HANY S.M. EL SAYED BASHANDY

Flavonoid Metabolomics in *Gerbera hybrida* and Elucidation of Complexity in the Flavonoid Biosynthetic Pathway

DEPARTMENT OF AGRICULTURAL SCIENCES
FACULTY OF AGRICULTURE AND FORESTRY
DOCTORAL PROGRAMME IN PLANT SCIENCES
UNIVERSITY OF HELSINKI

Department of Agricultural Sciences
Faculty of Agriculture and Forestry
University of Helsinki

Doctoral Program in Plant Sciences (DPPS)

Flavonoid metabolomics in *Gerbera hybrida* and elucidation of complexity in the flavonoid biosynthetic pathway

Hany S. M. ELSayed Bashandy

Department of Agricultural Sciences
Faculty of Agriculture and Forestry
University of Helsinki
Finland

Department of Genetics, Faculty of Agriculture
University of Cairo,
Egypt

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Agriculture and Forestry, University of Helsinki, for public examination in the lecture room B7, Viikki B building (Latokartanonkaari 9), on 23rd January 2016, at 12 o'clock noon.

Cover: Inflorescence of gerbera cultivars: Left, Terraregina; Right, President (photos By Hany Bashandy and Xianbao Deng).

أَلَمْ تَرَ أَنَّ اللَّهَ أَنزَلَ مِنَ السَّمَاء مَاء فَسَلَكَهُ يَنَابِيعَ فِي الْأَرْضِ ثُمَّ يُخْرِجُ بِهِ زَرْعًا مُّخْتَلِفًا أَلْوَانُهُ ثُمَّ يَهِيجُ فَتَرَاهُ مُصْفَرًّا ثُمَّ يَجْعَلُهُ حُطَامًا إِنَّ فِي ذَلِكَ لَذِكْرَى لِأُوْلِي الْأَلْبَابِ (الزمر 21)

(Do you not see that Allah sends down rain from the sky and makes it flow as springs [and rivers] in the earth; then He produces thereby crops of varying colours; then they dry and you see them turned yellow; then He makes them [scattered] debris. Indeed in that is a reminder for those of understanding.)

To my mother (أمي), my father, my wife, my kids and my family

Table of Contents

List of original publications

This thesis is based on the following publications, which are referred to in the text by their roman numerals. The publications have been reprinted with the kind permission of the publishers.

I   **Bashandy H**, Carvalho E, Martens S, Teeri TH. 2015. Flavonoid metabolomics in *Gerbera hybrida* cultivars reveal complexity in the biosynthetic network. Journal of Agricultural and Food Chemistry (submitted)

II   Deng X, **Bashandy H**, Ainasoja M, Kontturi J, Pietiäinen M, Laitinen RAE, Albert VA, Valkonen JPT, Elomaa P, Teeri TH. 2014. Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of *Gerbera hybrida*. New Phytologist, 201: 1469-1483.

III   **Bashandy H**, Pietiäinen M, Carvalho E, Lim KJ, Elomaa P, Martens S, Teeri TH. 2015. Anthocyanin biosynthesis in gerbera cultivar 'Estelle' and its acyanic sport 'Ivory'. Planta 242: 601-611.

IV   **Bashandy H**, Jalkanen S, Teeri TH. 2015. Within leaf variation is the largest source of variation in agroinfiltration of *Nicotiana benthamiana*. Plant Methods 11:47. doi: 10.1186/s13007-015-0091-5.

My contribution to the above publications:

I   I designed the experiment and carried out plant sampling and data analysis. I was involved the data interpretation and wrote the manuscript draft with THT and revised the manuscript together with the other co-authors.

II   I was involved in the experimental design and carried the biochemistry experiments (anthocyanin analysis) and shared in the molecular parts. I was involved in data interpretation, prepared and shared the writing of the manuscript and revised the manuscript together with the other co-authors.

III   I was involved in the experimental design and conducted cloning of different genes and measured the enzymatic activities. I was involved the data interpretation, wrote the manuscript with THT and revised the manuscript together with the other co-authors.

IV   I was involved in the experimental design. I and SJ constructed the plasmids and conducted the experiments. I, SJ, and THT analysed the data. I and THT wrote the manuscript. I was involved the data interpretation and revised the manuscript together with the other co-authors.

In addition, previously unpublished results are presented.

Abbreviations

4CL	4-Coumarate-CoA ligase
ANR	Anthocyanidin reductase
ANS	Anthocyanidin synthase
Ap	Apigenin
bHLH	Basic helix-loop-helix
C4H	Cinnamate 4-hydroxylase
CCR	Cinnamoyl-CoA reductase
CHI	Chalcone isomerase
CHS	Chalcone synthase
Cy	Cyanidin
DFR	Dihydroflavonol 4-reductase
DHK	Dihydrokaempferol
DHM	Dihydromyricetin
DHQ	Dihydroquercetin
E. coli	*Escherichia coli*
EST	Expressed sequence tag
F3'H	Flavonoid 3'-hydroxylase
F3'5'H	Flavonoid 3',5'-hydroxylase
F3H	Flavanone 3β-hydroxylase
FLS	Flavonol synthase
FNS	Flavone synthase
HCT	Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase
HPLC	High performance liquid chromatography
Km	Kaempferol
LDOX	leucoanthocyanidin dioxygenase
LAR	Leucoanthocyanidin reductase
LCy	Leucocyanidin
LPg	Leucopelargonidin
Lu	Luteolin
NADPH	Nicotinamide adenine dinucleotide phosphate
PAL	Phenylalanine ammonia-lyase
Pg	Pelargonidin
PKS	Polyketide synthase
Qu	Quercetin
TET	Tetcyclacis
TT	Transparent testa
TTG	Transparent testa glabra
UFGT	UDP-flavonoid glycosyl transferase
UPLC-MS/MS	Ultra performance liquid chromatography - tandem mass spectrometer

Abstract

The flavonoid biosynthetic pathway and flavonoid metabolites have been extensively studied because of their biological roles in plant and animal kingdoms. Flavonoid metabolites are involved in plant resistance, UV-protection, pollinator attraction and have antioxidant effects. Plant species synthesise a large number of secondary metabolites, including flavonoids, carotenoids, alkaloids, terpenoids and others.

The Asteraceae or Compositae family of flowering plants is large and contains more than 20 000 species. Common examples are *Lactuca sativa, Dendranthema grandiflorum, Dahlia pinnata, Helianthus annuus* and *Gerbera hybrida*, and many of them have a wide range of flower colours.

In this work, the ornamental model plant gerbera was used to investigate the flavonoid biosynthetic pathway and to clarify its complexities. RNAseq analysis of gerbera cultivars uncovered gerbera flavonoid pathway genes, which grouped into gene families of different sizes.

Flower colour is an important feature for marketing of gerbera. In gerbera, flavonoid metabolites accumulate in the adaxial epidermal layer of flower petals and lead to different flower colours and patterns. These metabolites consist of three subgroups, flavones, flavonols and anthocyanins in gerbera, which have also been used to characterize different cultivars. The cultivars show variation in flower colour according to their anthocyanin type. Pelargonidin cultivars are orange to pink (e.g., Terraregina and Estelle), cyanidin cultivars red to fuchsia (e.g., President and Parade) and pelargonidin/cyanidin cultivars deep red (e.g., Passion and Lamborghini). Additionally, some gerbera cultivars have acyanic flowers, which do not contain anthocyanins and are white (e.g., Ivory and Virginia) or contain carotenoids and are yellow to orange (e.g., Orange Dino and Athina).

Flavones, flavonols and anthocyanins in gerbera are either 4'-hydroxlated (apigenin, kaempferol and pelargonidin) or 3'4'-hydroxylated (luteolin, quercetin and cyanidin). All cultivars produce at early stages of petal development 3'4'-hydroxylated flavones and/or flavonols, but pelargonidin cultivars cease all 3'4'-

hydroxlation when anthocyanin biosynthesis starts. 3'-hydroxlation, evoked by flavonoid 3'-hydroxylase, has biphasic activity, which seems to be controlled post-transcriptionally.

The acyanic gerbera cultivars have flavonoid profiles comparable to the cyanic cultivars, except for the synthesis and accumulation of specific anthocyanins. Metabolite profiles indicated that all analysed acyanic cultivars have a block in a late stage of the anthocyanin pathway. Ivory, a sport of the pelargonidin-cultivar Estelle, has white flowers. Ivory has flavones and flavonols, but no anthocyanin. Gene expression of all flavonoid pathway genes was similar in Estelle and Ivory. However, both cultivars have two different alleles encoding dihydroflavonol 4-reductase and in Ivory one of them (*GDFR1-2*) was found to have a point mutation resulting in inactivation of the encoded enzyme. Still, Ivory expresses the second allele (*GDFR1-3*) and accumulates active DFR enzyme. The cyanidin cultivar President expresses only the *GDFR1-3* allele, but cannot synthesize pelargonidin. Therefore, GDFR1-2 contributes specifically to pelargonidin biosynthesis and GDFR1-3 to cyanidin biosynthesis. This could be explained by a coordinated biosynthesis of anthocyanins in multi-enzyme complexes, metabolons.

Gerbera chalcone synthases (GCHSs) belong to the superfamily of Type III polyketide synthase enzymes. GCHS1, 3 and 4 have different contributions to the flavonoid pathway, according to the tissue specific and post-transcriptional regulation. RNA interference of CHS encoding genes was used in different gerbera cultivars to show that *GCHS1* has the main contribution to anthocyanin accumulation in petal tissues. *GCHS4* was strongly expressed in petals but did not lead to anthocyanin accumulation. Still, *GCHS4* is expressed and encoded a functional enzyme in the vegetative tissues.

In summary, we have shown that while transcriptional regulation is known to be important for regulation of the flavonoid pathway, also post-transcriptional regulation takes place. Post- transcriptional regulation may occur at the level of translation or protein stability, but to explain results obtained for gerbera DFR enzymes, regulation where protein-protein interactions play an important role could be more important.

1. INTRODUCTION

1.1 Flavonoid metabolites in plants

Secondary metabolites are low molecular weight molecules that are, unlike primary metabolites, not critical to growth and development of the plant. Still, they have important roles, usually in the domain of interactions with the plant and the environment (Fraenkel, 1959; Wink, 2010). The biosynthetic pathway to flavonoids, a subgroup of plant secondary metabolites, is a branch of the general phenylpropanoid pathway (Saito *et al.,* 2013). Flavonoids comprise a large fraction of all secondary metabolites with approximately 9000 different molecular structures described (Ferrer *et al*., 2008; Wink, 2010, Tohge *et al.*, 2013). Of plant polyphenols, flavonoids are the largest class (65%) (Bravo, 1998), with a common backbone of diphenylpropane (C_6-C_3-C_6) consisting of two aromatic rings (A and B) connected by a heterocyclic ring (C) (Fig. 1). Flavonoids are divided into subgroups, namely flavones (e.g., apigenin and luteolin), flavonols (e.g., quercetin and kaempferol), isoflavonoids (e.g., genistein and daidzein), anthocyanidins (e.g., pelargonidinin and cyanidin), flavan-3-ols (e.g., catechin and epicatechin) and proanthocyanidins (also called condensed tannins) (Bravo, 1998; Weisshaar and Jenkins, 1998; Winkel-Shirley, 2001; Ferrer *et al.,* 2008; Halbwirth, 2010; Tohge *et al.,* 2013). Flavonoids have different functions in different plant species. They are important for plant reproduction, adaption to stress and function as signalling molecules during, e.g., symbiosis (Winkel-Shirley, 2001, Verpoorte and Memelink, 2002; Noel *et al.*, 2005; Tohge *et al.*, 2005; Tohge *et al.*, 2013; Sato, 2014). Flavonoids accumulate usually in the epidermal layers of plant tissues and organs, and can protect the plant by absorbing harmful radiation (Weisshaar and Jenkins, 1998). The accumulation of naringenin, quercetin, kaempferol and isoflavonoids is increased during pathogen attack and they display anti-microbial properties (Paiva, 2000; Winkel-Shirley, 2001). Flavonols are required for male fertility and pollen tube growth in *Zea mays* (maize), *Petunia hybrida* (petunia), *Nicotiana tabacum* (tobacco) and *Solanum lycopersicum* (tomato) (Weisshaar and Jenkins, 1998; Schijlen *et al*., 2007) however, not in *Arabidopsis thaliana* (arabidopsis) (Shirley *et al*., 1995) or *Gerbera hybrida* (gerbera) (Prof. Paula Elamaa, personal communication). Isoflavonoids have an important role in communication between plants

of the Fabaceae family and *Rhizobia*, and are secreted from the roots during initiation of symbiosis that leads to nitrogen fixation (Winkel-Shirley, 2001; Hirsch *et al.*, 2001; Tuteja *et al.*, 2004; Yi *et al.*, 2010; Tohge *et al.*, 2013). In many floral, fruit and vegetative tissues, flavonoids are responsible for pigmentation that helps plants to attract pollinators and seed spreaders. Colours are also one of the important marketing traits in cut flowers (Winkel-Shirley, 2001).

Nutritional and pharmaceutical experiments have established that flavonoids benefit human health. They are important in diets and can be used for medication to protect or recover from diseases (Dixon and Steele, 1999). Different sources of flavonoids have been used as health foods to reduce cardiovascular diseases and some types of cancers. For example, the health promoting properties of green tea has been explained by a high level of polyphenols (Dixon and Steele, 1999; Yao *et al.,* 2004; Butelli *et al.,* 2008). For long periods, the Chinese herbal *Radix puerariae* (kudzu) has been used because of its medicinal properties, e.g., as an anti-dipsotropic agent against alcohol addiction, which is due to its isoflavonoid content (daidzein) (Keung, 2003). As edible flowers, *Chrysanthemum indicum* (chrysanthemum), *Rosa hybrida* (rose), *Syringa vulgaris* (lilac) and *Viola tricolor* (pansy) have been used in health drinks and food decoration in different cultures (Mlcek and Rop, 2011). Most land plants contain flavonoids but their content and types vary. Of edible plants, particularly berries have high content of flavonoids and also high nutritional value for human health (Kähkönen *et al.*, 2003; Yao *et al.,* 2004; Bhagwat *et al.*, 2013).

In summary, flavonoids have different biological functions in plants and different effects in animals, but the details and molecular specificities for this are in most cases not understood. Although the flavonoid biosynthesis pathway is one of the best characterized plant secondary metabolite pathways (Winkel-Shirley, 2001; Tohge *et al.*, 2013; Sato, 2014), many details are still unclear. In-depth knowledge of the contribution of different enzymes in flavonoid biosynthesis is incomplete and particularly the control of competing branches of the pathway is not well understood.

1.2 The flavonoid biosynthetic pathway

The shikimate pathway generates the main metabolic flux for the flavonoid pathway, providing for the whole general phenylpropanoid pathway a link to primary metabolism in the form of phenylalanine. Animal cells do not share this pathway and need to get aromatic amino acids (phenylalanine, tyrosine and tryptophan) through their food, essentially from plants. The metabolic enzymes of the shikimate biosynthetic pathway are conserved between different species of plants, bacteria and fungi (Herrmann, 1995). The main enzyme that modulates flux through the shikimate pathway is 3-deoxy-D-arabino-heptulosonate 7-phoshpate synthase (DAHPS) (Tzin *et al.*, 2012).

The first enzyme acting on phenylalanine is phenylalanine ammonia lyase (PAL) followed by cinnamate 4-hydroxylase (C4H) and 4-coumarate-CoA ligase (4CL). These are enzymes of the early steps of the phenylpropanoid pathway that convert phenylalanine to a variety of secondary metabolites: lignins, lignans, coumarins, stilbenoids and flavonoids (Winkel-Shirley, 2001; Tohge *et al.*, 2013). Flavonoid pathway is a major branch of the phenylpropanoid pathway. The first dedicated enzyme is chalcone synthase (CHS), which condensates 4-coumaroyl-CoA with three acetate units from malonyl-CoA to generate naringenin chalcone, a precursor for all different flavonoids (Winkel-Shirley, 2001; Saito *et al.,* 2013; Tohge *et al.*, 2013). As an exception, the flavonoid pathway can be initiated form cinnamic acid in *Pinus* species (Kodan *et al.*, 2002) (Fig. 1).

Forward and reverse genetic molecular tools, such as mutations and RNA interference technology (RNAi) have been used to elucidate the enzymatic roles of different genes that contribute to the flavonoid biosynthetic pathway (Shirley *et al.*, 1995; Tohge *et al.*, 2013). The molecular basis of the regulation of flavonoid pathway has been studied in most detail using model plants such as petunia, *Antirrhinum majus* (snapdragon), maize, arabidopsis, chrysanthemum and gerbera (Shirley *et al.*, 1995; Winkel-Shirley, 2002; Koes *et al.,* 2005; Seitz *et al.,* 2006; Laitinen *et al.*, 2008; He *et al.*, 2013). New flower colours and patterns represent an important goal for the flower industry around the world (Forkmann and Martens, 2001; Tanaka *et al.*, 2005). Biotechnology and use of detailed knowledge of the pathway has extended the natural range

7

of flower colours. The production of orange flowers with pelargonidin (Pg)-type anthocyanins in petunia was the first example of flower colour engineering (Meyer *et al.*, 1987). In carnations, flower colour has been transformed from red to a deep purple (Winkel-Shirley, 2001; Tanaka *et al.*, 2005) and blue roses have been produced (Katsumoto *et al.*, 2007). More recently, transgenic blue chrysanthemums were produced as a result of overexpression of a *Cineraria hybrida* (cineraria) genes in chrysanthemum while endogenous genes were supressed using RNAi (He *et al.*, 2013; Noda *et al.*, 2013; Brugliera *et al*., 2013).

