCAAGAAAATAAGCATTACGGAAGTGAAGAGGGTAATAAATCAAATAATTCCC 2100
bHLH3           GTGAATGGCAAGAAAATAAGCATTACGGAAGTGAAGAGGGTAATAAATCAAATAATTCCC
                ************************************************************

ppa002884m      CAATCTGATTCCTAG 2115
bHLH3           CAATCTGATTCCTAG
                ***************
```

Amino acid sequence of bHLH3

```
ppa002884m      MAAPPSSSRLRSMLQASVQSVQWTYSLFWQICPQQGILVWSDGYYNGAIKTRKTVQPMEV
bHLH3           MAAPPSSSRLRSMLQASVQSVQWTYSLFWQICPQQGILVWSDGYYNGAIKTRKTVQPMEV
                ************************************************************

ppa002884m      SAEEASLQRSQQLRELYDSLSAGETNQPPARRPCASLSPEDLTESEWFYLMCVSFSFPPG
bHLH3           SAEEASLQRSQQLRELYDSLSAGETNQPPARRPCASLSPEDLTESEWFYLMCVSFSFPPG
                ************************************************************

ppa002884m      VGLPGKAYARRQHVWLTGANEVDSKTFSRAILAKSARIQTVVCIPLLDGVVELGTIERAE
bHLH3           VGLPGKAYARRQHVWLTGANEVDSKTFSRAILAKSARIQTVVCIPLLDGVVELGTIERVP
                ************************************************************.

ppa002884m      ------------------------------------------------------------
bHLH3           EDNALVQHVKTSFVDHNHHPPPKPALSEHSTSNPATSSDHPHFHSPHLPAMCPDPPLDAA

ppa002884m      --------------------SDSEAETGRNGQPVAPGPNPPAVSAPAAEPSELMQLEMSE
bHLH3           QEDEEEDEEEEDQEEDEEAESDSEAETGRNGQPVAPGPNPPAVSAPAAEPSELMQLEMSE
                                    ****************************************

ppa002884m      DIRLGSPDDASNNLDSDFHLLAVSQSRVNPADQQRQADSYRAESTRRWPSVQEPLSSGLQ
bHLH3           DIRLGSPDDASNNLDSDFHLLAVSQSRVNPADQQRQADSYRAESTRRWPSVQEPLSSGLQ
                ************************************************************

ppa002884m      PPPQGPLALEELTHDDTHYSETVSTILQRQATRWTDSSSTDQVAYSAQSAFIKWTTRVEH
bHLH3           PPPQGPLALEELTHGDTHYSETVSTILQRQATRWTDSSSTDQVAYSAQSAFIKWTTRVEH
                **************.*********************************************

ppa002884m      HMLVPIEGTSQWLLKYILFSVPFLHTKYRDENSPKSHEGDASTRLRKGTPQDELSANHVM
bHLH3           HMLVPIEGTSQWLLKYILFSVPFLHTKYRDENSPKSHEGDASTRLRKGTPQDELSANHVM
                ************************************************************

ppa002884m      AERRRREKLNERFIILRSLVPFVTKMDKASILGDTIEYVKQLRKKIQDLEARNVQMEDDQ
bHLH3           AERRRREKLNERFIILRSLVPFVTKMDKASILGDTIEYVKQLRKKIQDLEARNVQMEDDQ
                ************************************************************

ppa002884m      RSRSSGEIHRSNSMKELRSGLTVVERTRVGPPGSDKRKLRIVEGSGGAAVAKPKMMEESP
bHLH3           RSRSSGEIHRSNSMKELRSGLTVVERTRVGPPGSDKRKLRIVEGSGGAAVAKPKMMEESP
                ************************************************************

ppa002884m      PSPPPPPPQSSPTPMVTGTSLEVSIIESDGLLELQCPYREGLLLDIMQTLRELRIETTVV
bHLH3           PSPPPPPPQSSPTPMVTGTSLEVSIIESDGLLELQCPYREGLLLDIMQTLRELRIETTVV
                ************************************************************

ppa002884m      QSSLNNGFFVAELRAKVKENVNGKKISITEVKRVINQIIPQSDS
bHLH3           QSSLNNGFFVAELRAKVKENVNGKKISITEVKRVINQIIPQSDS
                *********************************************
```

Coding sequence of bHLH33

