Dissertationes bioscientiarum molecularium
Universitatis Helsingiensis in Viikki
32 / 2008

Secondary metabolites in *Gerbera hybrida*

Miia Ainasoja

Department of Applied Biology
Faculty of Agriculture and Forestry

Department of Biological and Environmental Sciences
Division of Genetics
Faculty of Biosciences

Helsinki Graduate School in Biotechnology and Molecular Biology

University of Helsinki
Finland

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Biosciences of the University of Helsinki, for public examination, in Viikki, Auditorium B3 (Forest Sciences building, Latokartanonkaari 9), on September 26th 2008, at 12 o`clock noon.

Supervised by
Professor Teemu Teeri
Department of Applied Biology
University of Helsinki

Reviewed by
Professor Sirpa Kärenlampi
Department of Biosciences
University of Kuopio

Docent Günter Brader
Department of Biological and Environmental Sciences
University of Helsinki

Opponent
Doctor of Philosophy
Markku Keinänen
Faculty of Biosciences
University of Joensuu

Custos
Professor Tapio Palva
Department of Biological and Environmental Sciences
University of Helsinki

ISSN 1795-7079
ISBN 978-952-10-4907-1 (paperback)
ISBN 978-952-10-4908-8 (pdf)
http://ethesis.helsinki.fi

Helsinki University Printing House
Helsinki 2008

Dedicated to the memory of my granddad Tauno Kemppainen

CONTENT

ABBREVIATIONS

ACS	acridone synthase
ANR	anthocyanidin reductase
ANS	anthocyanidin synthase
BAS	benzalacetone synthase
BBS	bibenzyl synthase
bHLH	basic helix-loop-helix
BIS	biphenyl synthase
BPS	benzophenone synthase
cDNA	complementary DNA
C4H	cinnamate-4-hydroxylase
CHI	chalcone isomerase
CHR	chalcone reductase
CHS	chalcone synthase
CoA	coenzyme A
CTAS	coumaroyl triacetic acid synthase
DFR	dihydroflavonol 4-reductase
F3H	flavanone 3-hydroxylase
F3'H	flavonoid 3' hydroxylase
FLS	flavonol synthase
FNS	flavone synthase
GM	genetic modification
3-GT	UDP-glucose:flavonoid 3-*O*-glucosyltransferase
5-GT	UDP-glucose:anthocyanin 5-*O*-glucosyltransferase
HEDS	homoeriodictyol/eriodictyol synthase
HCA	hydroxycinnamate
HPLC	high-performance liquid chromatography
IFS	isoflavone synthase
LAR	leucoanthocyanidin reductase
OKS	octaketide synthase
ORF	open-reading frame
PAL	phenylalanine ammonia lyase
PCA	principal component analysis
PCS	pentaketide chromone synthase
PKS	polyketide synthase
2PS	2-pyrone synthase
RT-PCR	reverse transcription polymerase chain reaction
SM	secondary metabolite
STS	stilbene synthase
TAL	triacetic acid lactone
TLC	thin-layer chromatography
VPS	phlorisovalerophenone synthase

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I **Ainasoja M**, Laitinen RAE and Teeri TH. Type III polyketide synthase family in *Gerbera hybrida*. Manuscript

II Laitinen RAE, **Ainasoja M**, Broholm SK, Teeri TH and Elomaa P (2008) Identification of target genes for a MYB-type anthocyanin regulator in *Gerbera hybrida*. Journal of Experimental Botany DOI 10.1093/jxb/ern216

III **Ainasoja M**, Pohjala L, Tammela P, Somervuo P, Vuorela P and Teeri TH (2008) Comparison of transgenic *Gerbera hybrida* lines and traditional varieties shows no differences in cytotoxicity or metabolic fingerprints. Transgenic Research DOI 10.1007/s11248-008-9165-8

IV Fallarero A, **Ainasoja M**, Sandberg M, Teeri T and Vuorela P. GT1-7 cell-based cytoxicity screening assay on 96- microtiter well plates as a platform for the safety assessment of genetically modified *Gerbera hybrida* extracts. Drug and Chemical Toxicology, in press

The publications are referred to in the text by their roman numerals.