Figure 1. Phenylpropanoid pathway in plants and the branch of flavonoid biosynthetic pathway. The core phenylpropanoid pathway (top) starts with phenylalanine, produced by the shikimate pathway. The flavonoid branch (bottom) starts with the enzyme CHS. Coloured end products are marked with their approximate colours. Enzyme names are abbreviated as follows: PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate-CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3β-hydroxylase; F3'H,

flavonoid 3′-hydroxylase; F3'5'H, flavonoid 3′5′-hydroxylase; DFR, dihydroflavonol 4-reductase; FNS, flavone synthase; FLS, flavonol synthase; ANS, anthocyanidin synthase; UFGT, UDP-flavonoid glycosyl transferase; CCR, cinnamoyl-CoA reductase; HCT, hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase; STS, stilbene synthase; ANR, anthocyanidin reductase; LAR, leucoanthocyanidin reductase. The aromatic and heterocyclic rings (A, B and C) of the flavonoid backbone with numbering of the carbon atoms in the red box.

Most enzymes of the flavonoid biosynthetic pathway are members of super families found elsewhere in primary or secondary metabolism. Such are L-amino acid ammonia lyases (PAL) (Calabrese *et al.*, 2004), acid-thiol ligases (4CL) (Wengenmayer *et al.*, 1976), polyketide synthases (CHS), NADPH-dependent reductases (DFR, LAR and ANR), UDP-glucose transferases (UFGTs) (Dixon and Steele, 1999; Tohge *et al.*, 2013), only few enzymes are unique to the pathway, such as CHI (Ngaki *et al.*, 2012). Hydroxylation occurs at several positions of the pathway. Members of the large cytochrome P450 monooxygenase superfamily have roles in hydroxylation of the B-ring of phenylpropanoids (C4H) and flavonoids (F3'H and F3'5'H) (Ferrer *et al.*, 2008). Fe^{2+}/2-oxoglutarate dependent (2-ODD) oxygenases are responsible for several other oxidation steps (F3H, ANS and FLS) (Cheng *et al.,* 2014). Interestingly, the flavone synthase of Apiaceae (FNS I) is a 2-ODD oxygenase, while in other species the reaction is catalysed by a cytochrome P450 monooxygenase (FNS II), first isolated from gerbera (Martens and Forkmann, 1998; 1999) (Fig. 1).

The genomewise most complete analysis of the flavonoid pathway has been done in arabidopsis. Arabidopsis tissues accumulate flavonols (quercetin), flavan-3-ols (epicatechin) and anthocyanins (cyanidin derivatives). Condensed tannins (proanthocyanidins) accumulate in the seed coat of arabidopsis, therefore a block in the flavonoid biosynthetic pathway or regulatory genes have a seed coat phenotype and the genes are often named as *TT* (*transparent testa*) (Shirley *et al.*, 1995; Peer *et al.*, 2001; Feller *et al.*, 2011). List of arabidopsis genes related to flavonoid biosynthesis is shown in Table 1.

Table 1. Flavonoid biosynthetic and regulatory genes in arabidopsis.

Locus	Locus identifier	Gene product	Locus	Locus identifier	Gene product
AtPAL1	At2g37040	PAL	*AtFLS1*	At5g08640	FLS
AtPAL2	At3g53260	PAL	*AtFLS2*	At5g63580	FLS
AtPAL3	At5g04230	PAL	*AtFLS3*	At5g63590	FLS
AtPAL4	At3g10340	PAL	*AtFLS4*	At5g63595	FLS
AtC4H	At2g30490	C4H	*AtFLS5*	At5g63600	FLS
At4CL1	At1g51680	4CL	*AtFLS6*	At5g43935	FLS
At4CL2	At3g21240	4CL	*TTG*	At5g24520	WD40 repeat protein
At4CL3	At1g65060	4CL	*TT1*	At5g24520	Zn-finger protein
At4CL5	At3g21230	4CL	*TT2*	At5g35550	MYB regulator
At4CL8	At5g38120	4CL	*TT8*	At4g09820	bHLH regulator
AtCHS, TT4	At5g13930	CHS	*TT10*	At5g48100	Laccase-like
AtCHI, TT5	At3g55120	CHI	*TT12*	At3g59030	Proton antiporter
AtF3H, TT6	At3g51240	F3H	*TT16*	At5g23260	MADS-box gene
AtF3'H, TT7	At5g07990	F3'H	*AtPAP1*	At1g56650	MYB75
AtDFR, TT3	At5g42800	DFR	*AtPAP2*	At1g66390	MYB90
AtANS, TT18	At4g22880	ANS	*AtMYB12*	At2g47460	MYB12
TT15	At1g43620	GT	*TT19*	At5g17220	Glutathione transferase-like

1.2.1 Phenylalanine ammonia-lyase

Phenylalanine ammonia-lyase (PAL, EC 4.3.1.24) is the first committed enzyme in the phenylpropanoid pathway and is involved in the biosynthesis of the polyphenol compounds lignins, lignans, isoflavonoids, flavonoids etc. in plants. PAL catalyses the reaction that converts phenylalanine into *trans*-cinnamic acid. The first PAL described was from barley (Koukol and Conn, 1961).

PAL is widely found encoded by a multigene family (2-6 genes) in plants (Wanner *et al.*, 1995). Fungi contain a similar enzyme that uses histidine as substrate and is called histidine ammonia lyase (HAL) (Noel *et al.*, 2005; Saito *et al.,* 2013). PAL enzymes in grasses and some species of fungi and bacteria have higher affinity for tyrosine than for phenylalanine and are therefore often named as tyrosine ammonia lyases (TAL) (Ferrer *et al.*, 2008).

Four PAL genes were found in the arabidopsis genome (*AtPAL1-AtPAL4*) and they all have different expression patterns (Cochrane *et al.*,

2004). *AtPAL1* and *AtPAL2* have the major contribution in the arabidopsis flavonoid pathway and the double mutant (*pal1/pal2*) causes reduction of flavonoid accumulation (Rohde *et al*., 2004).

1.2.2   Cinnamate 4-hydroxylase
Cinnamate 4-hydroxylase or *trans*-cinnamate 4-monooxygenase (C4H, EC 1.14.13.11) catalyses the hydroxylation step in the beginning of phenylpropanoid pathway and adds a hydroxyl group to *trans*-cinnamate at C-4 position to yield 4-coumaric acid (*p*-coumaric acid) (Springob *et al*., 2002; Saito *et al.,* 2013; Tohge *et al.*, 2013). C4H is the first plant cytochrome P450 enzyme studied in detail (Bell-Lelong *et al*., 1997). Like most other P450 enzymes, it is anchored to the endoplasmic reticulum membrane by its N-terminus. Today more than 80 members of the C4H family (CYP73A family) are listed in the cytochrome P450 home page http://drnelson.utmem.edu/CytochromeP450.html). Cytochrome P450 enzymes are classified using phylogenetic analysis into clans, families and subfamilies according to alignment similarities, without however investigating their enzymatic activity (Petersen *et al*., 2010). The first C4H was isolated form *Helianthus tuberosus* (Jerusalem artichoke) and carries the name CYP73A1 (Teutsch *et al*., 1993). C4H encoding genes are found as multigene families in most species (Lu *et al.*, 2006). However, in arabidopsis there is a single gene *AtC4H* (encoding CYP73A5). Complete loss-of-function mutants in *AtC4H* are lethal (Schilmiller *et al*., 2009). *AtC4H* mutants with compromised C4H activity have negative effects on arabidopsis growth, pollen maturation and flavonoid accumulation (Schilmiller *et al*., 2009; Saito *et al*., 2013).

1.2.3   4-Coumarate-CoA ligase
4-Coumarate-CoA ligase (4CL, EC 6.2.1.12) catalyses the conversion of 4-coumaric acid, with concomitant hydrolysis of ATP, to yield 4-coumaroyl-CoA. 4CL belongs to the family of acid-thiol ligases that catalyse the introduction of carbon-sulphur bonds.

Four paralogs of 4CL encoding genes were found in the arabidopsis genome. At4CL3 is involved in anthocyanin biosynthesis (Hamberger and Hahlbrock 2004; Tohge *et al.*, 2013), *At4CL1/2*, which are very closely related, contribute to lignin formation, as does *At4CL4* (Bassard *et al*., 2012).

1.2.4   Chalcone synthase

Chalcone synthase or naringenin-chalcone synthase (CHS, EC 2.3.1.74) is the first committed enzyme in the biosynthesis of flavonoids. CHS belongs to the superfamily of Type III polyketide synthases (PKS), which includes stilbene synthase (STS), 2-pyrone synthase (2PS), bibenzyl synthase (BBS), acridone synthase (ACS), coumaroyl triacetic acid synthase (CTAS) and many others (Flores-Sanchez and Verpoorte 2008; Saito *et al.,*2013). Type III PKS are homodimeric enzymes found mainly in higher plants, while Type I and II are found in bacteria and fungi (Austin and Noel, 2003; Austin *et al.*, 2004).

A 46-95% similarity in amino acid sequence is shared between different members of Type III PKSs in plants (Austin and Noel, 2003). CHS is the first and archetypical (first to be discovered and most widely found in plants) Type III PKS. The enzyme catalyses the formation of a polyketide intermediate during the condensation of 4-coumaroyl-CoA and three molecules of malonyl-CoA that then folds and aromatizes into naringenin chalcone (Austin and Noel, 2003; Abe and Morita, 2010; Saito *et al.,* 2013). STS generates the same polyketide intermediate, but folds it in a different pattern (Samappito *et al.*, 2003).

Mutations in the CHS encoding gene of arabidopsis (*AtCHS*, *TT4*) leads to a complete loss of flavonoids (Shirley *et al.*, 1995; Peer *et al.*, 2001). CHS is encoded by a single gene also in snapdragon (Sommer and Saedler, 1986; Shirley *et al.*, 1995; Burbulis *et al.*, 1996). However, there are multigene families encoding CHS in many plant species, for example in petunia 8-10 (Koes *et al.*, 1989), in soybean 9 (Tuteja and Vodkin, 2008), in *Ipomoea nil* (morning glory) 6 (Johzuka-Hisatomi *et al.*, 1999), in alfalfa 7, at least 5 of which are expressed in roots (Dixon *et al.*, 1996), in gerbera 3 (Helariutta *et al.*, 1995; Laitinen *et al.*, 2008) and in *Dahlia pinnata* (dahlia) 2 (Ohno *et al.*, 2011).

The loss of CHS enzyme activity results in reduced or complete lack of flavonoid accumulation (Napoli *et al.*, 1990; Elomaa *et al.*, 1993) and in some species to male sterility (Schijlen *et al.*, 2007). In arabidopsis, naringenin has been suggested to regulate auxin transport and distribution in plant cells (Brown *et al.,* 2001). In apple, downregulation of CHS by

RNAi leads to a dwarf phenotype and changes in auxin transport (Dare *et al.,* 2013).

CHSs are differentially regulated in different plant tissues and developmental stages, relating to flavonoid accumulation, plant defence or symbiosis processes (Tuteja *et al.*, 2004; Yi *et al.*, 2010; Dao *et al.*, 2011; Ohno *et al.*, 2011; Morita *et al.*, 2012). However, anthocyanin patterns in flowers of natural or cultivated plants are usually regulated downstream of CHS.

1.2.5 Chalcone isomerase

Chalcone isomerase, or chalcone-flavanone isomerase (CHI, EC 5.5.1.6) belongs to the family of isomerases. CHI catalyses the stereospecific cyclization of naringenin chalcone to (2*S*)-naringenin (NAR) (Saito *et al.,* 2013). CHI is a unique activity (Ngaki *et al.*, 2012) although CHI polypeptides in plants show similarity with peptides in fungi and bacterial cells (Gensheimer and Mushegian, 2004). In soybean, three CHIs were found, but only one contributes to the flavonoid pathway and uses tetrahydroxychalcones as substrate. The other two are responsible for isoflavonoid accumulation and they can use also 6'-deoxychalcones as substrate (Springob *et al.*, 2002).

In arabidopsis, *TT*5 is the gene encoding the enzyme (AtCHI). *tt5* mutants lack proanthocyanidins and show a lemon-yellow seed coat (Shirley *et al.*, 1995; Peer *et al.*, 2001). Petunia genome contains two *CHI* paralogs *CHIA* and *CHIB* that have highly similar sequences. However, only *CHIA* is translated to CHI (the functional enzyme) and is involved in flavonoid biosynthesis (Van Tunen *et al.*, 1991). The overexpression of the petunia *CHIA* gene led to 21-fold increased flavonol accumulation in transgenic tomato (Verpoorte and Memelink, 2002).

Morita *et al.* (2014) found mutations in morning glory that led to reduction of flavonoid accumulation and named the locus *EFP* (*Enhancer of Flavonoid Production*). *EFP* is widely distributed in petunia and *Torenia hybrida* (torenia) and its expression pattern correlated with genes of flavonoid pathway expression. *EFP* encodes a CHI-like protein but without CHI enzymatic activity.

1.2.6   Flavanone 3β-hydroxylase

Flavanone 3β-hydroxylase (F3H, syn. FHT, EC 1.14.11.9) activity is necessary for the production of both flavonols and anthocyanins. F3H is a 2-ODD oxygenase that converts (2*S*)-flavanones (NAR or eriodictyol, ERI) to dihydroflavonol (dihydrokaempferol, DHK or dihydroquercetin, DHQ) by adding a hydroxyl group at position 3 in the C-ring. FLS and ANS, two other 2-ODD oxygenase of the flavonoid pathway, can partially compensate a missing F3H activity *in planta* (Saito *et al.,* 2013). In arabidopsis, *tt6* mutants showed that *AtF3H* encodes for the single F3H enzyme in arabidopsis and the mutant has pale brown seed coat pigments and lacks flavonols (Shirley *et al.,* 1995; Pelletier and Shirley, 1996; Peer *et al.*, 2001).

1.2.7   Flavonol synthase

Flavonol synthase (FLS, EC 1.14.11.23) catalyses the biosynthesis of flavonols, a branch of the flavonoid pathway. FLS creates a double bond between C2 and C3 in the C-ring of dihydroflavonols. The enzyme yields different flavonols according to the hydroxylation pattern of the substrate, namely kaempferol from DHK, quercetin from DHQ and myricetin from DHM (Saito *et al.,* 2013). FLS competes for substrates, the dihydroflavonols, with dihydroflavonol 4-reductase (DFR). Mutations in *DFR* leads to reduction of anthocyanin and increases flavonol accumulation (Davies *et al.*, 2003). In arabidopsis, the six FLS encoding genes (*AtFLS1-AtFLS6*) have high sequence similarity, but only *AtFLS1* and *AtFLS3* encode an active enzymes (Owens *et al.*, 2008; Preuß *et al.*, 2009)

1.2.8   Flavone synthase

Flavone synthase (FNS, EC 1.14.11.22) is usually a member of the cytochrome P450 family (FNS II), but in Apiaceae belongs to the 2-ODD oxygenase family (FNS I). FNS dehydrogenates the substrate at C-2 and C-3 positions in the C-ring. FNS shares its substrate, naringenin, with F3H and yields flavones (Britsch *et al.*, 1981; Martens and Forkmann, 1998; 1999; Ferrer *et al.*, 2008).

1.2.9   Flavonoid 3'-hydroxylase

Flavonoid 3'-hydroxylase, or flavonoid 3'-monooxygenase, (F3'H, EC 1.14.13.21) is a member of the cytochrome P450 monooxygenases under

the group CYP75B. F3'H adds a hydroxyl group at 3'-position of the B-ring of flavonoids, naringenin and DHK resulting to eriodictyol and DHQ, respectively (Springob *et al.*, 2002; Seitz *et al.*, 2006, Saito *et al.,* 2013). The hydroxylation pattern of the B-ring is a determinant of the anthocyanin pigment type. F3'H is necessary for the biosynthesis of luteolin, quercetin and Cy-type anthocyanins (Winkel-Shirley, 2001; Peer *et al.*, 2001).

Tetcyclasis (TET) is a plant growth regulator that has been used experimentally as a chemical inhibitor for cytochrome P450 enzymes. TET is particularly effective as an F3'H inhibitor (Martens and Forkmann, 1998; He *et al.*, 2013).

*TT*7 (encoding CYP75B1) is the single F3'H in arabidopsis (Shirley *et al.*, 1995; Peer *et al.*, 2001). The gerbera gene (*GF3'H*) has been cloned and established for its enzymatic activity, too (Seitz *et al.*, 2006).

1.2.10 Flavonoid 3',5'-hydroxylase

Flavonoid 3'5'-hydroxylase (F3'5'H, EC 1.14.13.88) is a member of the cytochrome P450 monooxygenase group CYP75A, closely related to F3'H. F3'5'H is an important enzyme in plants that synthesize delphinidin-type anthocyanins, the flavonol myricetin and the flavone tricetin (Menting *et al.,* 1994). The enzyme catalyses the hydroxylation at both the 3' and 5' positions of the B-ring of flavonoids. Two loci (*Hf1* and *Hf2*) encode F3'5'H enzymes in petunia, and show 94% sequence identity (Menting *et al.,* 1994; Springob *et al.*, 2002).