```
ppa002645m      ATGGCTAATGGGACTCAAAACCATGAGCGGGTGCCAGAGAATCTGAGAAAACAGTTTGCT  60
bHLH33          ATGGCTAATGGGACTCAAAACCATGAGCGGGTGCCAGAGAATCTGAGAAAACAGTTTGCT
                ************************************************************

ppa002645m      GTTGCTGTGAGAAGTATTAAGTGGAGCTATGCAATTTTCTGGTCATTGTCAACTTCACAA 120
bHLH33          GTTGCTGTGAGAAGTATTAAGTGGAGCTATGCAATTTTCTGGTCATTGTCAACTTCACAA
                ************************************************************

ppa002645m      CAAGGGGTGCTGGAATGGTGTGAAGGGTACTACAATGGGGACATCAAAACCCGGAAGACT 180
bHLH33          CAAGGGGTGCTGGAATGGTGTGAAGGGTACTACAATGGGGACATCAAAACCCGGAAGACT
                ************************************************************

ppa002645m      GTTGAAGGCGTGGAACTTAAAACTGATAAAATGGGTTTAGAGAGGAATGCACAACTGAGG 240
bHLH33          GTTGAAGGCGTGGAACTTAAAACTGATAAAATGGGTTTAGAGAGGAATGCACAACTGAGG
                ************************************************************

ppa002645m      GAGCTGTATAAGTCTCTTTTAGAAGGCGAAACCGAGCCACAAGCTAAAGCGCCTTCTGCT 300
bHLH33          GAGCTGTATAAGTCTCTTTTAGAAGGCGAAACCGAGCCACAAGCTAAAGCGCCTTCTGCT
                ************************************************************

ppa002645m      GCATTGAATCCAGAGGATCTCTCAGATGCTGAGTGGTATTACTTGCTTTGCATGTCCTTT 360
bHLH33          GCATTGAATCCAGAGGATCTCTCAGATGCTGAGTGGTATTACTTGCTTTGCATGTCCTTT
                ************************************************************

ppa002645m      GTCTTCAATCCTGGCGAAGGTTTGCCAGGAAGAGCATTAGCAAATGGGCAAACCATTTGG 420
bHLH33          GTCTTCAATCCTGGCGAAGGTTTGCCAGGAAGAGCATTAGCAAATGGGCAAACCATTTGG
                ************************************************************

ppa002645m      TTATGTGATGCTCAATATGCAGATAGTAAAGTATTCTCTCGCTCTTTGCTGGCGAAGAGT 480
bHLH33          TTATGTGATGCTCAATATGCAGATAGTAAAGTATTCTCTCGCTCTTTGCTGGCGAAGAGT
                ************************************************************

ppa002645m      GCATCTATTCAGACTGTGGTCTGCTTTCCCTATCTGGGTGGTGTTGTTGAGCTAGGTGTG 540
bHLH33          GCATCTATTCAGACTGTGGTCTGCTTTCCCTATCTGGGTGGTGTTGTTGAGCTAGGTGTG
                ************************************************************

ppa002645m      ACTGAGCTGGTACCAGAGGACCTTAGTCTCATTCAACACATCAAGGCTTCTTTACTGGAT 600
PpbHLH2         ACTGAGCTGGTACCAGAGGACCTTAGTCTCGTTCAACACATCAAGGCTTCTTTACTGGAT
                ********************************* ********************************

ppa002645m      TTCTCAAAGCCCGATTGCTCGGAGAAATCTTCCTCTGCACCTCATAAGGCAGATGATGAT 660
bHLH33          TTCTCAAAGCCCGATTGCTCGGAGAAATCTTCCTCTGCACCTCATAAGGCAGATGATGAT
                ************************************************************

ppa002645m      TCAGACCAAGTGCTTGCCAAGGTTGACCATGAAATAGTTGATACATTGGCTTTAGAGAAC 720
bHLH33          TCAGACCAAGTGCTTGCCAAGGTTGACCATGAAATAGTTGATACATTGGCTTTAGAGAAC
                ************************************************************