ABSTRACT

Plants produce a diversity of secondary metabolites, i.e., low-molecular-weight compounds that have primarily ecological functions in plants. The flavonoid pathway is one of the most studied biosynthetic pathways in plants. In order to understand biosynthetic pathways fully, it is necessary to isolate and purify the enzymes of the pathways to study individual steps and to study the regulatory genes of the pathways. Chalcone synthases are key enzymes in the formation of several groups of flavonoids, including anthocyanins. In this study, a new chalcone synthase enzyme (GCHS4), which may be one of the main contributors to flower colour, was characterised from the ornamental plant *Gerbera hybrida*. In addition, four chalcone synthase-like genes and enzymes (*GCHS17, GCHS17b*, *GCHS26* and *GCHS26b*) were studied. Spatial expression of the polyketide synthase gene family in gerbera was also analysed with quantitative RT-PCR from 12 tissues, including several developmental stages and flower types. A previously identified MYB transcription factor from gerbera, *GMYB10,* which regulates the anthocyanin pathway, was transferred to gerbera and the phenotypes were analysed. Total anthocyanin content and anthocyanidin profiles of control and transgenic samples were compared spectrophotometrically and with HPLC. The overexpression of *GMYB10* alone was able to change anthocyanin pigmentation: cyanidin pigmentation was induced and pelargonidin pigmentation was increased. The gerbera 9K cDNA microarray was used to compare the gene expression profiles of transgenic tissues against the corresponding control tissues to reveal putative target genes for GMYB10. *GMYB10* overexpression affected the expression of both early and late biosynthetic genes in anthocyanin-accumulating transgenic tissues, including the newly isolated gene *GCHS4*. Two new MYB domain factors, named as *GMYB11* and *GMYB12*, were also upregulated.

Gene transfer is not only a powerful tool for basic research, but also for plant breeding. However, crop improvement by genetic modification (GM) remains controversial, at least in Europe. Many of the concerns relating to both human health and to ecological impacts relate to changes in the secondary metabolites of GM crops. In the second part of this study, qualitative and quantitative differences in cytotoxicity and metabolic fingerprints between 225 genetically modified *Gerbera hybrida* lines and 42 non-GM *Gerbera* varieties were compared. There was no evidence for any major qualitative and quantitative changes between the GM lines and non-GM varieties. The developed cell viability assays offer also a model scheme for cell-based cytotoxicity screening of a large variety of GM plants in standardized conditions.

1 INTRODUCTION

1.1 Secondary metabolites and their function

> "Secondary metabolites are compounds with a restricted occurrence in taxonomic groups that are not necessary for a cell (organism) to live, but play a role in the interaction of the cell (organism) with its environment, ensuring the survival of the organism in its ecosystem."

This definition of Dr. Robert Verpoorte (Verpoorte and Alfermann 2000) reflects the present understanding of the role of secondary metabolites (SM). SMs are low-molecular-weight compounds the formation of which is generally organ-, tissue-, and cell-specific (Wink 1999). In addition, the amount and types of compounds often differ between individuals from the same population of plants. SMs have many functions. They protect plants against stresses, both biotic (attack by fungi, bacteria, nematodes, insects or grazing by mammals) and abiotic (extremes of temperature and moisture, shading, injury or presence of heavy metals; Seigler 1998). In these cases, formation of SM can be influenced in both general and specific manner. SMs can act as allelochemicals that a plant uses to keep other plants from growing in too close proximity, or as kairomones that are used by the interacting organism to locate the host plant. They are also involved in processes such as pollination and fruit/seed dispersal as well as in flower pigmentation. From the human point of view, quite a few SMs are used as specialty chemicals such as drugs, flavours, fragrances, insecticides, and dyes, and thus have a great economic value (Verpoorte and Alfermann 2000).

In plants, SMs can be divided into three groups: terpenoids, polyketides, and phenylpropanoids based on their biosynthetic origin (Verpoorte and Alfermann 2000). In addition to these groups, an important class of SMs is alkaloids, which are nitrogenous organic molecules biosynthesized principally from amino acids, e.g., tryptophan, tyrosine, phenylalanine, lysine, and arginine, using many unique enzymes (Croteau et al. 2000). Many of the most important therapeutic agents are alkaloids. After the basic skeletons of SMs have been formed, further modifications result in plant species-specific compounds. The sites of biosynthesis are compartmentalised at cellular or subcellular level. However, SMs can be transported long distances and accumulate far from their location of synthesis.