Many flowering plants, such as are roses, carnations, chrysanthemums, lilies, gerbera and arabidopsis, lack flavonoid 3',5'-hydroxylase activity and therefor of delphinidin-type anthocyanins. (Brugliera *et al.*, 2000; Katsumoto *et al.,* 2007; Seitz *et al.,* 2006; Falginella *et al.,* 2010; He *et al.,* 2013). Some species of Asteraceae produce delphinidin and do contain F3'5'H activity. Interestingly, these enzymes are more related to F3'H than to F3'5'H of other species, indicating an independent evolutionary origin (Seitz *et al.,* 2006).

1.2.11 Dihydroflavonol 4-reductase

Dihydroflavonol 4-reductase (DFR, EC 1.1.1.219) catalyses the first committed reaction leading to anthocyanins (and proanthocyanidins).

DFR catalyses the reduction of the C-4 keto group in dihydroflavonols (DHK, DHQ and DHM) to the corresponding 4-hydroxyl group in leucoanthocyanidins (leucopelargonidin (LPg), leucocyanidin (LCy) and leucodelphinidin (LDp), respectively) (Springob *et al.*, 2002; Saito *et al.,* 2013). DFR encoding nucleotide sequences have been isolated, cloned and characterized from many different plant species, including snapdragon, maize, arabidopsis, petunia, *Vitis vinifera* (grape) and gerbera. Gerbera DFR cDNA was isolated and characterized from petals (Helariutta *et al.,* 1993; Martens *et al.,* 2002), and the crystal structure of grape DFR established the substrate and NADPH binding sites of the enzyme (Petit *et al.,* 2007).

In arabidopsis, *TT3* (*AtDFR*) was found to be a single gene that contributes to proanthocynidins in the seed, *tt3* mutant plants have a characteristic yellow seed coat (Shirley *et al.,* 1995; Peer *et al.*, 2001). Some DFRs, e.g., those from petunia and the orchid *Cymbidium hybrida* have substrate specificity and cannot reduce DHK to Pg-type anthocyanins (Forkmann and Ruhnau, 1987; Johnson *et al.,* 1999). Changing of single amino acids in the substrate binding site of DFR can, in fact, modify substrate specificity of the enzyme (Johnson *et al.,* 2001).

1.2.12 Anthocyanidin synthase

Anthocyanidin synthase, or leucoanthocyanidin dioxygenase (ANS or LDOX, EC 1.14.11.19), is a member of the 2-ODD oxygenase family that catalyses the conversion of colourless leucoanthocyanidins to coloured anthocyanidins close to the end of the anthocyanin pathway (Springob *et al.*, 2002). Saito *et al.* (1999) were the first to functionally express an active ANS in *E. coli*, isolated from *Perilla frutescens* (perilla). In arabidopsis, ANS is encoded by the locus *TT18* that contributes to anthocyanin and proanthocyanidin production. Mutants *tt11*, *tt17*, and *tds4* (*tannin deficient seed 4*) are described as allelic versions of *tt18* and inhibit ANS activity in arabidopsis seedlings (Shirley et al., 1995; Appelhagen *et al.*, 2011). ANS has been identified in many plant species including the *Prunus persica* (peach) (Ogundiwin *et al.,* 2008), in grape (Gollop *et al.,* 2001) and also in gerbera (Wellmann *et al.*, 2006).

Turnbull *et al.* (2000) showed that the arabidopsis ANS has a range of enzymatic activities that convert leucocyanidin to dihydroflavonols (DHQ) and traces of cyanidin and, also, ANS can convert dihydroflavonols to flavonols. Gerbera ANS showed the same ability to use different substrates *in vitro* and can convert (+)-catechin to cyanidin, quercetin, and a dimer of catechin (Wellmann *et al.*, 2006).

1.2.13 UDP-flavonoid glycosyl transferase

UDP-flavonoid glycosyl transferase (UFGT) is a common name for glycosyl transferases that are pivotal enzymes for flavonoid biosynthesis (Vogt and Jones 2000). Glycosylation stabilizes and increases water solubility of flavonoids and anthocyanins by adding sugar groups, in most cases to the C-ring, of the anthocyanidin skeleton but also A- and B-ring glycosylation (mainly of flavonoids) are known (Springob *et al.*, 2002). *UFGT* expression had similar patterns as the genes of the anthocyanin pathway in the red skin of grape (Kobayashi *et al.*, 2001) and in potato (Wei *et al.*, 2012), and lacking of the pigment accumulation correlated with the inactivation or absence of *UFGT* expression and activity, respectively.

Different *UFGT* members have specific names according to their specificity regarding the position and type of the sugar added. Flavonol 3-*O*-glucosyl transferase (3GT, EC 2.4.1.91) catalyses the glucoside formation of flavonoids at position 3 in the C-ring. UFGT are encoded by a number of different loci (Kobayashi *et al.,* 2001; Saito *et al.,* 2013). Small changes in the GT sequence are critical for substrate specificity, for example, 5-GTs in petunia and perilla have high sequence similarity but the petunia enzyme does not accept anthocyanidin 3-glucosides, while the perilla 5-GT accepts a wide range of substrates (Springob *et al.*, 2003).

In addition to glycosylation, both acylation and methylation of anthocyanins is supposed to have important functions for the stability of the molecules and both can serve to protect the glucosidic bonds against degradation (Springob *et al.*, 2003).

1.3   Regulation of flavonoid biosynthetic pathway

Flavonoid biosynthetic genes are regulated spatially and temporally during plant development by internal (developmental) and external signals (Koes *et al*., 1989; Jackson *et al*., 1992; Weisshaar and Jenkins, 1998; Feller *et al*., 2011; Zhou *et al.*, 2011). Transcriptional regulation of the pathway takes place in concert between several genes reacting to the same transcription factors. In monocots, the whole pathway to anthocyanins is regulated by the same transcription factors, however, in dicots the early and the late parts are regulated separately (Mol *et al*., 1998; Carey *et al*., 2004).

Different flavonoids are synthesized at different times during plant development (Kleindt *et al*., 2010; Song *et al*., 2014). The pathway has two branch points where different enzymes are competing for the same substrate (FNS and F3H for naringenin, FLS and DFR for dihydroflavonols). Control could take place based on timing, cell specificity, or within the cell where specific complexes of the flavonoid pathway enzymes (metabolons) would be formed (Winkel-Shirley, 2004).

1.3.1   Transcription factors

Regulation of the flavonoid pathway is based on endogenous elements (transcription factors) that modulate gene expression patterns of the pathway (Winkel-Shirley, 2002; Koes *et al*., 2005; Laitinen *et al.*, 2008; Stommel *et al*., 2009; Dixon *et al*., 2013).

Generally, transcription factors can have different regulatory roles as inducers or suppressors of transcription, and this is true also for regulators of the flavonoid biosynthesis. MYB, basic helix-loop-helix (bHLH) and WD40 repeat proteins are factors that interact to form the regulatory complex MYB-bHLH-WD40 that controls transcription of the genes in the anthocyanin pathway in many plants (Ramsay and Glover, 2005; Koes *et al*., 2005; Albert *et al*., 2014). Yeast two-hybrid experiments have demonstrated interaction between all pairs of MYB, bHLH and WD40 proteins (Koes *et al*., 2005, Dixon *et al*., 2013).

MYB proteins are characterized by three repeats of the conserved MYB DNA binding motif in mammalian cells (R1, R2 and R3) and one to four

in plants (Feller *et al*., 2011). The R2R3-MYB proteins C1 and Pl were characterized in maize where anthocyanin accumulation is dependent of C1 for seed pigmentation and Pl regulates phlobaphenes and flavone C-glycosides (Taylor and Briggs, 1990; Grotewold *et al.*, 1998; Stommel *et al*., 2009; Feller *et al*., 2011). In petunia, *AN2* encodes the anthocyanin regulator MYBPH3 that accumulates in the epidermal cell layer of petals, where flavonoid biosynthetic genes are actively expressed (Solano *et.al.*, 1995; Quattrocchio *et al*., 1999). The gerbera GMYB10 encodes a R2R3-MYB regulator that can induce anthocyanin biosynthesis in transgenic tobacco (Elomaa *et al*., 2003) and overexpression of *GMYB10* in transgenic gerbera plants significantly increased anthocyanin accumulation (Laitinen *et al*., 2008).

RNAi down regulation and ecotopic expression of petunia *MYB27* showed that MYB27 acts as a repressor of anthocyanin accumulation and flower pigmentation (Albert *et al*., 2014). The reppersor made its effect through its C-terminal EAR motif and targeted both anthocyanin pathway genes and one of the pathway regulators (Albert *et al*., 2014). The ectopic expression of the arabidopsis *AtMYB60* in transgenic lettuce (*Lactuca sativa*) inhibited anthocyanin accumulation through downregulation of DFR (Park *et al*., 2008).

The bHLH proteins (sometimes called MYC proteins) have a conserved basic helix-loop-helix (bHLH) domain consisting of a DNA-binding and a protein dimerization motif. bHLH proteins were first found in maize, where they are encoded by the *R* gene family (*R* and *B* loci). The *R* gene family encodes R-MYC proteins that contribute to anthocyanin accumulation in maize (Dooner *et al*., 1991). In petunia *AN1* encoding a bHLH protein that regulates the late part of the flavonoid pathway (Verpoorte and Memelink, 2002; Quattrocchio *et al*., 1999). In gerbera, GMYC1 was found to bind the gerbera *DFR* promoter in a transient assay and its expression correlated with anthocyanin accumulation in different cultivars (Elomaa *et al.*, 1998).

The WD40 repeat protein family contains the third component of the anthocyanin regulatory complex (Smith, 1999). WD proteins are encoded by the gene *PAC1* in maize, *AN11* in petunia, and *TTG1* in arabidopsis (Carey *et al*., 2004, DeVetten *et al*., 1997, Ramsay and Glover, 2005).

1.3.2   Multi-enzyme complexes (metabolons)

Helen Stafford was the first plant biochemist to suggest that multi-enzyme complexes, later termed metabolons by Paul Srere (Srere, 1985), would be responsible for plant secondary metabolite synthesis. Her hypothesis stated that enzymes of the secondary metabolism interact, bind together and channel the intermediate substrates from enzyme to enzyme until the final product is released (Stafford, 1974). Channelling would increase efficiency and protect the intermediates that may be unstable or toxic (Mathews, 1993). Multi-enzyme complexes and channelling have been suggested to be involved also in primary metabolite pathways. Enzymes in the Calvin-Benson cycle (Gontero *et al.*, 1988) and the tricarboxylic acid cycle (Meyer *et al.*, 2011; Bartholomae *et al.*, 2014) represent well characterized multi-enzyme complexes. Stafford's hypothesis has gained support and has been investigated through experiments performed both *in vivo* and *in vitro* (Winkel-Shirley, 2004).

Fractionation experiments provided the first evidence of multi-enzyme complexes and protein-protein interactions between the flavonoid pathway enzymes. C4H is a membrane associated protein, but also the operationally soluble enzymes (PAL, CHS and GT) were found to be membrane associated in cell suspension of *Haplopappus gracilis* (slender goldenweed) (Fritsch and Grisebach, 1975) and in petals of *Hippeastrum* (amaryllis) (Wagner and Hrazdina, 1984).

Sewalt *et al.* (1997) analysed transgenic lines of *Nicotiana tabacum* (tobacco), where PAL and C4H were separately downregulated. Both transgenic lines showed not only reduction of lignin accumulation, but also modification of lignin in respect to monolignol composition ratios, however in an opposite way. These changes proposed that monolignol biosynthesis occurs by more than one route in the cells, with separate substrate channelling between PAL and C4H in separate multi-enzyme complexes (metabolons).

PAL is generally considered as a soluble enzyme, but it occurs both as a soluble and a membrane-bound form. The association of PAL with the endoplasmic reticulum (ER), together with C4H, has been further established by immunological methods, making effective channelling of

metabolites through the phenylpropanoid pathway plausible (Rasmussen and Dixon, 1999; Sato *et al.*, 2004).

When ^{3}H-L-phenylalanine was fed to tobacco cell cultures undergoing flavonoid biosynthesis, the pathway intermediate ^{3}H-*trans*-cinnamic acid did not equilibrate with the cytoplasmic cinnamate pool. This was interpreted as proof for channelling between PAL and C4H in the pathway (Rasmussen and Dixon, 1999). Channelling of unstable intermediates in isoflavonoid biosynthesis between the membrane associated *O*-methyltransferase and isoflavone synthase was observed in alfalfa. Here, channelling protected intermediates, so that the final product *in vivo* was different from the product made *in vitro* (Liu and Dixon, 2001).

Yeast two-hybrid and co-immunoprecipitation experiments have been used to determine that arabidopsis enzymes of the flavonoid pathway interact with each other. Yeast two-hybrid showed AtCHS, AtCHI and AtDFR interaction, and co-immunoprecipitation demonstrated the interaction between AtCHS, AtCHI and AtF3H (Burburlis and Winkel-Shirley, 1999; Saslowsky and Winkel-Shirley, 2001).

Owens *et al.* (2008) showed using yeast two-hybrid assay that AtFLS1, AtFLS3 and AtFLS5 all interact with AtCHS. As AtFLS5 is catalytically inactive, they brought up the interesting hypothesis that non-catalytic forms of enzyme molecules may have roles in metabolon formation (compare with involvement of non-catalytic CHI forms described in section 1.2.5).

Förster resonance energy transfer between two different fluorescent proteins has been used to demonstrate interaction between enzymes of the flavonoid pathway in arabidopsis cells. Both AtFLS1 and AtDFR interact with AtCHS. The interaction between AtCHS and AtDFR, or AtFLS1, was weak in the presence of AtFLS1 or AtDFR, respectively, demonstrating that AtFLS1 and AtDFR compete for the same binding site (Crosby *et al.*, 2011).

The metabolon hypothesis has often been used as a suitable concept to explain different aspects of flavonoid metabolism. Not only does it

explain efficient passing of toxic and unstable reaction intermediates from one enzyme to the other, but presence of different coexisting metabolons allows control of substrate competition within the pathway (Winkel-Shirley, 2001; Bassard *et al*., 2012). However, the precise nature of how proteins interact to form multi-enzyme complexes *in vivo* has been very difficult to demonstrate except in few cases (Crosby *et al*., 2011). Engineering of PAL and C4H into yeast did not show any evidence for multi-enzyme complexes (Ro and Douglas, 2004). Reasons for these difficulties could be in the stability of metabolon structure *in planta* compared to *in vitro*, or to using foreign hosts that may lack some important components. Finally, if metabolons are essential in efficient flavonoid biosynthesis, mutations in interaction components or interfaces should lead to compromised biosynthesis or incorrect pathway branching. This kind of mutations have not been found.

2.   AIMS OF THE STUDY

The aims of the PhD study presented here were to investigate how the ornamental plant gerbera manages the branched pathway of flavonoid biosynthesis, specifically:

1. Does the hydroxylation pattern of flavones, flavonols and anthocyanins correlate in different gerbera cultivars.
2. Are different allelic or paralogous forms of genes encoding flavonoid related enzymes dedicated to different branches of the pathway. Particularly, are cyanidin and pelargonidin branches using the same or different enzymatic forms.

In addition, we wanted to optimize heterologous production of gerbera proteins in *Nicotiana benthamiana* for enzymatic assays.

3. MATERIALS AND METHODS

The materials and methods described in publications I-IV are summarized in Table 2.

Table 2. Materials and methods that have been used in this study. The Roman numerals refer to the four original publications.

Material or method	Publication
Nicotiana benthamiana plants	III, IV
Gerbera plants	I, II, III
Total RNA isolation and cDNA synthesis	I, II, III
Quantitative real time PCR	(II)
Isolation of polysome RNA	(II)
Construction of TRV vectors	(II)
DFR constructs	III
Vacuum agroinfiltration	(II), III
Syringe agroinfiltration	III, IV
CHS enzyme assay	(II)
DFR enzyme assay	III
HPLC analysis of flavonoids	III
UPLC/MS/MS analysis	(I,III)
Western blotting	(II), III
In situ hybridization	(II)
Nested analysis	IV
Tetcyclasis treatment	III

The publication in brackets denotes that the methods were conducted by co-authors in the respective publications.

F3'H enzyme assay in gerbera flower extracts

Microsomal membrane fractions from gerbera petals were isolated according to Diesperger *et al*. (1974) and used as a source for F3'H. Protein concentration was quantified by the method of Bradford (1976). F3'H enzymatic activity was assayed in a reaction mixture (Heller *et al*., 1985) of 100 µL, consisting of 100 mM Tris-HCl (pH 7.5), 0.1 mM DHK, 1 mM NADPH, 10 µg of microsomal proteins and in treated samples 0.1 mM TET (dissolved in 50% ethanol). The assay mixture was incubated at 30 ºC for 1 h. Reaction was stopped with 10 µL glacial acetic acid, then extracted twice with 100 µL ethyl acetate. Extracts were combined, dried and dissolved in 80 µL methanol for HPLC analysis, which detected the presence of the reaction product DHQ.

RNAseq preparation

TRIzol®reagent (Invitrogen) was used according to the manufacturer's instructions for extracting the total RNA from gerbera inflorescence tissues according to Table 3.