ppa002645m      CTATATTCCCCCTCAGAAGAAATCAAATTTGATCCGATGGGAATCAATGATTTACATGGA 780
bHLH33          CTATATTCCCCCTCAGAAGAAATCAAATTTGATCCGATGGGAATCAATGATTTACATGGA
                ************************************************************

ppa002645m      AACTATGAGGAGTTCAACATGGACTCTCCTGAGGAATGTTCTAATGGTTGTGAGCACAAT 840
bHLH33          AACTATGAGGAGTTCAACATGGACTCTCCTGAGGAATGTTCTAATGGTTGTGAGCACAAT
                ************************************************************

ppa002645m      CATCAGACAGAAGACTCCTTTATGCCTGAAGGTATCAATGATGGGGCTTCTCAAGTTCAG 900
bHLH33          CATCAGACAGAAGACTCCTTTATGCCTGAAGGTATCAATGATGGGGCTTCTCAAGTTCAG
                ************************************************************

ppa002645m      AGTTGGCATTTCATGGATGAAGACTTCAGCATTGGTGTTCAAGATTCCATGAATTCTAGT 960
bHLH33          AGTTGGCATTTCATGGATGAAGACTTCAGCATTGGTGTTCAAGATTCCATGAATTCTAGT
                ************************************************************

ppa002645m      GACTGCATATCTGAAGCTTTTGTTAATAAAAAAAGGGCTCAATCTTCCCCTAGACACGAG 1020
bHLH33          GACTGCATATCTGAAGCTTTTGTTAATAAAAAAAGGGCTCAATCTTCCCCTAGACACGAG
                ************************************************************

ppa002645m      AGTGTCAACCGTAACCATTTAAAGGAACTTGAAAACCTCAATGACACAAAATTTAGCTCC 1080
bHLH33          AGTGTCAACCGTAACCATTTAAAGGAACTTGAAAACCTCAATGACACAAAATTTAGCTCC
                ************************************************************

ppa002645m      TTGGATCTTGGACCTGCTGATGATCATATACACTACACAAGAACTCTTTCTAATATTCTG 1140
bHLH33          TTGGATCTTGGACCTGCTGATGATCATATACACTACACAAGAACTCTTTCTAATATTCTG
                ************************************************************

ppa002645m      GGAAGCTCAACAAGGTTGACTGAAAACCCATGTTCTTGCGATGGAGATTGCAAATCCAGT 1200
bHLH33          GGAAGCTCAACAAGGTTGACTGAAAACCCATGTTCTTGCGATGGAGATTGCAAATCCAGT
                ************************************************************

ppa002645m      TTTGTGACATGGAAGAAAGGAGTCGTTGATAATTGTAGGCCAACAGTACATCAGAAAATA 1220
bHLH33          TTTGTGACATGGAAGAAAGGAGTCGTTGATAATTGTAGGCCAACAGTACATCAGAAAATA
                ************************************************************

ppa002645m      CTAAAGAAGATTTTGTTTACAGTTCCTTTGATGTGTGGTGCGAGCTCCCAAAATACTATA 1320
bHLH33          CTAAAGAAGATTTTGTTTACAGTTCCTTTGATGTGTGGTGCGAGCTCCCAAAATACTATA
                ************************************************************

ppa002645m      CAAGATGGGCTCTCGAAGCTGCAAAGTGATGATATTCACAAGGGACATGTTATGCCTGAT 1380
bHLH33          CAAGATGGGCTCTCGAAGCTGCAAAGTGATGATATTCACAAGGGACATGTTATGCCTGAT
                ************************************************************

ppa002645m      AAACTGAAAGAGAATGAAAAATTGCTGGTCCTGAGGTCAATGGTTCCTTCTATCAGTGAG 1440
bHLH33          AAACTGAAAGAGAATGAAAAATTGCTGGTCCTGAGGTCAATGGTTCCTTCTATCAGTGAG
                ************************************************************
```