As opposed to secondary metabolites, primary metabolites are found in all plants and perform essential metabolic roles, e.g., by participating in nutrition and reproduction (Croteau et al. 2000). Sometimes it is hard to make difference between primary and secondary metabolites. For example, both primary and secondary metabolites are found among the terpenoids and the same compound may have both primary and secondary roles.

1.2 Major biosynthetic pathways of secondary metabolites

1.2.1 Terpenoids

The largest family of plant SMs is the terpenoids; over 30000 individual compounds of this class have been described from plants (Broun and Somerville 2001). Terpenoids are extremely variable in structure possessing hundreds of different carbon skeletons and a large assortment of functional groups. The classification is based on the number of isoprenoid units (C_5) present in their structure (Wink 1999). Although biosynthesis is based on a unit of five carbon atoms, terpenoid nomenclature is based on a unit of ten carbon atoms. C_5 compounds are called hemiterpenes, C_{10} compounds monoterpenes, C_{15} compounds sesquiterpenes, and so on.

Terpenoids, both primary and secondary metabolites, play diverse functional roles in plants as hormones (e.g. gibberellins, abscisic acid), photosynthetic pigments (e.g. phytol, carotenoids), electron carriers (e.g. ubiquinone, plastoquinone), mediators of polysaccharide assembly (e.g. polyprenyl phosphates), and structural components of membranes (phytosterols; McGarvey and Croteau 1995). In addition, many specific terpenoid compounds serve in communication and defence of plants. Terpenoids can be found in large amounts as essential oils, resins and waxes. Members of the terpenoid group include industrially useful polymers (e.g. natural rubber) and a number of pharmaceuticals and agrochemicals. Terpene-based natural products that have been used in treating human diseases include, e.g., the antimalarial drug artemisinin and the anticancer drug taxol (Martin et al. 2003).

The biosynthetic pathway to terpenoids comprises four stages (McGarvey and Croteau 1995). The first step involves the synthesis of the C_5 isoprene unit isopentenyl diphosphate (IPP). The classical route for IPP is via the reactions of the acetate-mevalonate pathway in the cytoplasm and mitochondria. Lichtenthaler et al. (1997) discovered that the basic C_5 units can also be constructed via the plastidic glyceraldehyde phosphate-pyruvate pathway. In fact, most of the plant terpenes are synthesised in this latter pathway. In the second stage, the C_5 units condense to generate three larger prenyl diphosphates: geranyl diphosphate (GPP, C_{10}), farnesyl diphosphate (FPP, C_{15}) and geranylgeranyl diphosphate (GGPP, C_{20}). In the third stage, the C_{10}-C_{20} diphosphates undergo a wide range of cyclisations and rearrangements to produce the parent carbon skeletons of each terpene class. Subsequent transformations of the basic parent skeletons involve oxidations, reductions, isomerisations, conjugations and other transformations to form thousands of distinct terpene metabolites.

Terpenoid biosynthesis is regulated both spatially and temporally (McGarvey and Croteau 1995). Formation of volatile compounds is spatially regulated (Cheng et al. 2007). Several studies have shown that terpenoid volatiles are often synthesised *de novo* in some special physical structures. Specialised secretory structures are usually required when large amounts of hydrophobic mono-, sesqui- and diterpenes are produced and accumulated (Gershenzon et al. 1989; Lewinsohn et al. 1991). Biosynthesis of some terpenoids is restricted to specific tissues where they are utilised (Bartley and Scolnik

1995; Pichersky et al. 1994). There are at least three distinct semi-autonomous subcellular compartments where terpenoids are synthesised: the cytosol/ER for sesquiterpenoids and triterpenoids (Goodwin 1979), plastids for mono-, di- and tetraterpenoids (Kleinig 1989) and mitochondria (and/or Golgi apparatus) for ubiquinone biosynthesis (Lütke-Brinkhaus et al. 1984). The rhythmic and developmental regulation of volatile terpenoid emission from different species has been reported (Pichersky et al. 1994; Cheng et al. 2007). Many terpenoids are also specifically induced in defence response to elicitors (Chappell and Nable 1987) or wounding (Lewinsohn et al. 1991). In indirect defense, a blend of terpenoid volatiles has been shown to attract the carnivores of herbivores, after the herbivore attacking (Mumm et al. 2008).