Table 3. Gerbera tissue samples collected for RNA extraction. The developmental stages are according to Helariutta *et al.* (1993) (Fig. S1).

Cultivar	Tissue	Developmental stage	Number of paired end reads
Terraregina	Petal	3	5 242 135
	Petal	6	4 339 960
	Scape	3	13 159 893
	Anther	8	12 259 493
Terraregina-MYB10	Petal	3	7 645 542
	Petal	6	7 075 139
	Scape	3	8 190 898
	Anther	8	5 228 370
Estelle	Petal	5	10 738 063
	Petal	7	2 196 824
Ivory	Petal	5	18 739 129
	Petal	7	7 089 932

Samples were treated with DNase (Fermentas, Leon-Rot, Germany) to remove any residues of genomic DNA. The next steps, including poly-A RNA fractionation on oligo-dT columns and subsequent sequencing of paired end reads by Illumina technology was done at the Biomedicum Functional Genomics Unit (University of Helsinki, Finland). RNA sequencing yielded 2-18 million read pairs of 140 bp per library (Table 3). Reads, together with previously generated 16994 gerbera ESTs (Laitinen *et al.*, 2008) and 300177 gerbera Roche 454 reads (unpublished), were assembled in Ghent University, Belgium.

4. RESULTS AND DISCUSSION

4.1 Flavonoid metabolites in gerbera cultivars (I)

Gerbera is a member of the Asteraceae family, which is one of the largest families of angiosperm plants. Asteraceae includes a full l0 % of flowering plant species and many of the economically important species, such as *Helianthus annuus* (sunflower), *Calendula officinalis* (pot marigold), and chrysanthemum belong to Asteraceae (Funk *et al*., 2005). Gerbera is a good model to study the Asteraceae family because of variation in the inflorescence structure and in flavonoid metabolites (Teeri *et al.*, 2006), and because of the possibility to apply genetic transformation (Elomaa *et al*., 1993). The flower colours of gerbera have a wide range and are derived from carotenoids and anthocyanins (Asen, 1984; Tyrach and Horn, 1997).

Flavonoid metabolites in gerbera consist of three subgroups: flavones, flavonols and anthocyanins. Most of the anthocyanins accumulate in the adaxial epidermi of the petals and give the different cultivars their distinctive colour (Helariutta *et al.,* 1995; Fig. 2 in Paper I).

Flavonoid metabolites in different gerbera cultivars (Table 1 in Paper I; Fig. 3 in Paper III) and through different developmental stages of petal development were analysed using Ultra Performance Liquid Chromatography - Tandem Mass Spectrometer (UPLC-MS/MS) (Vrhovsek *et al*., 2012; Paper III). Both cyanic (i.e., anthocyanin containing) and acyanic cultivars were found to have similar flavonoid profiles, except for anthocyanin accumulation. All gerbera cultivars accumulated flavonols (kaempferol or quercetin) and the cyanic cultivars anthocyanins pelargonidin and/or cyanidin (Figs 3 and 4 in Paper I; Fig. 3 in Paper III). This indicates that the acyanic cultivars must have a block after dihydroflavonol formation, which is a common precursor to both flavonols and anthocyanins (Fig. 2) (Davies *et al.,* 2003; Tyrach and Horn, 1997).

Flavonoids were present nearly exclusively as glycosides in all samples, only apigenin, luteolin and kaempferol were also found in their aglycone form. The position C-3 in the C-ring of the flavonoid structure was the most common position of glycosylation, (except for flavones where position C-7 is used instead), but also glycosides at position C-5 in the A-ring were found for pelargonidin and cyanidin (as a 3,5-diglycoside), at position C-4' in the B-ring for quercetin (as a 3,4'-diglycoside) and position C-3 in the C-ring as 3-rutinoside for flavonols (Fig. 3 and 4 in Paper I).
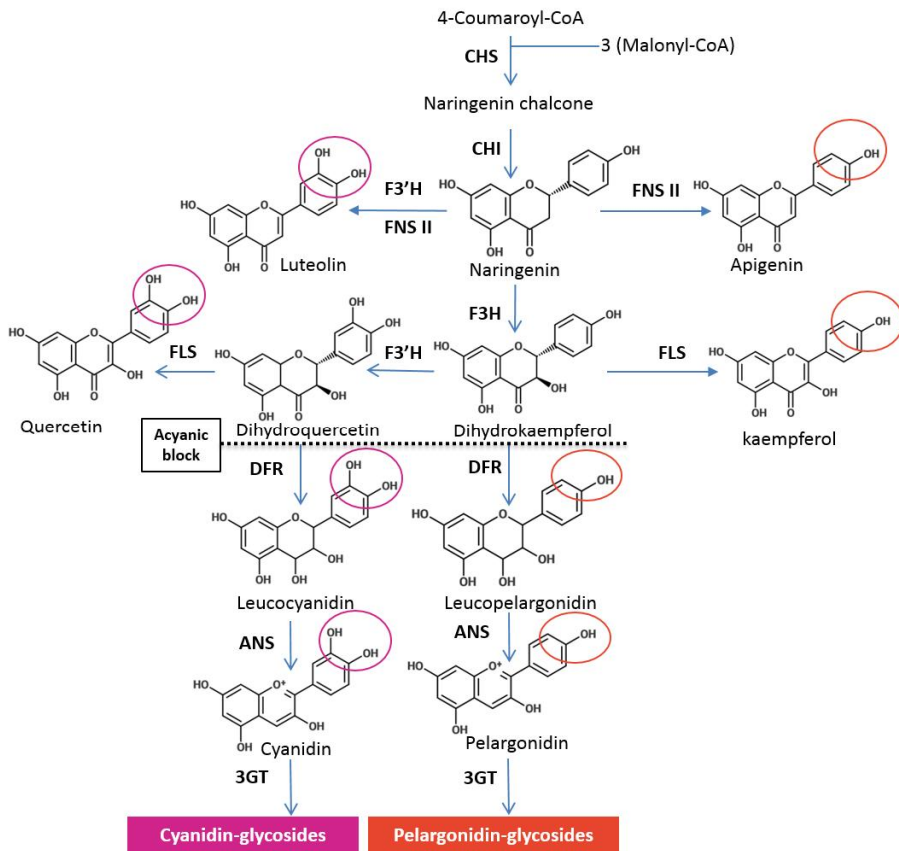
Figure 2: The core of flavonoid pathway in gerbera cultivars. The pathway is complete in cyanic gerbera cultivars but in acyanic cultivars there is a block in the late stages of the pathway. The circles show the shared hydroxylation patterns between apigenin, kaempferol and pelargonidin in Pg-cultivars and luteolin, quercetin and cyanidin in Cy-cultivars. Enzyme abbreviations are shown in Figure 1.

UPLC/MS/MS analysis showed also presence of other phenolic compounds (Fig. 5 in Paper I) such as chlorogenic acids, not described before in gerbera, but common in many plants like apple (Hulme, 1953) and coffee (Lallemand *et al*., 2012). Chlorogenic acids are known for their high antioxidant capacity (Leiss *et al*., 2009).

Gerbera cultivars can be divided according to their anthocyanin content to acyanic, Pg-, Cy- and Pg/Cy-types. Most Pg-cultivars contained, in addition to pelargonidin, apigenin and kaempferol, Cy-cultivars luteolin and quercetin, and Pg/Cy cultivars all four flavones and flavonols (Paper I). In these, the hydroxylation patterns correlate between the three flavonoid groups. However, the Pg-types Salsa and Cariba contained quercetin in their fully opened petals (Fig. 6 in Paper I). In fact, all analysed Pg-types showed some accumulation of quercetin during early stages of flower development (Fig. 7 in Paper I). This indicated that F3'H enzymatic activity is present at early stages, but absent when anthocyanins are synthesized. Separate early and late activity of F3'H can be observed also in Cy-type cultivars, which therefore show a biphasic pattern of the B-ring hydroxylation of the colourless flavonols.

4.2 Gerbera *F3'H*

Hydroxylation pattern of the B-ring determines the flavonoid type for flavones, flavonols and anthocyanins (Halbwirth, 2010). The enzyme F3'H is responsible for adding the hydroxyl group to carbon position 3' in all three types (Seitz *et al*., 2006; Saito *et al*., 2013).

The gerbera cDNA for F3'H (*GF3'H*) has been isolated and shown to encode for an active enzyme (Seitz *et al*., 2006). The gerbera RNAseq data generated in this work showed that most likely this is the only F3'H encoding gene expressed in petals (Fig. S2).

RNAseq (Fig. S2) and microarray data (Laitinen *et al*., 2008) showed that expression of *GF3'H* is upregulated during development of petals in the cultivar Terraregina. Particularly, petals at stage 6 contain ten times more *GF3'H* transcripts than petals at stage 3 (Fig. S2), although 3'-hydroxylation does not take place at stage 6 (Fig. 7 in Paper I). In Terraregina-MYB10, a transgenic line overexpressing the anthocyanin regulator *GMYB10*, GF3'H is upregulated both in floral and vegetative tissues.

The enzymatic activity of GF3'H was determined in petals at stage 6/7 of three cultivars or lines: the Pg-type Terraregina, the Cy-type President, and Terraregina-MYB10 (Fig. 3). TET, which is a specific chemical inhibitor for F3'H (Martens *et al.*, 2003), was used to inhibit F3'H activity in the assays and discriminate between F3'H and background hydroxylase activity. As summarized in Figure 3, Terraregina petals did not contain any F3'H activity, Terraregina-MYB10 petals contained some activity and President petals many times more (Fig. 3). Spraying of young inflorescences with TET blocked anthocyanin accumulation in the Cy-type cultivar President (Fig. 4), which was reflected to the F3'H enzyme levels in petal microsomes (Fig. 3). Interestingly, Cy-type cultivars do not start accumulating pelargonidin after TET treatment. This is discussed further in Paper III.
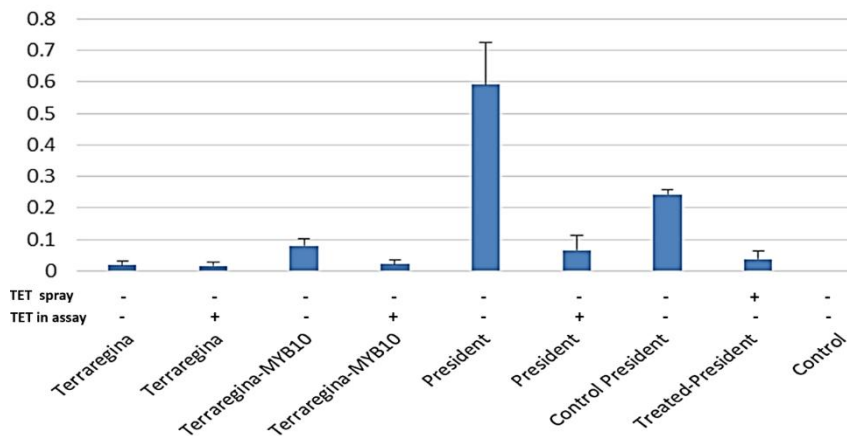
Figure 3: F3'H enzymatic activity in microsomal extract of petals. Terraregina showed only the background hydroxylation of the substrate DHQ. The last two samples were collected from flowers shown in Figure 4, after spraying of young flower with TET. For others, TET was added in the reaction mixture. The control contained only substrate without protein extract. DHQ concentration is shown in μM.

Figure 4: Treatment of the Cy-type cultivar President inflorescences. The treated inflorescence was sprayed with TET (1mg/ml in 50% ethanol) (right), and the control with 50% ethanol only (left). Both inflorescences were treated at early stages, when petals were green.

In summary, although *GF3'H* transcripts increasingly accumulate during petal development, there is no enzymatic activity in the Terraregina petals at stage 6/7.

4.3 Flavonoid biosynthetic pathway genes in gerbera

RNAseq data expanded our knowledge about gene families encoding enzymes of the flavonoid pathway in gerbera (Papers II and III).

The more than 100 million paired end reads generated in this work, together with previously generated gerbera transcriptome sequences were assembled into 295 331 contigs (of which 50 245 were longer than 500 bp). The contigs were individually annotated by comparison to published sequences using the tblastx algorithm (Altschul *et al.*, 1997). Assemblies encoding enzymes for the flavonoid pathway were extracted by keyword searches in the annotation and by a similarity blastx search with known sequences. Similar assemblies were grouped using the mcl algorithm (Van Dongen, 2000 ; Enright *et al.*, 2002) and a threshold of 95% nucleotide sequence identity was used to categorize similar sequences to allelic variants (>95%) or paralogs (<95%). This, of course, does not group together all fragmental assemblies. Gerbera RNAseq reads generated in this work and by others in the laboratory (Teemu Teeri and Paula Elomaa, unpublished) were mapped to the set of gerbera flavonoid related

assemblies using the blastn algorithm, automatized in the DOS shell LocalBlast (Teemu Teeri, unpublished). LocalBlast also extracts read counts at specified stringencies defined post mapping. This feature helps in characterisation of expression profiles for close paralogs. Profiles at stringency 1e-50 are shown in appendixes.

The RNAseq (Table 4; Fig. S1) and microarray data (Laitinen *et al.*, 2008) showed upregulation of many genes that relate to flavonoid biosynthesis, for example *GPAL*, *GC4H*, *GCHS3*, *GCHS4*, *GCHI1*, *GF3H, GF3'H* and *GDFR1*. However, no changes in expression of *4CL, GCHS1*, *GFLS* and *GFNS* were observed (Laitinen *et al.*, 2008; Table 4). Genes that have high expression in petals (petal abundant) or are up-regulated in the Terraregina-MYB10 line, might be involved in the flavonoid pathway. These were isolated as full length cDNA molecules and sequenced (Table 4; Fig. S2). Most of the cloned sequences (16 genes) had high level of alignment identity with assembly fragments sequences, close to 100%.

Gerbera RNAseq assembly showed that flavonoid biosynthetic gene families contain sometimes single genes like most flavonoid gene families in arabidopsis (Saito *et al.*, 2013), but in other cases multiple genes were found, which can be expressed differently in the floral and vegetative tissues. Some gene paralogs were not upregulated in Terrearegina-MYB10 and were not petal abundant, and could therefore be involved in other phenylpropanoid pathway branches, like *GPAL*, *GC4H* and *G4CL* in lignin biosynthesis (Table 4; Fig. S1). The exact number of genes in the families is difficult to conclude since all gene fragments in the assembly do not necessary represent different genes.

Table 4: Gene families of flavonoid biosynthetic pathway in gerbera according to assembled Illumina reads.

Gene family	Properties
GPAL1-9	All full length or nearly full length. *GPAL2* and *GPAL3* induced by *GMYB10*, others not.
GC4H1-7	Only *GC4H1* and *GC4H2* are full length. *GC4H1* is induced by *GMYB10*, *GC4H2* not. Others not well expressed.
G4CL1-14	At least seven are full length. Only *G4CL1* and *G4CL2* are induced by *GMYB10*.

GCHS1, 3 **and** *4*	All are full length. *GCHS4* and *GCHS3* are induced by *GMYB10*.
GCHI1-3	*GCHI1* is the only one with strong expression and it is induced by *GMYB10*.
GFNSII	Strong expression in petals, not induced by *GMYB10*.
GF3H1-2	Only *GF3H1* has strong expression and is induced by *GMYB10*.
GF3'H1	Strong expression in petals, induced by *GMYB10*.
GFLS1	Expressed in petals, not induced by *GMYB10*.
GDFR1	Expressed in petals and induced by *GMYB10*.
GANS1	Expressed in petals and induced by *GMYB10*.

4.4 Gerbera *CHS* (**II**)

In arabidopsis and snapdragon, *CHS* is represented by single genes (Sommer and Saedler, 1986; Burbulis *et al*., 1996). Gerbera has three *CHS* genes, *GCHS1* and *GCHS3* were reported by Helariutta *et al*. (1995) and *GCHS4* by Laitinen *et al*. (2008). Our RNAseq assembly did not uncover additional CHS encoding sequences, although new *CHS*-related sequences were identified that encode 2-pyrone synthases (Eckermann et al., 1998; Pietiäinen, Kontturi and Teeri, unpublished).

Helariutta *et al*. (1995) showed that *GCHS1* is abundant in petals and *GCHS3* in pappus bristles. Anti-*GCHS1* transgenic gerbera plants showed lack of anthocyanin pigmentation (Elomaa *et al*., 1996), indicating that *GCHS1* is a main contributor to anthocyanin biosynthesis in gerbera.

However, microarray analysis (Laitinen *et al*., 2008) and our RNAseq data (Table 4; Fig. S2) showed no significant changes in the expression level of *GCHS1* in Terraregina-MYB10 that overexpresses the anthocyanin regulator. Instead *GCHS4*, another petal abundant *CHS* gene, was strongly upregulated in this line (Fig. 1a in Paper II; Laitinen *et al*., 2008).

Virus-induced gene silencing (VIGS) had been recently developed for gerbera transient gene silencing (Deng *et al.*, 2012). It was possible to develop gene specific fragments for VIGS that resulted in down regulation of *GCHS1* and *GCHS4* separately. Only *GCHS1* downregulation led to loss of pigmentation in cultivar Terraregina.