```
ppa002645m      GTTGATAAAGCATCGGTCCTGGATGACACAATTAAGTACTTGAAAGAGCTTGAGGCAAGA 1500
bHLH33          GTTGATAAAGCATCGGTCCTGGATGACACAATTAAGTACTTGAAAGAGCTTGAGGCAAGA
                ************************************************************

ppa002645m      GCAGAAGAGATGGAATCCTGCATGGACACCGTGGAAGCGATAGCTAGAAGGAAATACCTG 1560
bHLH33          GCAGAAGAGATGGAATCCTGCATGGACACCGTGGAAGCGATAGCTAGAAGGAAATACCTG
                ************************************************************

ppa002645m      GACAGGGCAGAGAAGACATCAGATAACTATGATAAAATAAAGATGGATAATGTTAAAAAG 1620
bHLH33          GACAGGGCAGAGAAGACATCAGATAACTATGATAAAATAAAGATGGATAATGTTAAAAAG
                ************************************************************

ppa002645m      CCTTGGCTAAACAAGAGAAAGGCCTGTGACATTGACGAAACTGACCCGGATCTCAATAGG 1680
bHLH33          CCTTGGCTAAACAAGAGAAAGGCCTGTGACATTGACGAAACTGACCCGGATCTCAATAGG
                ************************************************************

ppa002645m      CTTGTTCCCCGAGAAAGCTTGCCACTAGATGTGAAAGTCATTTTAAAAGAGCAGGAGGTT 1740
bHLH33          CTTGTTCCCCGAGAAAGCTTGCCACTAGATGTGAAAGTCATTTTAAAAGAGCAGGAGGTT
                ************************************************************

ppa002645m      CTGATAGAGATGAGATGCCCTTATAGGGAATATATCTTGCTTGATATAATGGATGCCATT 1800
bHLH33          CTGATAGAGATGAGATGCCCTTATAGGGAATATATCTTGCTTGATATAATGGATGCCATT
                ************************************************************

ppa002645m      AACAATCTGTACTTAGATGCTCACTCAGTCCAATCATCCACTCTTGATGGTGTTCTCACA 1860
bHLH33          AACAATCTGTACTTAGATGCTCACTCAGTCCAATCATCCACTCTTGATGGTGTTCTCACA
                ************************************************************

ppa002645m      TTGAGCCTTACATCAAAGTTTCGAGGAGCAGCGGTTGCACCAGTTGGGATGATAAAACAG 1920
bHLH33          TTGAGCCTTACATCAAAGTTTCGAGGAGCAGCGGTTGCACCAGTTGGGATGATAAAACAG
                ************************************************************

ppa002645m      GCGCTTTGGAAAATTGCCGGTAAGTGTTGA 1950
bHLH33          GCGCTTTGGAAAATTGCCGGTAAGTGTTGA
                ******************************
```

Amino acid sequence of bHLH33

```
ppa002645m      MANGTQNHERVPENLRKQFAVAVRSIKWSYAIFWSLSTSQQGVLEWCEGYYNGDIKTRKT
bHLH33          MANGTQNHERVPENLRKQFAVAVRSIKWSYAIFWSLSTSQQGVLEWCEGYYNGDIKTRKT
                ************************************************************

ppa002645m      VEGVELKTDKMGLERNAQLRELYKSLLEGETEPQAKAPSAALNPEDLSDAEWYYLLCMSF
bHLH33          VEGVELKTDKMGLERNAQLRELYKSLLEGETEPQAKAPSAALNPEDLSDAEWYYLLCMSF
                ************************************************************

ppa002645m      VFNPGEGLPGRALANGQTIWLCDAQYADSKVFSRSLLAKSASIQTVVCFPYLGGVVELGV
bHLH33          VFNPGEGLPGRALANGQTIWLCDAQYADSKVFSRSLLAKSASIQTVVCFPYLGGVVELGV
                ************************************************************

ppa002645m      TELVPEDLSLIQHIKASLLDFSKPDCSEKSSSAPHKADDDSDQVLAKVDHEIVDTLALEN
bHLH33          TELVPEDLSLVQHIKASLLDFSKPDCSEKSSSAPHKADDDSDQVLAKVDHEIVDTLALEN
                **********:*************************************************