1.2.2 Polyketides

Polyketide metabolites are found in bacteria, fungi, and plants (Hopwood and Sherman 1990). They have many biological activities, which have been utilised in medicinal and agricultural sectors as antimicrobials, immunosuppressants, antiparasitics, and anticancer agents (Seigler 1998). Their structures are strikingly diverse, yet they share a common pattern of biosynthesis. They are synthesized by polyketide synthase enzymes (PKSs) using simple precursors such as acetyl-CoA, 4-coumaroyl-CoA and malonyl-CoA in a manner similar to fatty acid biosynthesis. The chain length of the polyketide intermediates increases by units of two carbon atoms to various lengths. Cyclisation of intermediates occurs by an intramolecular aldol or Claisen condensation. Several types of reactions follow the initial cyclisations: addition and removal of hydroxyl and methoxyl groups, oxidation of methyl groups to alcohols, aldehydes, and acids, the reverse of this series, and decarboxylation. PKSs can be divided into three groups (Austin and Noel 2003). PKS types I and II are found in fungi and in some bacteria, and resemble fatty acid synthase enzymes. Type III PKSs are nearly completely plant-specific, only few are found in bacteria. They have similarity to chalcone synthases (CHSs), indicative of a common origin in evolution. Type I PKSs consist of multi-domain polyproteins that form large multi-functional biosynthetic complexes. Type II PKS activities are catalysed by putative multienzyme complexes consisting of discrete, separable proteins. Type III PKSs are homodimers catalyzing various iterative condensation reactions utilizing a diverse set of starter molecules.

1.2.3 Phenylpropanoids

Phenylpropanoids, i.e. phenolic compounds, with a C_6-C_3 skeleton are products of the general phenylpropanoid (phenylalanine/hydroxycinnamate) pathway, which includes as first steps reactions from *L*-phenylalanine to the hydroxycinnamic acids (HCAs; Wink 1999). The first enzyme of the pathway is phenylalanine ammonia lyase (PAL; Koukol and Conn 1961) that catalyses deamination of *L*-phenylalanine to (*E*)-cinnamate, which is hydroxylated to 4-coumarate by cinnamate 4-hydroxylase (Gabriac et al. 1991). The 4-

coumarate:CoA ligases catalyze the activation of HCAs (4-coumarate, caffeate, ferulate, and sinapate) to the respective CoA esters (4-coumaroyl-CoA, caffeoyl-CoA, feruloyl-CoA, and sinapoyl-CoA; e.g. Hamberger and Hahlbrock 2004). The synthesis of caffeate, ferulate and sinapate take place through hydroxylations and methylations at different levels of the pathway.

The biosynthetic pathways of natural compounds derived from phenylpropanoids are among the most thoroughly investigated biosynthetic pathways in plants (Seigler 1998). They include, e.g., pathways leading to coumarins, lignin, lignans, hydrolyzable tannins (gallotannins and ellagitannins) and flavonoids, which are discussed in more detail.

1.3 Flavonoids

To date, over 6000 naturally occurring flavonoids have been isolated and identified; many are common in higher plants (Austin and Noel 2003). All flavonoids share a basic C_6-C_3-C_6 structural skeleton, consisting of two aromatic C_6 rings (A and B) and a heterocyclic ring (C) that contains one oxygen atom (Figure 1A). According to the oxidation level of the central heterocyclic ring C, flavonoids are grouped into the following structural classes: flavanones, dihydroflavonols, leucoanthocyanidins (flavan-3,4-diols), flavan-3-ols, anthocyanidins, anthocyanins, flavones, flavonols and isoflavones (Figure 2). All these classes of flavonoids have a basic structural skeleton that is further modified to produce hundreds of different compounds of each class. Chalcones, dihydrochalcones, stilbenes and aurones lack the typical flavonoid structure, but are biosynthetically closely related to flavonoids. Most flavonoid compounds are glycosides, which are often accumulated in the vacuoles of plant cells. Glycosides can be either *O*- or *C*-linked. The variety of flavonoid glycosides is based on the number of positions on the flavonoid for glycosylation, the level of glycosylation and the number of different sugars that can be involved. In a single species, dozens of different flavonoids may be present. Of the several hundred aglycones isolated from plants, only eight are distributed widely (Seigler 1998). The eight most common flavonoid nuclei are the flavonols kaempferol, quercetin and myricetin, the anthocyanidins pelargonidin, cyanidin and delphinidin, and the flavones apigenin and luteolin. Most genes for the biosynthetic enzymes of flavonoid pathway have been isolated, and the biochemical reactions catalyzed by these enzymes have been characterized. Many of the enzymes, although not all, are encoded by small gene families.