Although petals that express *GCHS4* (but not *GCHS1*) are not able to synthesize anthocyanins, GCHS4 is not an active enzyme. *GCHS4* expression was found to be the dominant *CHS* in reddish vegetative tissues of Terraregina, which accumulate cyanidin (Fig. 5 in Paper II). Terraregina petals accumulate pelargonidin, so we wanted to test if GCHS4 is a cyanidin specific chalcone synthase. For this, we selected the cyanidin accumulating cultivar President for VIGS experiments. Like in Terraregina, only *GCHS1* down regulation led to loss pigmentation, *GCHS4* played no role.

Close investigation of the expression of *CHS* genes showed different patterns for the two genes through the developmental stages of the inflorescence of Terraregina (Fig. 4a in Paper II). *GCHS1* expression correlates with anthocyanin accumulation in petals. It starts at early stages when the petals are green, has highest expression level around stage 6, when anthocyanins accumulation is on its way, and then decreases during the later stages. Instead, *GCHS4* expression starts later and increases gradually until very late stages (Fig. 4a in Paper II). Clearly, *GCHS1* expression is more related to anthocyanin biosynthesis in petals than *GCHS4* expression is.

Nevertheless, *GMYB10* upregulates *GCHS4*, and in transgenic lines where *GMYB10* is downregulated (35S-anti-*GMYB10*), the normally red petioles of Terraregina leaves are green (Fig. 7a in Paper II). The petals of Terraregina-MYB10 (the overexpressing line) accumulate more anthocyanins than nontransgenic lines and, particularly, petals accumulate cyanidin that nontransgenic lines do not (Laitinen *et al.*, 2008). Interestingly, only the normally unpigmented parenchymal tissue in petals accumulates cyanidin, the epidermal layer of Terraregina-MYB10 has only pelargonidin (Fig. 5). It is peculiar that Terraregina petal epidermi cannot accumulate cyanidin, even when *GF3'H* is upregulated by *GMYB10*. A possible dominant pelargonidin specifying

genetic factor described by Tyrach and Horn (1997) is discussed in more detail in Paper III.
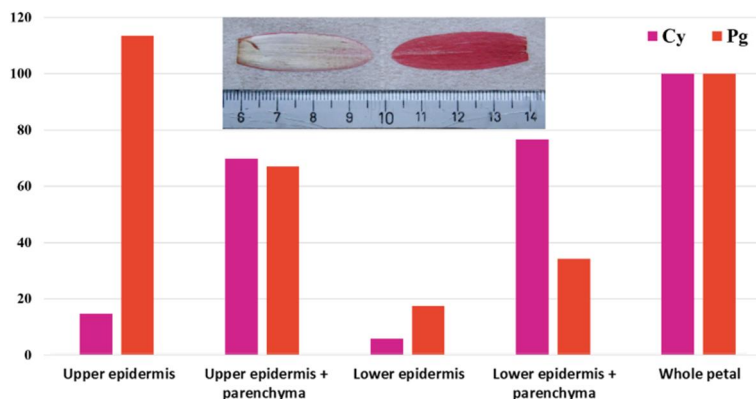
Figure 5: Anthocyanin accumulation in petal layers of Terraregina-MYB10. Most of the Pg-type anthocyanins accumulate in the upper (adaxial) epidermis, while Cy-type anthocyanins accumulate in the parenchyma tissues. Cy and Pg contents were scaled to whole petal content. In set shows a petal of Terraregina peeled for its upper epidermis (right).

4.5 Gerbera *DFR* (III)

The first gerbera DFR encoding cDNA (*GDFR1*) was isolated and characterized from Terraregina by Helariutta *et al*. (1993). *GDFR1* is expressed in both epidermal layers of petals and upregulated in Terraregina-MYB10 (Laitinen *et al*., 2008; Fig. S2).

The acyanic cultivar Ivory is a sport of the pelargonidin containing pink cultivar Estelle, i.e., it originates from an acyanic branch of Estelle (Laitinen, 2006). Ivory gives the impression of being a transposon mutant of Estelle, since revertant sectors are regularly observed (Fig. 6). Still, RNA gel blotting (Paula Elomaa, unpublished), microarray experiments (Laitinen, 2006) and RNAseq data generated in this work (Paper III) showed that there were no transcripts, flavonoid related or not, that were differently expressed in Ivory compared to Estelle. In addition, no unusual rearrangements that would indicate presence of transposons in Ivory transcripts were observed.

Figure 6: Gerbera cultivars Estelle and Ivory. The cultivar Ivory with revertant sectors in this picture (right) originates from a mutant branch of Estelle (left).

Flavonoid profiles were investigated in Estelle and Ivory in order to identify the block in the pathway to anthocyanins in Ivory. Samples were collected from whole petals and petal upper epidermi at stage 5 and 7. As expected, the results showed that pelargonidin derivatives were missing in Ivory. While apigenin was present at the same level in both cultivars, kaempferol levels were higher in Ivory than in Estelle (Fig. 3 in Paper III). This indicated that the flavonoid pathway has a block in Ivory after the stage of DHK formation, that is, in the late stages of the pathway (Fig. 6 in Paper I).

RNAseq showed the presence of three *DFR* alleles in the cultivars sequenced (Terraregina, Estelle and Ivory). The *GDFR1* sequence published by Helariutta *et al.* (1993) was assigned as *GDFR1-1*. In addition to *GDFR1-1*, Terraregina petals expressed another, 99% identical allele, *GDFR1-2*, and Estelle expressed *GDFR1-2* and *GDFR1-3* (98% identity to *GDFR1-2* and *GDFR1-1*). Ivory carries a single base pair mutation in *GDFR1-2*, the mutant allele assigned as *GDFR1-2m*. The mutation causes an amino acid change at the substrate and NADPH binding site of the enzyme, resulting in loss of activity (Fig. 5 in Paper III) and causing Ivory to carry white petals. What is interesting is that the white Ivory petals still express *GDFR1-3*, an enzyme that shows activity when expressed in *E. coli* and, in fact, Ivory petal extracts show DFR activity of about half of that of Estelle petals. Apparently, *GDFR1-3* cannot contribute to anthocyanin biosynthesis in Ivory – this anthocyanin would be pelargonidin.

GDFR1-3 was found also in cultivar President, with 98% nucleotide identity and encoding an enzyme with 100% amino acid sequence identity with the Estelle/Ivory GDFR1-3. Based on sequencing of eight DFR encoding cDNA molecules from President, *GDFR1-3* seems to be the only allele in this cultivar. President accumulates cyanidin in its petals (Papers I, II, III). The difference between cyanidin, compared to pelargonidin, is in hydroxylation of the B-ring, evoked by the enzyme F3'H (Fig. 2). We saw already in section 4.2 that application of the F3'H inhibitor TET during development of petals in cultivar President does not lead to pelargonidin in place of cyanidin, instead the petals turn acyanic (Fig. 4).

The different GDFR1 enzymes (encoded by the three different alleles) do not show strong substrate preference for DHK vs. DHQ *in vitro* (Paper III). It is possible that they have specificity *in vivo* or, alternatively, pelargonidin and cyanidin are synthesized in completely separate enzyme complexes, metabolons.

4.6 Nested analysis and agroinfiltration (IV)

Production of eukaryotic proteins in bacterial hosts has often problems due to low production levels, protein folding or protein toxicity (Frommer and Ninnemann, 1995). Even when proteins are produced they may miss post-translational modifications. A eukaryotic production host is therefore preferred, such as yeast or, for plant proteins, plant cells that can be easily transformed. DFR from gerbera and other species has been produced in tobacco protoplast to investigate their enzymatic activity (Martens *et al*., 2002). In this work we used agroinfiltration instead of protoplasts for heterologous protein production.

Syringe agroinfiltration, a method of transient *Agrobacterium*-mediated DNA transfer, was used to produce enzymatically active GDFRs in the leaves of *Nicotiana benthamiana* (Paper III). We observed sometimes variation in agroinfiltration, and as part of the experiments reported here, wanted to address the source of the variation. In order to dissect the variance into its components, we used a hierarchical or nested experimental design (Krzywinski *et al*., 2014; Tichopad *et al*., 2009) (Paper IV).

Using firefly luciferase as a reporter, we showed that the main variation is between individual sampling spots of the agroininfiltrated leaf (Fig. 3 in Paper IV). Second largest source is position of the leaf on the plant, but variation between individual plants or repetitions of the experiment was small.

5.   CONCLUSIONS AND PROSPECTS

Gerbera (my favourite plant) was used as a model to study flavonoid metabolites in the Asteraceae family. We wanted to understand how the plant controls the branched pathway to flavonoids and analysed in detail the transcripts, enzymes and flavonoid metabolites in different gerbera cultivars. We found that although flavones, flavonols and anthocyanins can be differently hydroxylated, this can be explained by temporally regulated presence of F3'H activity during petal development. As has been shown many times before, transcript levels do not always reflect enzymatic activity in the cell. F3'H activity could not be shown during the stages of pelargonidin accumulation in cultivar Terraregina although transcript levels were rising. Similarly, we could show that transcripts for an active CHS enzyme (GCHS4) may accumulate in petals, but no anthocyanins are made.

Even enzymatic activity (defined in tissue extracts) does not always reflect the biosynthetic capacity of the cell. For reasons not very well understood yet, the cell may not conduct biosynthesis even when all enzymatic activities are present, as shown for gerbera DFR. One possibility is that enzymes may exhibit substrate specificities *in vivo* differently from activities measured *in vitro*. For the allelic forms of *GDFR1*, complementation studies in petunia may provide insight to this question.

Alternatively, the different DFR enzymes of gerbera may participate in different metabolons, a hypothesis often raised to explain this kind of phenomena, but difficult to prove in practise. Plant cells may also contain factors participating in metabolite (in our case anthocyanin) biosynthesis that have not been identified before. Generation of RNAseq data of Cy- and Pg/Cy-type gerbera cultivars, in comparison to the Pg-type cultivars sequenced in this work, may shed light to this hypothesis in the future.

6. SUPPLEMENTAL DATA

Figure S1. The developmental stages of Terraregina inflorescences, as described by Helariutta *et al*. (1993). The figure was prepared by my co-author Dr. Xianbao Deng.

GPAL family

GPAL1

GPAL2

GPAL3

GC4H family

GC4H1

GC4H2

G4CL family

G4CL1

G4CL2
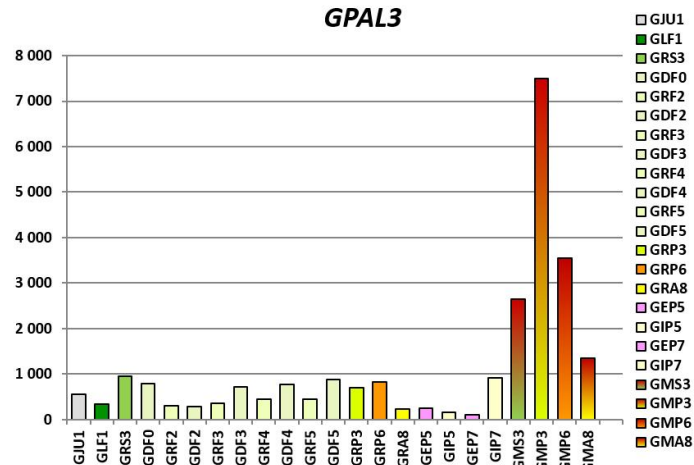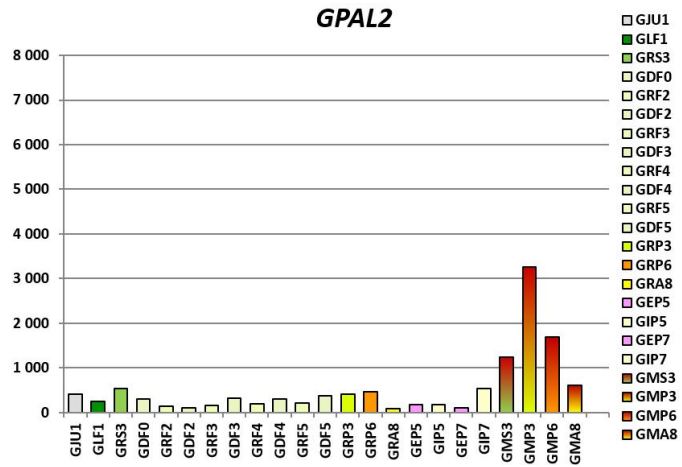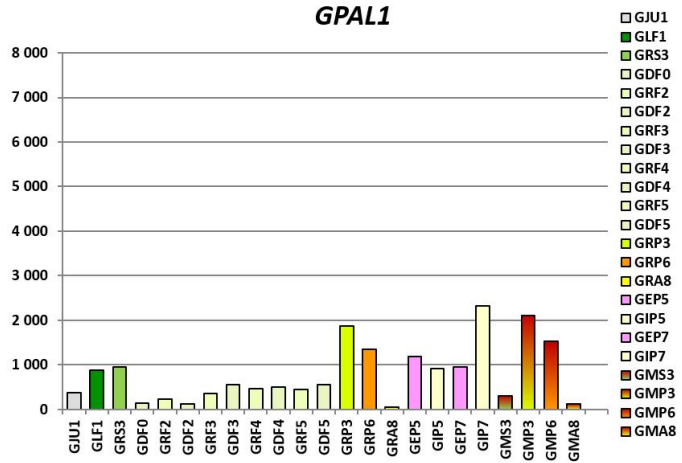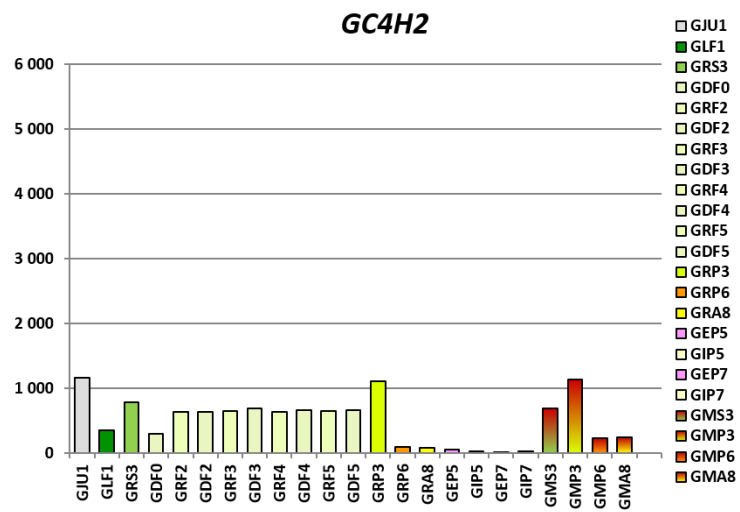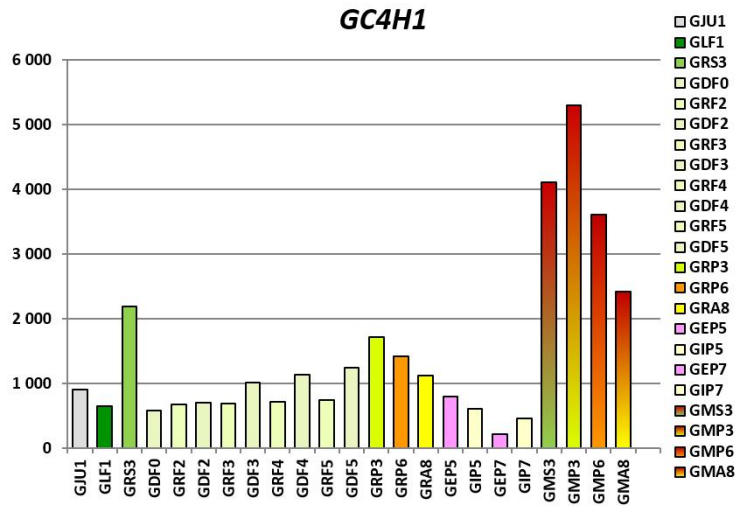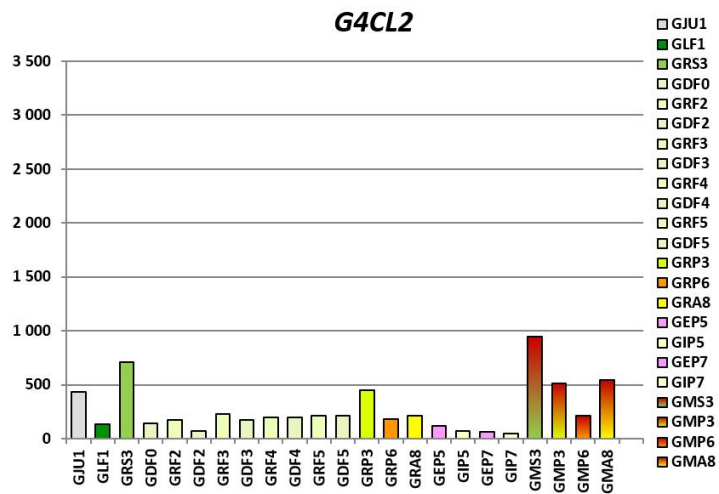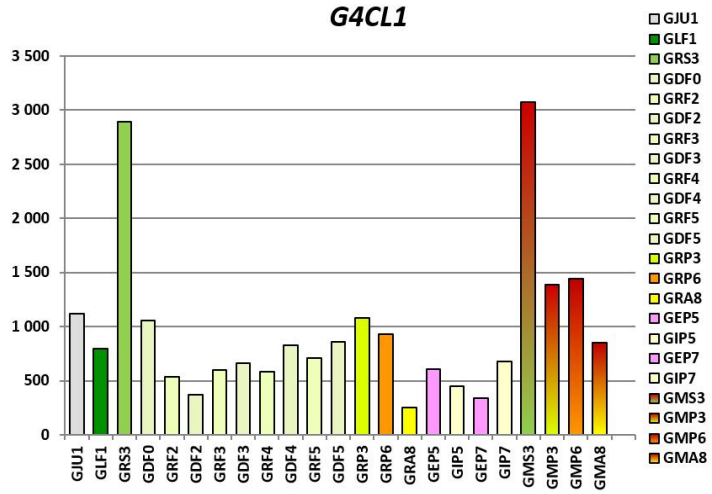
GCHS family
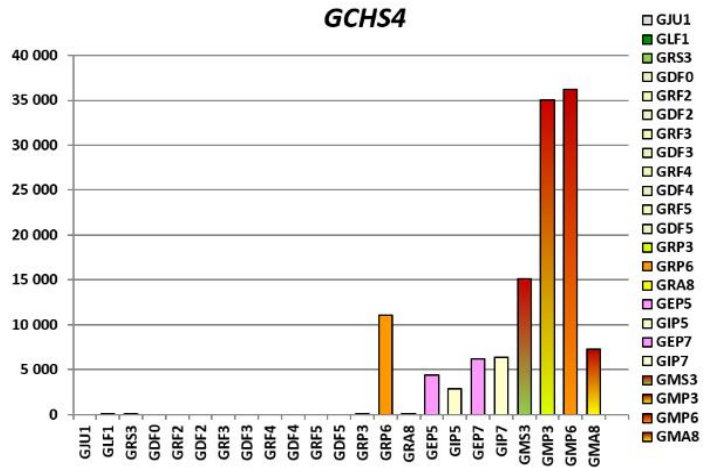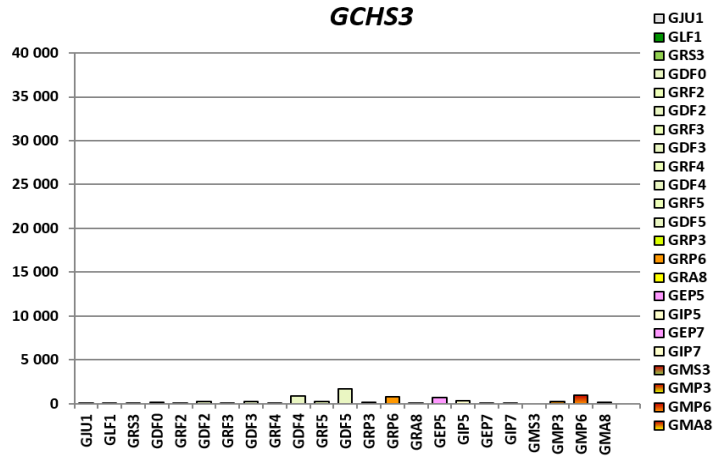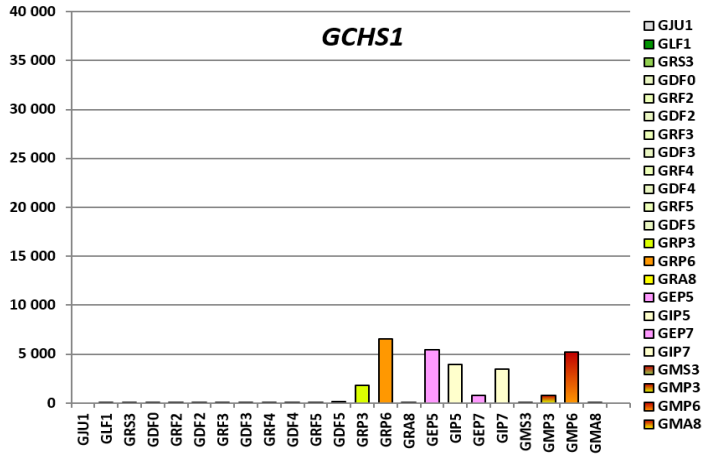
GCHI family
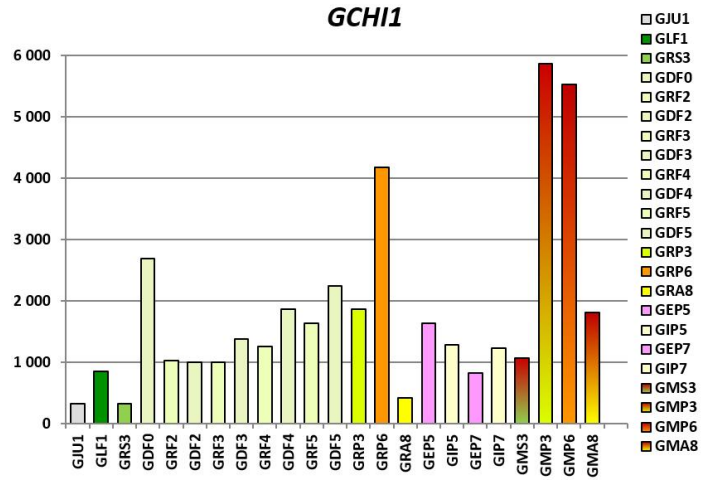
GFNS II family