ppa002645m      LYSPSEEIKFDPMGINDLHGNYEEFNMDSPEECSNGCEHNHQTEDSFMPEGINDGASQVQ
bHLH33          LYSPSEEIKFDPMGINDLHGNYEEFNMDSPEECSNGCEHNHQTEDSFMPEGINDGASQVQ
                ************************************************************

ppa002645m      SWHFMDEDFSIGVQDSMNSSDCISEAFVNKKRAQSSPRHESVNRNHLKELENLNDTKFSS
bHLH33          SWHFMDEDFSIGVQDSMNSSDCISEAFVNKKRAQSSPRHESVNRNHLKELENLNDTKFSS
                ************************************************************

ppa002645m      LDLGPADDHIHYTRTLSNILGSSTRLTENPCSCDGDCKSSFVTWKKGVVDNCRPTVHQKI
bHLH33          LDLGPADDHIHYTRTLSNILGSSTRLTENPCSCDGDCKSSFVTWKKGVVDNCRPTVHQKI
                ************************************************************

ppa002645m      LKKILFTVPLMCGASSQNTIQDGLSKLQSDDIHKGHVMPDKLKENEKLLVLRSMVPSISE
bHLH33          LKKILFTVPLMCGASSQNTIQDGLSKLQSDDIHKGHVMPDKLKENEKLLVLRSMVPSISE
                ************************************************************

ppa002645m      VDKASVLDDTIKYLKELEARAEEMESCMDTVEAIARRKYLDRAEKTSDNYDKIKMDNVKK
bHLH33          VDKASVLDDTIKYLKELEARAEEMESCMDTVEAIARRKYLDRAEKTSDNYDKIKMDNVKK
                ************************************************************

ppa002645m      PWLNKRKACDIDETDPDLNRLVPRESLPLDVKVILKEQEVLIEMRCPYREYILLDIMDAI
bHLH33          PWLNKRKACDIDETDPDLNRLVPRESLPLDVKVILKEQEVLIEMRCPYREYILLDIMDAI
                ************************************************************

ppa002645m      NNLYLDAHSVQSSTLDGVLTLSLTSKFRGAAVAPVGMIKQALWKIAGKC
bHLH33          NNLYLDAHSVQSSTLDGVLTLSLTSKFRGAAVAPVGMIKQALWKIAGKC
                **************************************************
```

Coding sequence of GL3

```
ppa002762m      ATGGCTACTAGGCTTCAGAACCAGGACCGGGTGCCAGAGAACCTGAGAAAACAGCTTGCT 60
GL3             ATGGCTACTAGGCTTCAGAACCAGGACCGGGTGCCAGAGAACCTGAGAAAACAGCTTGCT
                ************************************************************

ppa002762m      CTTGCTGTGAGAAGCATCGAATGGAGCTATGCAATCTTCTGGTCGATTTCAGCAAGACAA 120
GL3             CTTGCTGTGAGAAGCATCGAATGGAGCTATGCAATCTTCTGGTCGATTTCAGCAAGACAA
                ************************************************************

ppa002762m      CCAGGGGTGTTGGAGTGGGGTGATGGATACTACAATGGAGATATCAAGACAAGAAAAACA 180
GL3             CCAGGGGTGTTGGAGTGGGGTGATGGATACTACAATGGAGATATCAAGACAAGAAAAACA
                ************************************************************

ppa002762m      GTTCAAGCCATAGAACTTAATGCTGACCAAATGGGATTGCAAAGGAGTGAACAATTGAGA 240
GL3             GTTCAAGCCATAGAACTTAATGCTGACCAAATGGGATTGCAAAGGAGTGAACAATTGAGA
                ************************************************************

ppa002762m      GAACTTTACGAGTCCCTCTCGGCTGGTGAAGCAAGTCCACAAGCTAGAAGGCCTTCGGCA 300
GL3             GAACTTTACGAGTCCCTCTCGGCTGGTGAAGCAAGTCCACAAGCTAGAAGGTCTTCGGCA
                ************************************************** ********