1.3.1 Functions of flavonoids in nature and human uses

Flavonoids have many key functions in plant growth and development; many of these tasks are critical for survival (Bohm 1998; Gould and Lister 2006). In some species, they are required for the attraction of animal vectors for pollination and seed dispersal (e.g. anthocyanins) as well as germination of pollen grains and successful pollen tube growth (e.g. flavonols kaempferol and quercetin 3´- methyl ether). Flavonoids induce the activation of bacterial nodulation genes (nitrogen fixation; e.g. flavone luteolin and

isoflavones daidzein and genistein) and the oviposition response of some insects. They are involved in resorption of mineral nutrients from senescing leaves. Some flavonoids help plants that grow under suboptimal environments, e.g., they enhance tolerance to a variety of abiotic stressors, play a role in defense against herbivores and pathogens and form the basis for allelopathic interactions with other plants. Flavonoids are also a regular part of our food and have applications in food industry as sweeteners or food colours. As flower pigments and co-pigments, they contribute to the garden and cut flower industry. Co-pigmentation is the enhancement of flower colour due to the association of anthocyanins with colourless co-pigments, e.g., the flavone apigenin. Many flavonoids have potential beneficial effects on human health. They have been found to have, e.g., antioxidant, antiviral, antibacterial, or antifungal properties and have been tested as potential medicinal agents against human diseases including malaria and HIV. Field legumes, on the other hand, produce several isoflavonoid derivatives that may have serious negative effects upon grazing livestock owing to their capacity to simulate the action of animal hormones.