GF3H family

GF3H1

GF3H2

GF3'H family

GF3'H1

GFLS family

GFLS1

46
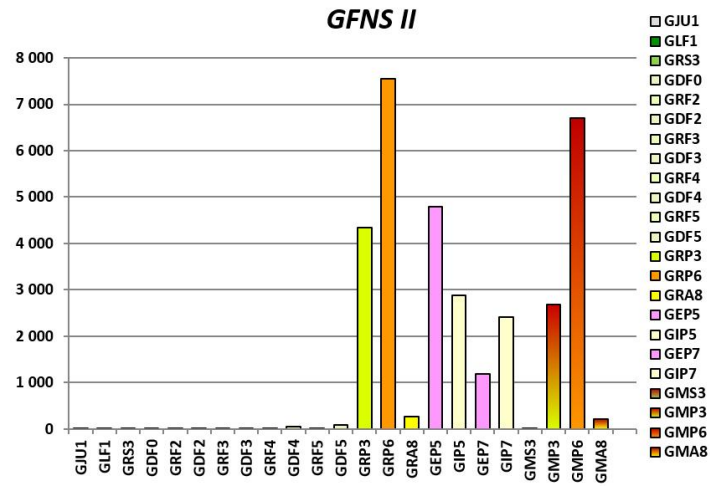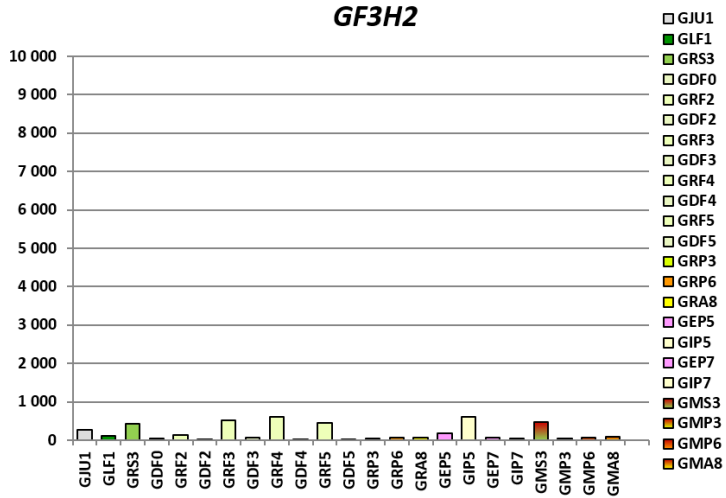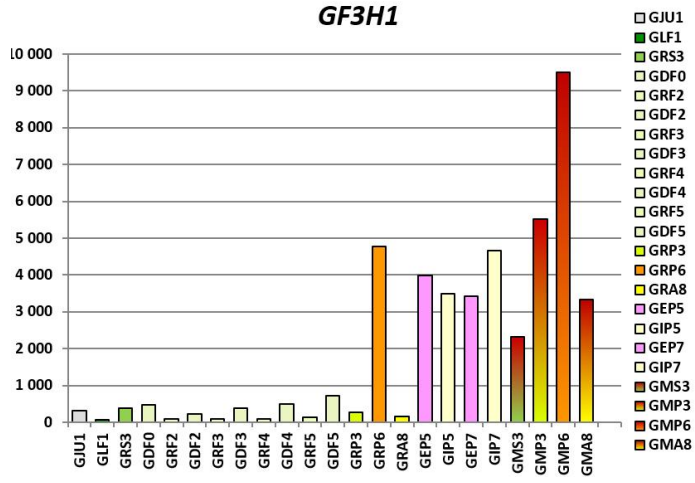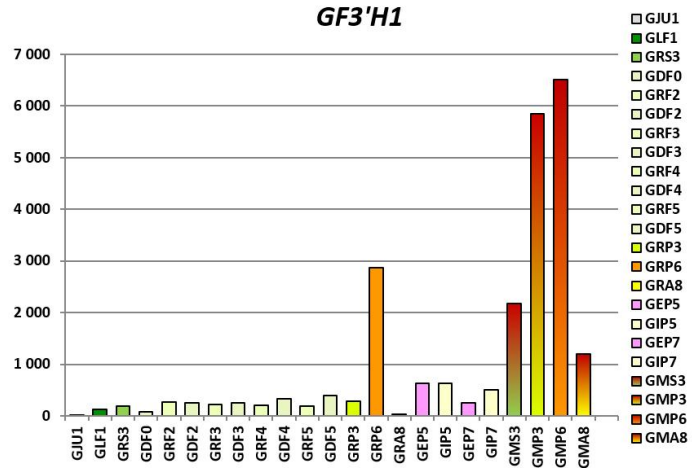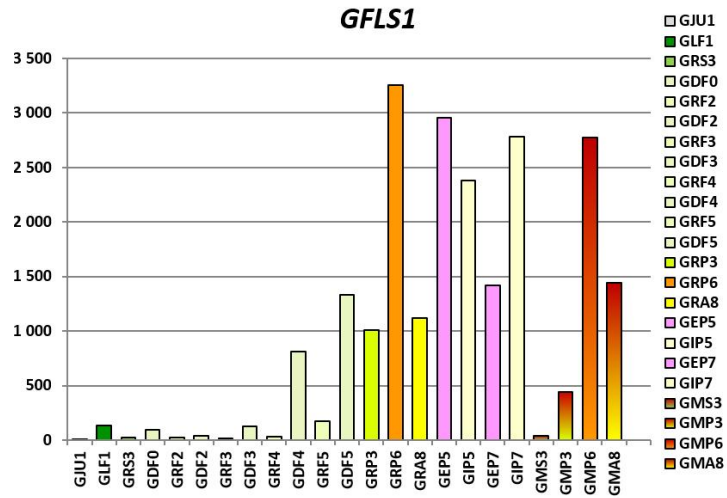
GDFR family
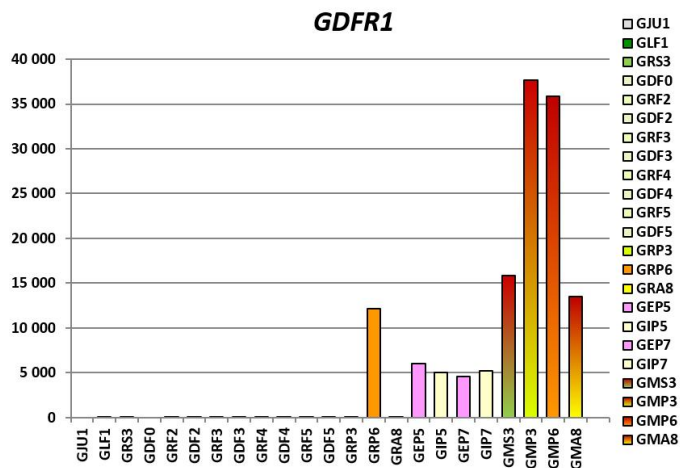
GANS family
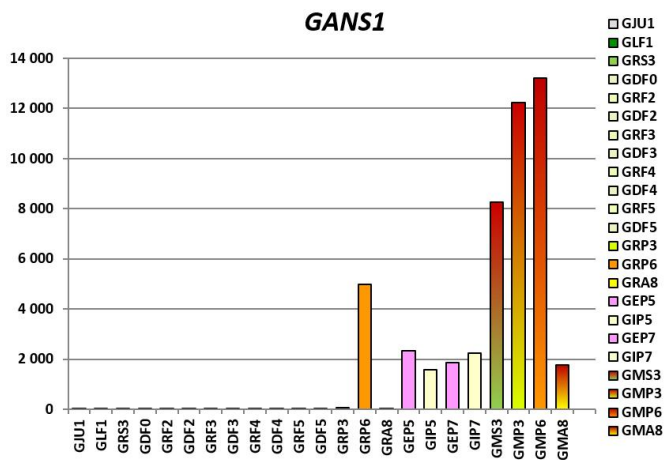
Figure S2. Expression of gerbera Flavonoid biosynthesis genes in RNAseq libraries representing different tissues and developmental stages. The score on y-axis shows the number of reads matching the target gene at stringency 1e-50, normalized to target length and library size, finally scaled with average of library sizes. Libraries are abbreviated according to the following table:

Code	Sample	Cultivar
GJU1	Root	Terraregina
GLF1	Leaf mix of young and old leaves	Terraregina
GRS3	Scape stage 3	Terraregina
GDF0	Disc Flower stage 0	Terraregina
GRF2	Ray Flower stage 2	Terraregina
GDF2	Disc Flower stage 2	Terraregina
GRF3	Ray Flower stage 3	Terraregina
GDF3	Disc Flower stage 3	Terraregina
GRF4	Ray Flower stage 4	Terraregina
GDF4	Disc Flower stage 4	Terraregina
GRF5	Ray Flower stage 5	Terraregina
GDF5	Disc Flower stage 5	Terraregina
GRP3	Petal stage 3	Terraregina
GRP6	Petal stage 6	Terraregina
GRA8	Anther stage 8	Terraregina
GEP5	Petal stage 5	Estelle
GIP5	Petal stage 5	Ivory
GEP7	Petal stage 7	Estelle
GIP7	Petal stage 7	Ivory
GMS3	Scape stage 3	Terraregina-MYB10
GMP3	Petal stage 3	Terraregina-MYB10
GMP6	Petal stage 6	Terraregina-MYB10
GMA8	Anther stage 8	Terraregina-MYB10

Ulla-Maija; green house: Sanna, Sini, Terttu and Jouko (kind man). Also, kiitos our active secretaries: Sari and Riikka, Laura and Seija Oikarinen. Thanks to Kaj-Roger (Radioactive man), the reception lady: Anne (smiling lady).

I would like to thank the breeding company Terra Nigra B.V. (The Netherlands) for supplying the gerbera cultivars, Dr. Lieven Sterck at Ghent University for helping in RNAseq data analysis and Paula Kristo (Haartman Institute Sequencing Unit) for her nice help.

I know Karen (Dr. Karen Sims-Huopaniemi) looking for her name. Everyone should to thank our nice, lovely and active lady and leaders of plant science program, Professors Tapio and Jaakko. Also, I did not forget to thank our YEB coordinator Karin Hemmann and our school leader Prof. Jari for arranging and funding a nice courses and meetings.

I would like to thank all my friends in Finland, Kiitos: Dr. A. Geneid, Taha Hussein, Osman, Yasser, M. Hamdy, M. Löst, E. Botros, Maiser, Maged, Saad, Ismael, El-Assal, M. Hussein, Mahmoud Hassan, Rayan, Hossam Elhusseiny, M. Adel, Ayman Bayoumi, Mohammed Ali, Khaled, Qasem, Zuhdi, Yehia, Mahmoud, Drs: Fahim, Nader, M. Soliman and Almgary.

Also, thanks to my colleagues and friends in Egypt: All genetics staff professors and doctors Abdel Kader, Sawsan, Ebtissam, M.Ali, M. Hussain, Mona, Reda, Samir, Basita, Elassal, Haggag, Gaber, El Awady, Hady, Sherine, Engy, El Soda, Ahmed, Heba, Nagya, Dalia, Naima, Sara, Hanaa, Samad, Mai… and AGRI group: Drs. Gihan, Ismail, Attia, Hosam, Dessouky, Shafik and Tarik, Ayob. And professor Rafic (Veterinary medicine).

Thanks to my wonderful and lovely persons in my life: my mother, my father, my wife, my kids Salma, Ahmed, Mohamed and my brothers and sisters, Also my Friends: Eng. M. Nasr, M.Taha, Eng. Nabil, Dr. Galal, Dr. Emad, Dr. Abdallah Soubhy, Dr. Ramadan Bakr, Sayed, Ekram, Dr. Wahid Awad, El Hadidi and Khalid Hashim.

الحمد لك الحمد لله رب العالمين... الحمد لله تتم الصالحات.... والحمد لله على نعمة الحمد

كل الشكر لكى يا أمى رحمة الله عليكى فى الفردوس الأعلى, كنت اتمنى حضورك يا أمى ولكن إنت فى أفضل مكان. بارك الله فيك يا أبى وجزاك واعطاك حتى ترضى من خير الدنيا والأخرة. بارك الله لأخواتى و أحبائى فى مصر : هناء و خلف و أبو سمرة وعمور و أبو حميد وحوده و محمد و سارة و كذلك المهندس رمضان* والأستاذ عبد الله والبحيرى و علاء وحسوبه وحسبو و رمضان و سلوى و شويكار و داليا و خيرات و غزال و تحياتى و شكرى للجميع من تذكرت و لم اتذكره وسامحونى.