ppa002762m      TCATTATCACCTGAAGATCTAGCTGATACGGAGTGGTATTACCTAGTTTGCATGTCATTC 360
GL3             TCATTATCACCTGAAGATCTAGCTGATACGGAGTGGTATTACCTAGTTTGCATGTCATTC
                ************************************************************

ppa002762m      GTCTTCAACGTTGGACAAGGGTTGCCAGGACGAACATTAGCAAATGGTCAACCTATCTGG 420
GL3             GTCTTCAACGTTGGACAAGGGTTGCCAGGACGAACATTAGCAAATGGTCAACCTATCTGG
                ************************************************************

ppa002762m      CTATGCAATGCTCACTATGCTGATAGTAAAGTGTTTACTCGCTCTCTCCTGGCAAAG--- 480
PpbHLH3         CTATGCAATGCTCACTATGCTGATAGTAAAGTGTTTACTCGCTCTCTCCTGGCAAAGGGC
                ************************************************************

ppa002762m      ------------ACTGTGGTATGCTTTCCATTTTTGGGAGGTGTGATTGAGCTGGGTGTG 540
GL3             GCCTCCATTCAGACTGTGGTATGCTTTCCATTTTTGGGAGGTGTGATTGAGCTGGGTGTG
                            ************************************************

ppa002762m      ACTGAGTTGGTTATGGAGGACCCTGATCTCATTCAGCATGTTAAAACATCTTTCTTGGAG 600
GL3             ACTGAGTTGGTTATGGAGGACCCTGATCTCATTCAGCATGTTAAAACATCTTTCTTGGAG
                ************************************************************

ppa002762m      GTTCCATATCCCATAGCTTCCAAGAAAACCAATCCTAGTGTAGGAAGCACAAGAAATGAC 660
GL3             GTTCCATATCCCATAGCTTCCAAGAAAACCAATCCTAGTGTAGGAAGCACAAGAAATGAC
                ************************************************************

ppa002762m      AATGATCTTGCCTGCACTGTGCTTGATCATGATGTTATGGACGCCAAGTTAATTCCTGTT 720
GL3             AATGATCTTGCCTGCACTGTGCTTGATCATGATGTTATGGACGCCAAGTTAATTCCTGTT
                ************************************************************

ppa002762m      GTAGGATGTGAAGAAATGAATGCGACTTCACCTAATAACAGTTCAAATGGTTTGGGGCTC 780
GL3             GTAGGATGTGAAGAAATGAATGCGACTTCACCTAATAACAGTTCAAATGGTTTGGGGCTC
                ************************************************************

ppa002762m      AATCAACCAGCTGACGATTCATTCATGGTTGAAGGGATGAATGGCGGAGCTTCTCAAGTG 840
GL3             AATCAACCAGCTGACGATTCATTCATGGTTGAAGGGATGAATGGCGGAGCTTCTCAAGTG
                ************************************************************

ppa002762m      CAAAGCTGGCAATTTATGGATGATGAGTTCAGTAATTTTGTACATCATTCTATGGATTCC 900
GL3             CAAAGCTGGCAATTTATGGATGATGAGTTCAGTAATTTTGTACATCATTCTATGGATTCC
                ************************************************************

ppa002762m      AGTGACTGTATATCTCAAACTTTGGTATATCCTGAAAAGGTTCCCTTGGGTCCTAAGGCT 960
GL3             AGTGACTGTATATCTCAAACTTTGGTATATCCTGAAAAGGTTCCCTTGGGTCCTAAGGCT
                ************************************************************

ppa002762m      GAAAAGGCAAGTGACCATTGCCTGCATGATCTTAAAGAGCGCAATAGCACAAAACTGACC 1020
GL3             GAAAAGGCAAGTGACCATTGCCTGCATGATCTTAAAGAGCGCAATAGCACAAAACTGACC
                ************************************************************

ppa002762m      TCTTTAGGTCCCCAAGGCACTGACTTGCAGTATCAGAGTGTTCTTTCTGCCCTTTTAAAG 1080
GL3             TCTTTAGGTCCCCAAGGCACTGACTTGCAGTATCAGAGTGTTCTTTCTGCCCTTTTAAAG
                ************************************************************