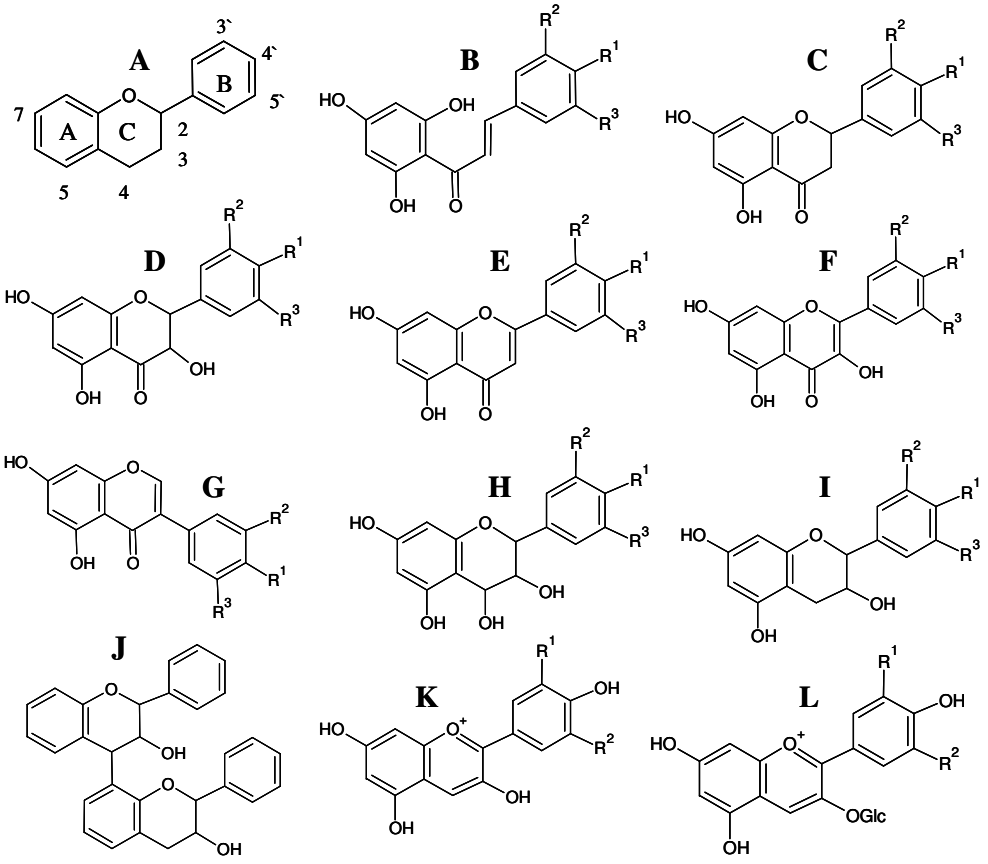

Figure 1 *Structures of flavonoids. Basic structure of flavonoid (A), chalcone (B), flavanone (C), dihydroflavonol (D), flavone (E), flavonol (F), isoflavone (G), leucoanthocyanidin (H), flavan (I), proanthocyanidin (J), anthocyanidin (K) and anthocyanin (L).*