شكر خاص لزوجتى و اولادى سلمى وأحمد و محمد اعلم كم كنت مقصر معكم فأنتم أفضل ثمار حياتى بارك الله فيكم و هداكم و اصلح حالكم دنيا و دين.

Kiitos isäni, äitini, minun ihana perhe, THT ja everyone may I forget...forgive me
Hany Bashandy 2016   هانى بشندى 2016

REFERENCES

Abe I, Morita H. 2010. Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. Natural Product Reports 27:809-838.

Albert N, Davies KM, Lewis D, Zhang H, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE. 2014. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. The Plant Cell 26:962-980.

Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W, Lipman D.1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research 25:3 389-3 402.

Appelhagen I, Jahns O, Bartelniewoehner L, Sagasser M, Weisshaar B, Stracke R. 2011. Leucoanthocyanidin dioxygenase in *Arabidopsis thaliana*: Characterization of mutant alleles and regulation by MYB–BHLH–TTG1 transcription factor complexes. Gene 484:61-68.

Asen S. 1984. High pressure liquid chromatographic analysis of flavonoid chemical markers in petals from *Gerbera* flowers as an adjunct for cultivar and germplasm identification. Phytochemistry 11:2523-2526.

Austin MB, Izumikawa M, Bowman ME, Udwary DW, Ferrer J-L, Moore BS, Noel JP. 2004. Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. The Journal of Biological Chemistry 279:45162-45174.

Austin MB, Noel JP. 2003. The chalcone synthase superfamily of Type III polyketide synthases. Natural Product Reports 20:79-110.

Bartholomae M, Meyer FM, Commichau FM, Burkovski A, Hillen W, Seidel G. 2014. Complex formation between malate dehydrogenase and isocitrate dehydrogenase from *Bacillus subtilis* is regulated by tricarboxylic acid cycle metabolites. Federation of European Biochemical Societies Journal 281:1132-1143.

Bassard J-E, Richert L, Geerinck J, Renault H, Duval F, Ullmann P, Schmitt M, Meyer E, Mutterer J, Boerjan W, Jaeger GD, Mely Y, Goossens A, Werck-Reichharta D. 2012. Protein-Protein and Protein-Membrane associations in the lignin pathway. The Plant Cell 24:4465-4482.

Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C. 1997. Cinnamate 4-hydroxylase expression in arabidopsis regulation in response to development and the environment. Plant Physiology 113:729-738.

Bhagwat S, Haytowitz DB, Holden JM (Ret.). 2013. USDA database for the flavonoid content of selected foods, release 3.1. U.S. department of

agriculture, agricultural research service. Nutrient data laboratory home page: http://www.ars.usda.gov/nutrientdata/flav

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry 72:248-254.

Bravo L. 1998. Polyphenols: Chemistry, dietary sources, metabolism, andnutritional significance. Nutrition Reviews 56:317-333.

Britsch L, Heller W, Grisebach H. 1981. Conversion of flavanone to flavone, dihydroflavonol and flavonol with an enzyme system from cell cultures of *Parsley*. Zeitschrift für Naturforschung C 36:742-750.

Brown DE, Rashotte AM, Murphy AS, Normanly J, Tague BW, Peer WA, Taiz L, Muday GK. 2001. Flavonoids act as negative regulators of auxin transport *in vivo* in arabidopsis. Plant Physiology 126:524-535.

Brugliera F, Tao G-Q, Tems U, Kalc G, Mouradova E, Price K, Stevenson K, Nakamura N, Stacey I, Katsumoto Y, Tanaka Y, Mason JG. 2013. Violet/Blue chrysanthemums- metabolic engineering of the anthocyanin biosynthetic pathway results in novel petal colors. Plant Cell and Physiology 54:1696-1710.

Burbulis I, Winkel-Shirley B. 1999. Interactions among enzymes of the arabidopsis flavonoid biosynthetic pathway. Proceedings of the National Academy of Sciences 96:12929-12934.

Burbulis IE, Iacobucci M, Shirley BW. 1996. A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in arabidopsis. The Plant Cell 8: 1013-1025.

Butelli E, Titta L, Giorgio M, Mock H-P, Matros A, Peterek S, Schijlen EG, Hall RD, Bovy AG, Luo J, Martin C. 2008. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nature Biotechnology 26:1301- 1308.

Calabrese JC, Jordan DB, Boodhoo A, Sariaslani S, Vannelli T. 2004. Crystal structure of phenylalanine ammonia lyase: multiple helix dipoles implicated in catalysis. Biochemistry 43:11403-11416.

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.

Cheng A-X, Han X-J, Wu Y-F, Lou H-X. 2014. The function and catalysis of 2-oxoglutarate-dependent oxygenases involved in plant flavonoid biosynthesis. International Journal of Molecular Sciences 15:1080-1095.

Cochrane FC, Davin LB, Lewis NG. 2004. The *Arabidopsis phenylalanine ammonia lyase* gene family: kinetic characterization of the four PAL isoforms. Phytochemistry 65:1557-1564.

Crosby KC, Pietraszewska-Bogiel A, Gadella TWJ, Winkel BSJ. 2011. Förster resonance energy transfer demonstrates a flavonoid metabolon in living plant cells that displays competitive interactions between enzymes. Federation of European Biochemical Societies Letters 585:2193-2198.

Dao TTH, Linthorst HJM, Verpoorte R. 2011. Chalcone synthase and its functions in plant resistance. Phytochemistry Reviews 10:397-412.

Dare AP, Tomes S, Jones M, McGhie TK, Stevenson DE, Johnson RA, Greenwood DR, Hellen RP. 2013. Phenotypic changes associated with RNA interference silencing of *chalcone synthase* in apple (*Malus × domestica*). The Plant Journal 74:398-410.

Davies KM, Schwinn KE, Deroles SC, Manson DG, Lewis DH, Bloor SJ, Bradley J. 2003. Enhancing anthocyanin production by altering competition for substrate between flavonol synthase and dihydroflavonol 4-reductase. Euphytica 131:259-268.

Deng X, Elomaa P, Nguyen CX, Hytönen T, Valkonen JPT, Teeri TH. 2012. Virus-induced gene silencing for Asteraceae - a reverse genetics approach for functional genomics in *Gerbera hybrida*. Plant Biotechnology Journal 10:970-978.

DeVetten 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 and Development. 11:1422-1434.

Diesperger H, Müller C, Sandermann, H. 1974. Rapid isolation of a plant microsomal fraction by Mg^{2+}-precipitation. Federation of European Biochemical Societies Letters 43:155-158.

Dixon RA, Lamb CJ, Masoud S, Sewalt VJH, Paiva NL. 1996. Metabolic engineering: prospects for crop improvement through the genetic manipulation of phenylpropanoid biosynthesis and defense responses. Gene 179:61-71.

Dixon RA, Liu C, Jun JH. 2013. Metabolic engineering of anthocyanins and condensed tannins in plants. Current Opinion in Biotechnology 24:329-335.

Dixon RA, Steele CL.1999. Flavonoids and isoflavonoids - a gold mine for metabolic engineering. Trends in plant science 4:394-400.

Dooner IK, Robbins TP, Jorgensen RA. 1991. Genetic and developmental control of anthocyanin biosynthesis. Annual Review of Genetics 25:173-199.

Eckermann S, Schröder G, Schmidt J, Strack D, Edrada RA, Helariutta Y, Elomaa P, Kotilainen M, Kilpeläinen I, Proksch P, Teeri TH, Schröder J. 1998. New pathway to polyketides in plants. Nature 396:387-390.

Elomaa P, Helariutta Y, Kotilainen M, Teeri TH.1996. Transformation of antisense constructs of the *chalcone synthase* gene superfamily into *Gerbera hybrida*: differential effect on the expression of family members. Molecular Breeding 2:41-50.

Elomaa P, Honkanen J, Puska R, Seppänen P, Helariutta Y, Mehto M, Kotilainen M, Nevalainen L, Teeri TH. 1993. *Agrobacterium*-mediated transfer of antisense *chalcone synthase* cDNA to *Gerbera hybrida* inhibits flower pigmentation. Biotechnology 11:508-511.

Elomaa P, Mehto M, Kotilainen M, Helariutta Y, Nevalainen L, Teeri TH. 1998. A bHLH transcription factor mediates organ, region and flower type specific signals on *dihydroflavonol-4-reductase* (*dfr*) gene expression in the inflorescence of *Gerbera hybrida* (Asteraceae). The Plant Journal 16:93-99.

Elomaa P, Uimari A, Mehto M, Albert VA, Laitinen RAE, 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 Physiology 133:1831-1842.

Enright A, Van Dongen S, Ouzounis C. 2002. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Research 30: 1575-1584.

Falginella L, Castellarin SD, Testolin R, Gambetta GA, Morgante M, Gaspero GD. 2010. Expansion and subfunctionalisation of flavonoid 3',5'-hydroxylases in the grape vine lineage. BioMed Central Genomics 11:1-18.

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

Ferrer J-L, Austin MB, Stewart CJ, Noel JP. 2008. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiology and Biochemistry 46:356-370.

Flores-Sanchez IJ, Verpoorte R. 2008. PKS activities and biosynthesis of cannabinoids and flavonoids in *Cannabis sativa* L. plants. Plant Cell Physiology 49:1767-1782.

Forkmann G, Martens S. 2001. Metabolic engineering and applications of flavonoids. Current Opinion in Biotechnology12:155-160.

Forkmann G, Ruhnau B. 1987. Distinct substrate specificity of dihydroflavonol-4-reductase from flowers of *Petunia hybrida*. Zeitschrift für Naturforschung C 42:1146-1148.

Fraenkel GS. 1959. The raison d'Etre of secondary plant substances. Science 129:1466-1470.

Fritsch and Grisebach 1975. Biosynthesis of cyanidin in cell cultures of *Haplopappus gracilis*. Phytochemistry 14: 2437-2442.

Frommer W, Ninnemann O. 1995. Heterologous expression of genes in bacterial, fungal, animal, and plant cells. Annual Review of Plant Physiology and Plant Molecular Biology 46:419-444.

Funk VA, Bayer RJ, Keeley S, Chan R, Watson L, Gemeinholzer B, Schilling E, Panero JL, Baldwin BG, Garcia-Jacas N, Susanna A, Jansen RK. 2005. Everywhere but Antarctica: Using a super tree to understand the diversity and distribution of the Compositae. Biologiske Skrifter 55:343-374.

Gensheimer M, Mushegian A. 2004. Chalcone isomerase family and fold: No longer unique to plants. Protein Science 13:540-544.

Gollop R, Farhi S, Perl A. 2001. Regulation of the leucoanthocyanidin dioxygenase gene expression in *Vitis* v*inifera*. Plant Science 161:579-588.

Gontero B, Cardenas ML, Ricard J. 1988. A functional five-enzyme complex of chloroplasts involved in the Calvin cycle. European Journal of Biochemistry 173:437-443.

Grotewold E, Chamberlin M, Snook M, Siame B, Butler L, Swenson J, Maddock S, Clair G, Bowen B. 1998. Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. The Plant Cell 10:721-740.

Halbwirth H. 2010. The creation and physiological relevance of divergent hydroxylation patterns in the flavonoid pathway. International Journal of Molecular Sciences 11: 595-621.

Hamberger B, Hahlbrock K. 2004. The 4-coumarate: CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate activating and three commonly occurring isoenzymes. Proceedings of the National Academy of Sciences 101:2209-2214.

He H, Ke H, Keting H, Qiaoyan X, Silan D. 2013. Flower colour modification of chrysanthemum by suppression of *F3'H* and overexpression of the exogenous *Senecio cruentus F3'5'H* gene. PLOS ONE 8: e74395.

Helariutta Y, Elomaa P, Kotilainen M, Griesbach R, Schrder J, Teeri TH.1995. Chalcone synthase-like genes active during corolla development are

differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (Asteraceae). Plant Molecular Biology 28:47–60.

Helariutta Y, Elomaa P, Kotilainen M, Seppänen P, Teeri TH. 1993. Cloning of cDNA coding for dihydroflavonol-4-reductase (DFR) and characterization of *dfr* expression in the corollas of *Gerbera hybrida* var.Regina (Compositae). Plant Molecular Biology 22:183-193.

Heller W, Forkmann G, Britsch L, Grisebach H. 1985. Enzymatic reduction of (+)-dihydroflavonols to flavan-3,4-cis-diols with flower extracts from *Matthiola incana* and its role in anthocyanin biosynthesis. Planta 165:284-287.

Herrmann KM. 1995. The shikimate pathway as an entry to aromatic secondary metabolism. Plant Physiology. 107:7-12.

Hirsch AM, Lum MR, Downie JA. 2001. What makes the *Rhizobia*-legume symbiosis so special? Plant Physiology 127:1484-1492.

Hulme AC. 1953. The isolation of chlorogenic acid from the apple fruit. Biochemical Journal 53: 337-340.

Jackson D, Roberts K, Martin C. 1992. Temporal and spatial control of expression of anthocyanin biosynthetic genes in developing flowers of *Antirrhinum majus*. The Plant Journal 2:425-434.

Johnson ET, Ryu S, Yi H, Shin B, Cheong H, Choi G. 2001. Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. The Plant Journal 25:325-333.

Johnson ET, Yi H, Shin B, Oh B-J, Cheong H, Choi G. 1999. *Cymbidium hybrida* dihydrofavonol 4-reductase does not efficiently reduce dihydrokaempferol to produce orange pelargonidin-type anthocyanins. The Plant Journal 19:81-85.

Johzuka-Hisatomi Y, Hoshino A, Mori T, Habu Y, Iida S. 1999. Characterization of the chalcone synthase genes expressed in flowers of the common and Japanese morning glories. Genes and Genetic Systems 74:141-147.

Katsumoto Y, Fukuchi-Mizutani M, Fukui Y, Brugliera F, Holton TA, Karan M, Nakamura N, Yonekura-Sakakibara K, Togami J, Tao A, Nehra NS, Lu C-Y, Dyson BK, Tsuda S, Ashikari T, Kusumi T, Mason JG, Tanaka Y. 2007. Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating delphinidin. Plant and Cell Physiology 48:1589-1600.

Keung WM. 2003. Anti-dipsotropic isoflavones: The potential therapeutic agents for alcohol dependence. Medicinal Research Reviews 23:669-696.

Kleindt C, Stracke R, Mehrtens F, Weisshaar B. 2010. Expression analysis of flavonoid biosynthesis genes during *Arabidopsis thaliana* silique and seed development with a primary focus on the proanthocyanidin biosynthetic pathway. BioMed Central Research Notes, 3:255-255.

Kobayashi S, Ishimaru M, Ding CK, Yakushiji H, Goto N. 2001. Comparison of *UDP-glucose*: *flavonoid 3-O-glycosyl transferase* (*UFGT*) gene sequences between white grapes (*Vitis* v*inifera*) and their sports with red skin. Plant Science 160:543-550.

Kodan A, Kuroda H, and Sakai F. 2002. A stilbene synthase from Japanese red pine (*Pinus densiflora*): Implications for phytoalexin accumulation and down-regulation of flavonoid biosynthesis. Proceedings of the National Academy of Sciences 99:3335-3339.

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

Koes RE, Spelt CE, van den Elzen PJ, Mol JN. 1989. Cloning and molecular characterization of the *chalcone synthase* multigene family of *Petunia hybrida*. Gene 81:245-57.

Koukol J, Conn EE. 1961. The metabolism of aromatic compounds in higher plants. IV. Purification and properties of the phenylalanine deaminase of *Hordeum vulgare*. The Journal of Biological Chemistry 236:2692-2698.

Krzywinski M, Altman N, Blainey P. 2014. Nested designs. Nature Methods 11:977-978.

Kähkönen MP, Heinämäki J, Ollilainen V, Heinonen M. 2003. Berry anthocyanins: isolation, identification and antioxidant activities. Journal of the Science of Food and Agriculture 83:1403-1411.

Laitinen RAE, Ainasoja M, Broholm SK, Teeri TH, Elomaa P. 2008. Identification of target genes for a MYB-type anthocyanin regulator in *Gerbera hybrida*. Journal of Experimental Botany 59:3691-3703.

Laitinen RAE. 2006. The Gerbera cDNA Microarray: A Tool for Large-Scale Identification of Genes Involved in Flower Development. PhD thesis, University of Helsinki, http://urn.fi/URN:ISBN:952-10-3403-3

Lallemand LA, Zubieta C, Lee SG, Wang Y, Acajjaoui S, Timmins J, McSweeney S, Jez JM, McCarthy JG, McCarthy AA. 2012. A structural basis for the biosynthesis of the major chlorogenic acids found in coffee. Plant Physiology 160:249-260.

Leiss K, Maltese F, Choi Y, Verpoorte R, Klinkhamer P. 2009. Identification of chlorogenic acid as a resistance factor for thrips in chrysanthemum. Plant Physiology 150:1567-1575.