ppa002762m      GGCTCACACCAATTGATTTTGGGCCCAAACTTCCAAAATTGTCATCAGGAATCTAACTTT 1140
GL3             GGCTCCCACCAATTGATTTTGGGCCCAAACTTCCAAAATTGTCATCAGGAATCTAACTTT
                ***** ************************************************************

ppa002762m      GTCAGTTGGAAGAGAGGAGGATTTGTAAAATGCCGGAAACAAAGAGGTGGAAGCCCGCAA 1200
GL3             GTCAGTTGGAAGAGAGGAGGATTTGTAAAATGCCGGAAACAAAGAGGTGGAAGCCCGCAA
                ************************************************************

ppa002762m      AAATTATTGAAGCAGATTTTGTTTGAAGTTCCTCGGATGCATGTTGATTGTGTGCTCGAG 1260
GL3             AAATTATTGAAGCAGATTTTGTTTGAAGTTCCTCGGATGCATGTTGATTGTGTGCTCGAG
                ************************************************************

ppa002762m      TCCCCAGAAGATAACAGTAATAGAAATGGAGTTTGGAGACCAGAGGCTGATGAAATTGGT 1320
GL3             TCCCCAGAAGATAACAGTAATAGAAATGGAGTTTGGAGACCAGAGGCTGATGAAATTGGT
                ************************************************************

ppa002762m      ATGAATCATGCATTATCTGAGAGAAGGCGAAGGGAAAAACTAAATGAAAGATTTTGCGTT 1380
GL3             ATGAATCATGCATTATCTGAGAGAAGGCGAAGGGAAAAACTAAATGAAAGATTTTGCGTT
                ************************************************************

ppa002762m      TTAAAATCGATGGTCCCCTCGATTAGCAAGGATGACAAAGTATCCATATTAGATGATGCA 1440
GL3             TTAAAATCGATGGTCCCCTCGATTAGCAAGGATGACAAAGTATCCATATTAGATGATGCA
                ************************************************************
```

```
ppa002762m      ATAGAATATTTGAAAGATCTTGAGAAAAGAGTTGAAGAGTTGGAATCCTGCCGGGAGCCG 1500
GL3             ATAGAATATTTGAAAGATCTTGAGAAAAGAGTTGAAGAGTTGGAATCCTGCCGGGAGCCG
                ************************************************************

ppa002762m      TCAGATTTAGAAGCCAAAATAAAGAGGAAAATCCAAGATACTATTGAGAGAACATCTGAC 1560
GL3             TCAGATTTAGAAGCCAAAATAAAGAGGAAAATCCAAGATACTATTGAGAGAACATCTGAC
                ************************************************************

ppa002762m      AACTGTTGCAACACCAAAATTAGTAATGGGAAGAAGCCATTAGTTTACAAGAGAAAGGCC 1620
GL3             AACTGTTGCAACACCAAAATTAGTAATGGGAAGAAGCCATTAGTTTACAAGAGAAAGGCC
                ************************************************************

ppa002762m      AGTGACATTGATGAAACGGAGCCAGAAATCAGTTATGTTGTATCAAAGCACGGTTCAAGT 1680
GL3             AGTGACATTGATGAAACGGAGCCAGAAATCAGTTATGTTGTATCAAAGCACGGTTCAAGT
                ************************************************************

ppa002762m      GATAATATAACAGTGAATATGAACAAGAAGGATGTTCTAATTGAGATGAAATTTCCTTGG 1740
GL3             GATAATATAACAGTGAATATGAACAAGAAGGATGTTCTAATTGAGATGAAATTTCCTTGG
                ************************************************************

ppa002762m      AGGGAAGGAGTGTTGCTAGAGATCATGGATGCCACAAGCCGTCTCCAATTAGATACTCAC 1800
GL3             AGGGAAGGAGTGTTGCTAGAGATCATGGATGCCACAAGCCGTCTCCAATTAGATACTCAC
                ************************************************************

ppa002762m      TCAGTTCAATCATCCACAGCAGATGGGATTCTTTCCGTCACTATTAAATCCAGGTTCAAG 1860
GL3             TCAGTTCAATCATCCACAGCAGATGGGATTCTTTCCGTCACTATTAAATCCAGGTTCAAG
                ************************************************************

ppa002762m      GGGTCGACTGTCGCATCAGCAGGGACAATCCAGCAAGCACTTCAGCGAATCGCCAGAAGC 1920
GL3             GGGTCGACTGTCGCATCAGCAGGGACAATCCAGCAAGCACTTCAGCGAATCGCCAGAAGC
                ************************************************************

ppa002762m      TGTTGA 1926
GL3             TGTTGA
                ******
```

Amino acid sequence of GL3