Liu C, Dixon R. 2001. Elicitor-induced association of isoflavone *O*-methyltransferase with endomembranes prevents the formation and 7-*O*-methylation of daidzein during isoflavonoid phytoalexin biosynthesis. The Plant Cell 13:2643-2643.

Lu S, Zhou Y, Li L, Chiang VL. 2006. Distinct roles of *cinnamate 4-hydroxylase* genes in *Populus*. Plant and Cell Physiology 47: 905-914.

Martens S, Forkmann G.1998. Genetic control of flavone synthase II activity in flowers of *Gerbera hybrids*. Phytochemistry 49:1953-1958.

Martens S, Forkmann G.1999. Cloning and expression of *flavone synthase II* from *Gerbera* hybrids. The Plant Journal 20:611-618.

Martens S, Knott J, Seitz CA, Janvari L, Yu S-N, Forkmann G. 2003. Impact of biochemical pre-studies on specific metabolic engineering strategies of flavonoid biosynthesis in plant tissues. Biochemical Engineering Journal 14:227-235.

Martens S, Teeri T, Forkmann G. 2002. Heterologous expression of *dihydroflavonol 4-reductase* from various plants. Federation of European Biochemical Societies Letters 531:453-458.

Mathews CK. 1993.The cell-bag of enzymes or network of channels? Journal of Bacteriology 175:6377-6381.

Menting J, Scopes RK, Stevenson TW. 1994. Characterization of flavonoid 3',5'-hydroxylase in microsomal membrane fraction of *Petunia hybrida* flowers. Plant Physiology 106: 633-642.

Meyer FM, Gerwig J, Hammer E, Herzberg C, Commichau FM, Völker U, Stülke J. 2011. Physical interactions between tricarboxylic acid cycle enzymes in *Bacillus subtilis*: Evidence for a metabolon. Metabolic Engineering 13:18-27.

Meyer P, Heidmann I, Forkmann G, Saedler H. 1987. A new petunia flower colour generated by transfrmation of a mutant with a maize gene. Nature 330:677-678.

Mlcek J, Rop O. 2011. Fresh edible flowers of ornamental plants - A new source of nutraceutical foods. Trends in Food Science and Technology 22: 561-569.

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

Morita Y, Saito R, Ban Y, Tanikawa N, Kuchitsu K, Ando T, Yoshikawa M, Habu Y, Ozeki Y, Nakayama M. 2012. Tandemly arranged *chalcone synthaseA*

genes contribute to the spatially regulated expression of siRNA and the natural bicolor floral phenotype in *Petunia hybrida*. The Plant Journal 70:739-749.

Morita Y1, Takagi K, Fukuchi-Mizutani M, Ishiguro K, Tanaka Y, Nitasaka E, Nakayama M, Saito N, Kagami T, Hoshino A, Iida S. 2014. A chalcone isomerase-like protein enhances flavonoid production and flower pigmentation. The Plant Journal 78:294-304.

Napoli C, Lemieux C, Jorgensen R. 1990. Introduction of a chimeric *chalcone synthase* gene into *Petunia* results in reversible co-suppression of homologous genes intrans. The Plant Cell 2:279-289.

Ngaki M, Louie G, Philippe R, Manning G, Pojer F, Bowman M, Li L, Larsen E, Wurtele E, Noel J. 2012. Evolution of the chalcone-isomerase fold from fatty-acid binding to stereospecific catalysis. Nature 485:530-536.

Noda N, Aida R, Kishimoto S, Ishiguro K, Fukuchi-Mizutani M, Tanaka Y, Ohmiya A. 2013. Genetic engineering of novel bluer-colored chrysanthemums produced by accumulation of delphinidin-based anthocyanins. Plant Cell Physiology 54:1684-1695.

Noel JP, Austin MB, Bomat EK. 2005. Structure–function relationships in plant phenylpropanoid biosynthesis. Current Opinion in Plant Biology 8:249-253.

Ogundiwin EA, Peace CP, Nicolet CM, Rashbrook VK, Gradzie TM, Bliss FA, Parfitt D, Crisosto CH. 2008. *Leucoanthocyanidin dioxygenase* gene (*PpLDOX*): a potential functional marker for cold storage browning in peach. Tree Genetics and Genomes 4:543-554.

Ohno S, Hosokawa M, Kojima M, Kitamura Y, Hoshino A, Tatsuzawa F, Doi M, Yazawa S. 2011. Simultaneous post-transcriptional gene silencing of two different *chalcone synthase* genes resulting in pure white flowers in the octoploid *Dahlia*. Planta 234:945-958.

Owens DK, Alerding AB, Crosby KC, Bandara AB, Westwood JH, Winkel BSJ. 2008. Functional analysis of a predicted *flavonol synthase* gene family in arabidopsis. Plant Physiology 147:1046-1061.

Paiva NL. 2000. An Introduction to the biosynthesis of chemicals used in plant-microbe Communication. The Journal of Plant Growth Regulation 19:131-143.

Park JS, Kim JB, Cho KJ, Cheon CI, Sung MK, Choung MG, Roh KH. 2008. *Arabidopsis* R2R3-MYB transcription factor AtMYB60 functions as a transcriptional repressor of anthocyanin biosynthesis in lettuce (*Lactuca sativa*). Plant Cell Reports 27:985-94.

Peer W, Brown D, Tague B, Muday G, Taiz L, Murphy A. 2001. Flavonoid accumulation patterns of *transparent testa* mutants of arabidopsis. Plant Physiology 126:536-548.

Pelletier MK, Shirley BW.1996. Analysis of flavanone 3-hydroxylase in arabidopsis seedlings coordinate regulation with chalcone synthase and chalcone isomerase. Plant Physiology 111: 339-345.

Petersen, M, Hans J, Matern U. 2010. Biosynthesis of phenylpropanoids and related compounds, in Annual Plant Reviews Volume 40: Biochemistry of plant secondary metabolism, second edition (ed M. Wink), Wiley-Blackwell, Oxford, UK.

Petit P, Granier T, Langlois d'Estaintot B, Manigand C, Bathany K, Schmitter J-M, Lauvergeat V, Hamdi S, Gallois B. 2007. Crystal structure of grape dihydroflavonol 4-reductase, a key enzyme in flavonoid biosynthesis. Journal of Molecular Biology 368:1345-1357.

Preuß A, Stracke R, Weisshaar B, Hillebrecht A, Matern U, Martens S. 2009. *Arabidopsis thaliana* expresses a second functional flavonol synthase. Federation of European Biochemical Societies Letters 583:1981-1986.

Quattrocchio F, Wing J, Woude K, Souer E, 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.

Ramsay NA, Glover BJ. 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends in Plant Science10:63-70.

Rasmussen S, Dixon RA. 1999. Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. The Plant Cell 11: 1537-1551.

Ro D, Douglas C. 2003. Reconstitution of the entry point of plant phenylpropanoid metabolism in yeast (*Saccharomyces cerevisiae*). The Journal of Biological Chemistry 279:2600-2607.

Rohde A, Morreel K, Ralph J, Goeminne G, Hostyn V, De Rycke R, Kushnir S, Van Doorsselaere J, Joseleau JP, Vuylsteke M, Van Driessche G, Van Beeumen J, Messens E, Boerjan W. 2004. Molecular phenotyping of the *pal1* and *pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. The Plant Cell 16: 2749-2771.

Saito K, Kobayashi M, Gong Z, Tanaka Y, Yamazaki M. 1999. Direct evidence for anthocyanidin synthase as a 2-oxoglutarate-dependent oxygenase: molecular cloning and functional expression of cDNA from a red forma of *Perilla frutescens*. The Plant Journal 17: 181-189.

Saito K, Yonekura-Sakakibara K, Nakabayashi R, Higashi Y, Yamazaki M, Tohge T, Fernie AR. 2013. The flavonoid biosynthetic pathway in arabidopsis: Structural and genetic diversity. Plant Physiology and Biochemistry 72: 21-34.

Samappito S, Page JE, Schmidt J, De-Eknamkul W, Kutchan TM. 2003. Aromatic and pyrone polyketides synthesized by a stilbene synthase from *Rheum tataricum.* Phytochemistry 62: 313–323.

Saslowsky D, Winkel-Shirley B. 2001. Localization of flavonoid enzymes in arabidopsis roots. The Plant Journal 27:37-48.

Sato F. 2014. Plant mecondary metabolism. In: eLS. John Wiley and Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0001812.pub2.

Sato T, Takabe K, Fujita, M. 2004. Immunolocalization of phenylalanine ammonia-lyase and cinnamate-4-hydroxylase in differentiating xylem of poplar. Comptes Rendus Biologies 327: 827-836.

Schijlen EGWM, Ric de Vos CH, Martens S, Jonker HH, Rosin FM, Molthoff JW, Tikunov YM, Angenent GC, Tunen AJ, and Bovy AG. 2007. RNA interference silencing of chalcone synthase, the first step in the flavonoid biosynthesis pathway, leads to parthenocarpic tomato fruits. Plant Physiology 144:1 520-1 530.

Schilmiller AL, Stout J, Weng J-K, Humphreys J, Ruegger MO, Chapple C. 2009. Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in arabidopsis. The Plant Journal 60:771-782.

Seitz C, Eder C, Deiml B, Kellner S, Martens S, Forkmann G. 2006. Cloning, functional identification and sequence analysis of *flavonoid 3'-hydroxylase* and *flavonoid 3',5'-hydroxylase* cDNAs reveals independent evolution of *flavonoid 3',5'-hydroxylase* in the Asteraceae family. Plant Molecular Biology 61:365-381.

Sewalt VJH, Ni W, Blount JW, Jung HC, Masoud SA, Howles PA, Lamb C, Dixon RA. 1997. Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of *L-phenylalanine ammonia-lyase* or *cinnamate 4 hydroxylase*. Plant Physiology 115: 41-50.

Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM. 1995. Analysis of arabidopsis mutants deficient in flavonoid biosynthesis. The Plant Journal 8:659-71.

Smith IE. 1999. The WD repeat: A common architecture for diverse functions. Trends in Biochemical Sciences 24:181-185.

Solano R, Nieto C, Avila J, Cafias L, Diaz I, Paz-Ares1 J. 1995. Dual DNA binding specificity of a petal epidermis specific MYB transcription factor

(MYB.PH3) from *Petunia hybrid*a. The European Molecular Biology Organization Journal. 14: 1773-1784.

Sommer H, Saedler H. 1986. Structure of the *chalcone synthase* gene of *Antirrhinum majus*. Molecular Genetics and Genomics 202:429-434.

Song H, Ryu H, Lee K, Jeong I, Kim D, Oh S. 2014. Metabolomics investigation of flavonoid synthesis in soybean leaves depending on the growth stage. Metabolomics 10:833-841.

Springob K, Nakajima J-I, Yamazaki M, Saito K. 2003. Recent advances in the biosynthesis and accumulation of anthocyanins. Natural Product Reports 20:288-303.

Srere PA. 1985. The metabolon. Trends in Biochemical Sciences 10:109.

Stafford HA 1974. Possible multi-enzyme complexes regulating the formation of C6- C3 phenolic compounds and lignins in higher plants. Recent Advances in Phytochemistry 8:53-79.

Stommel J, Lightbourn G, Winkel B, Griesbach R. 2009. Transcription factor families regulate the anthocyanin biosynthetic pathway in *Capsicum annuum*. Journal of the American Society for Horticultural Science 134: 2244-2251.

Tanaka Y, Katsumoto Y, Brugliera F, Mason J. 2005. Genetic engineering in floriculture. Plant Cell, Tissue and Organ Culture 80:1-24.

Taylor LP, Briggs W. 1990. Genetic regulation and photocontrol of anthocyanin accumulation in maize seedlings. The Plant Cell 2:115-127.

Teeri TH, Elomaa P, Kotilainen M, Albert VA. 2006. Mining plant diversity: gerbera as a model system for plant development and biosynthetic research. BioEssays 28:756–767.

Teutsch GH, Hasenfratz PM, Lesot A, Stoltz C, Garnier J-M, Jeltscht J-M, Durst F, Werck- Reichhart D.1993. Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. Proceedings of the National Academy of Sciences 90:4102-4106.

Tichopad A, Kitchen R, Riedmaier I, Christiane B, Ståhlberg A, Kubista M. 2009. Design and optimization of reverse-transcription quantitative PCR experiments. Clinical Chemistry 55:1816-1823.

Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J-I, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki1M, Saito K. 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.

Tohge T, Watanabe M, Hoefgen R, Fernie AR. 2013. The evolution of phenylpropanoid metabolism in the green lineage. Critical Reviews in Biochemistry and Molecular Biology 48: 123-152.

Turnbull JJ, Sobey WJ, AplinRT, Hassan A, Firmin JL, Schofield CJ, Prescott AG. 2000. Are anthocyanidins the immediate products of anthocyanidin synthase? Chemical Communications 2473-2474.

Tuteja JH, Clough SJ, Chan W-C, Vodkin LO. 2004. Tissue-specific gene silencing mediated by a naturally occurring *chalcone synthase* gene cluster in *Glycine max*. The Plant Cell 16: 819-835.

Tuteja JH, Vodkin LO. 2008. Structural features of the endogenous *CHS* silencing and target loci in the soybean genome. Crop Science 48: S49–S68.

Tyrach A, Horn W.1997. Inheritance of flower color and flavonoid pigments in gerbera. Plant Breeding 116: 377–381.

Tzin V, Malitsky S, Moyal Ben Zvi M, Bedair M, Sumner L, Aharoni A, Galili G. 2012. Expression of a bacterial feedback-insensitive *3-deoxy-D-arabinoheptulosonate 7-phosphate synthase* of the shikimate pathway in arabidopsis elucidates potential metabolic bottlenecks between primary and secondary metabolism. New Phytologist 194: 430-439.

Van Dongen, S. Graph Clustering by Flow Simulation. PhD thesis, University of Utrecht, May 2000. (http://www.library.uu.nl/digiarchief/dip/diss/1895620/inhoud.htm).

van Tunen A, Mur L A., Recourt K, Gerats AGM, Mol JNM.1991. Regulation and manipulation of flavonoid gene expression in anthers of petunia: The molecular basis of the *Po* mutation. The Plant Cell 3:39-48.

Verpoorte R, Memelink J. 2002. Engineering secondary metabolite production in plants. Current Opinion in Biotechnology 13:181-187.

Vogt T, Jones P. 2000. Glycosyltransferases in plant natural product synthesis: Characterization of a supergene family. Trends in Plant Science 5:1360-1385.

Vrhovsek U, Masuero D, Gasperotti M, Franceschi P, Caputi L, Viola R. 2012. A versatile targeted metabolomics method for the rapid quantification of multiple classes of phenolics in fruits and beverages. Journal of Agricultural and Food Chemistry 60:8831-8840.

Wagner GJ, Hrazdina G. 1984. Endoplasmic reticulum as a site of phenylpropanoid and flavonoid metabolism in *Hippeastrum*. Plant Physiology 74: 901-906.

Wanner LA, Li G, Ware D, Somssich IE, Davis KR.1995. The *phenylalanine ammonia-lyase* gene family in *Arabidopsis thaliana*. Plant Molecular Biology 27:327-338.

Wei Q, Wang Q-Y, Feng Z-H, Wang B, Zhang Y-F, Yang Q. 2012. Increased accumulation of anthocyanins in transgenic potato tubers by overexpressing the *3GT* gene. Plant Biotechnology Reports 6:69-75.

Weisshaar B, Jenkins GI. 1998. Phenylpropanoid biosynthesis and its regulation. Current Opinion in Plant Biology 1:251-257.

Wellmann F, Griesser M, Schwa W, Martens S, Eisenreich W, Matern U, Lukačina R. 2006. Anthocyanidin synthase from *Gerbera hybrida* catalyzes the conversion of (+)-catechin to cyanidin and a novel procyanidin. Federation of European Biochemical Societies Letters 580:1642-1648.

Wengenmayer H, Ebel J, Grisebach H.1976. Enzymic synthesis of lignin precursors. Purification and properties of a cinnamoyl-CoA: NADPH reductase from cell suspension cultures of soybean (*Glycine max*). European Journal of Biochemistry 65:529-536.

Wink M. 2010. Introduction: Biochemistry, physiology and ecological functions of secondary metabolites, in Annual Plant Reviews Volume 40: Biochemistry of plant secondary metabolism, second edition (ed M. Wink), Wiley-Blackwell, Oxford, UK.

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

Winkel-Shirley B. 2002. Biosynthesis of flavonoids and effects of stress. Current Opinion in Plant Biology 5:218-223.

Winkel-Shirley B. 2004. Metabolic channelling in plants. Annual Review of Plant Biology 55:85-107.

Yao LH, Jiang YM, Shi J, Tomas-Barbera FA, Datta N, Singanusong R, Chen SS. 2004. Flavonoids in food and their health benefits. Plant Foods for Human Nutrition 59:113-122.

Yi J, Derynck MR, Li X, Telmer P, Marsolais F, Dhaubhade S. 2010. A single-repeat *MYB* transcription factor, *GmMYB*176, regulates *CHS*8 gene expression and affects isoflavonoid biosynthesis in soybean. The Plant Journal 62:1019-1034.

Zhou L, Wang Y, Peng Z. 2011. Molecular characterization and expression analysis of *chalcone synthase* gene during flower development in tree peony (*Paeonia suffruticosa*). African Journal of Biotechnology 10:1275-1284.

Recent Publications in this Series