```
ppa002762m      MATRLQNQDRVPENLRKQLALAVRSIEWSYAIFWSISARQPGVLEWGDGYYNGDIKTRKT
GL3             MATRLQNQDRVPENLRKQLALAVRSIEWSYAIFWSISARQPGVLEWGDGYYNGDIKTRKT
                ************************************************************

ppa002762m      VQAIELNADQMGLQRSEQLRELYESLSAGEASPQARRPSASLSPEDLADTEWYYLVCMSF
GL3             VQAIELNADQMGLQRSEQLRELYESLSAGEASPQARRSSASLSPEDLADTEWYYLVCMSF
                ************************************.***********************

ppa002762m      VFNVGQGLPGRTLANGQPIWLCNAHYADSKVFTRSLLAK-----TVVCFPFLGGVIELGV
GL3             VFNVGQGLPGRTLANGQPIWLCNAHYADSKVFTRSLLAKGASIQTVVCFPFLGGVIELGV
                ****************************************     *****************

ppa002762m      TELVMEDPDLIQHVKTSFLEVPYPIASKKTNPSVGSTRNDNDLACTVLDHDVMDAKLIPV
GL3             TELVMEDPDLIQHVKTSFLEVPYPIASKKTNPSVGSTRNDNDLACTVLDHDVMDAKLIPV
                ************************************************************

ppa002762m      VGCEEMNATSPNNSSNGLGLNQPADDSFMVEGMNGGASQVQSWQFMDDEFSNFVHHSMDS
GL3             VGCEEMNATSPNNSSNGLGLNQPADDSFMVEGMNGGASQVQSWQFMDDEFSNFVHHSMDS
                ************************************************************

ppa002762m      SDCISQTLVYPEKVPLGPKAEKASDHCLHDLKERNSTKLTSLGPQGTDLQYQSVLSALLK
GL3             SDCISQTLVYPEKVPLGPKAEKASDHCLHDLKERNSTKLTSLGPQGTDLQYQSVLSALLK
                ************************************************************

ppa002762m      GSHQLILGPNFQNCHQESNFVSWKRGGFVKCRKQRGGSPQKLLKQILFEVPRMHVDCVLE
GL3             GSHQLILGPNFQNCHQESNFVSWKRGGFVKCRKQRGGSPQKLLKQILFEVPRMHVDCVLE
                ************************************************************

ppa002762m      SPEDNSNRNGVWRPEADEIGMNHALSERRRREKLNERFCVLKSMVPSISKDDKVSILDDA
GL3             SPEDNSNRNGVWRPEADEIGMNHALSERRRREKLNERFCVLKSMVPSISKDDKVSILDDA
                ************************************************************

ppa002762m      IEYLKDLEKRVEELESCREPSDLEAKIKRKIQDTIERTSDNCCNTKISNGKKPLVYKRKA
GL3             IEYLKDLEKRVEELESCREPSDLEAKIKRKIQDTIERTSDNCCNTKISNGKKPLVYKRKA
                ************************************************************

ppa002762m      SDIDETEPEISYVVSKHGSSDNITVNMNKKDVLIEMKFPWREGVLLEIMDATSRLQLDTH
GL3             SDIDETEPEISYVVSKHGSSDNITVNMNKKDVLIEMKFPWREGVLLEIMDATSRLQLDTH
                ************************************************************

ppa002762m      SVQSSTADGILSVTIKSRFKGSTVASAGTIQQALQRIARSC
GL3             SVQSSTADGILSVTIKSRFKGSTVASAGTIQQALQRIARSC
                *****************************************
```