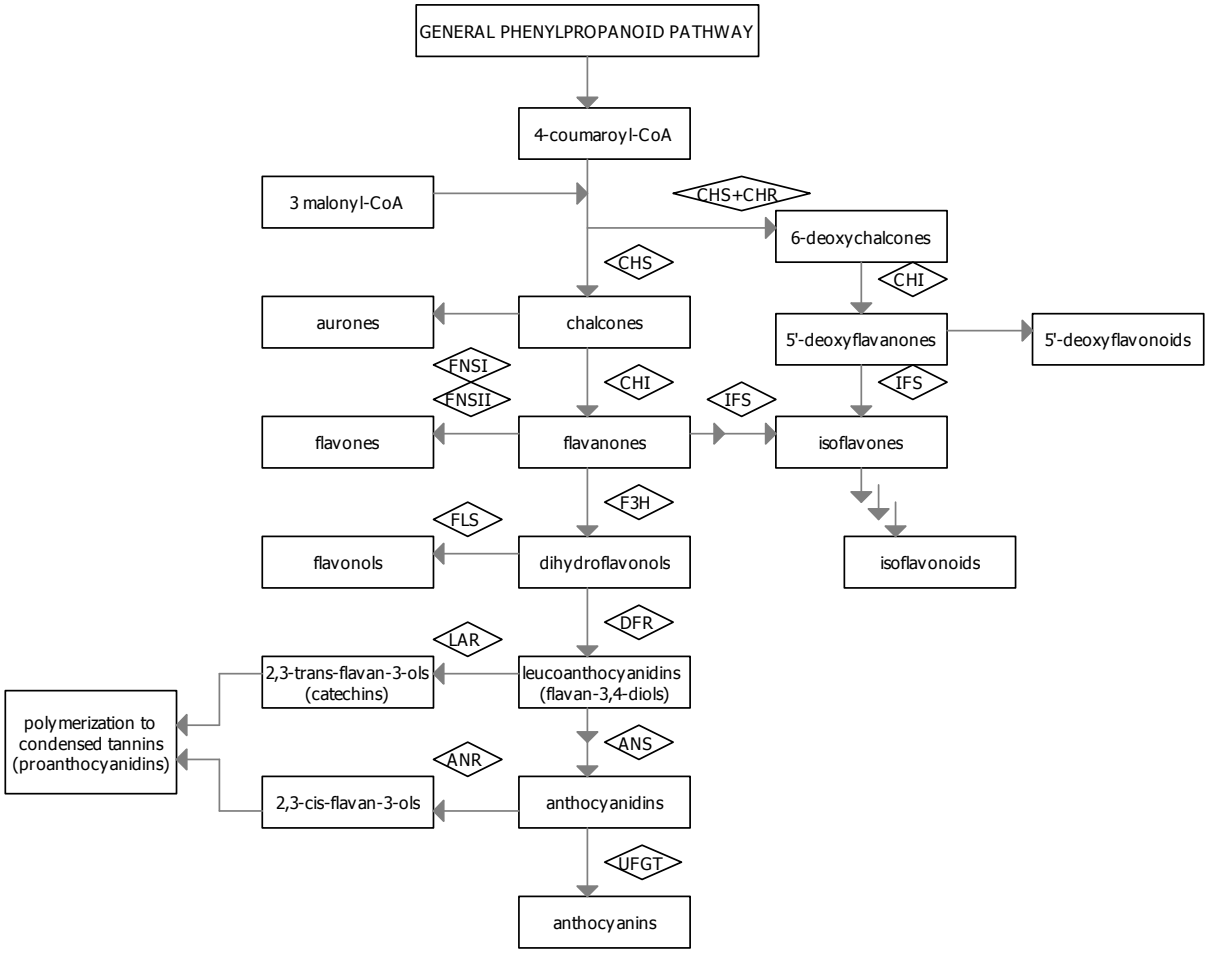

Figure 2 *Major branch pathways of flavonoid biosynthesis*

1.3.2 Flavonoid biosynthesis

1.3.2.1 Chalcones and chalcone synthases

Chalcones are key intermediates in the formation of several major groups of flavonoids. They lack the typical flavonoid structure, but are biosynthetically closely related to flavonoids. The chalcones have two aromatic rings, A and B, which are linked by a C_3-bridge (Figure 1B). This bridge gives rise to the heterocyclic ring C in all other flavonoid classes.

The synthesis of chalcones is catalyzed by chalcone synthases (CHSs; Figure 2). This activity was first demonstrated from cell cultures of *Petroselinum crispum* (parsley; Kreuzaler and Hahlbrock 1972). CHS is the key enzyme of the flavonoid pathway and requires no cofactors. CHS condenses 4-coumaroyl-CoA with three malonyl-CoA derived acetate units to form a tetraketide, followed by a Claisen-type cyclization reaction leading

to formation of naringenin chalcone (4,2',4',6'-tetrahydroxychalcone; Heller and Hahlbrock 1980). The molecular weight of the CHS dimer ranges from 75 to 88 kDa. The first *CHS* gene was isolated by Reimold et al. (1983). Ferrer et al. (1999) published the first three-dimensional structure of CHS2 from the legume *Medicago sativa* (alfalfa). They found three conserved residues (Cys164, His303, and Asn336) likely to participate in the multiple decarboxylation and condensation reactions necessary during the course of the CHS-catalysed mechanism. In addition, 15 inert active site residues were proposed to provide steric guidance to the linear polyketide. CHS has been found from all gymnosperms and angiosperms examined thus far and most plants contain multiple copies of *CHS* genes, which are expressed in various tissues at different developmental stages (Austin and Noel 2003). CHSs belong to plant type III polyketide synthase enzymes.

Plant type III PKSs produce a number of biologically important compounds. They are involved in the biosynthesis of flavonoids as well as a variety of other plant secondary metabolites (Austin and Noel 2003; Schröder 1997). Because plant PKSs were discovered soon after CHSs, known type III plant PKSs are often referred to as the chalcone synthase/stilbene synthase superfamily. Stilbene synthases (STSs) are CHS-like enzymes that synthesize the same linear tetraketide intermediate from 4-coumaroyl-CoA and malonyl-CoA, but release resveratrol as an end product of aldol condensation. A third CHS-like enzyme, acridone synthase (ACS), catalyzes three condensations of malonyl-CoA to an *N*-methylanthraniloyl-CoA starter, followed by a CHS-like intramolecular Claisen cyclisation of the tetraketide intermediate (Baumert et al. 1994). Approximately 100 naturally occurring acridone alkaloids are presently known.

Recently, several new enzymes related to CHSs have been characterised, often via cloning of corresponding cDNA molecules using methods based on DNA hybridisation, and their functional expression in *E. coli*. These enzymes differ from CHSs in their preference for substrates, the number of acetyl unit additions they catalyse, and their mechanism of chain termination. 2-pyrone synthase (2PS) is a triketide synthase found from the ornamental plant gerbera performing two condensation reactions with acetyl-CoA and malonyl-CoA producing triacetic acid lactone (TAL, 6-methyl-4-hydroxy-2-pyrone; Eckermann et al. 1998; Helariutta et al. 1995). Jez et al. (2000b) determined the structure of gerbera 2PS complexed with the reaction intermediate acetoacetyl-CoA. 2PS and CHS shared a common three-dimensional fold, a set of conserved catalytic residues, and similar CoA binding sites, but the active site cavity in 2PS was shown to be smaller, in concordance with preference of a smaller substrate in 2PS. Benzalacetone synthase (BAS) is a diketide synthase that catalyses a single step decarboxylative condensation of 4-coumaroyl-CoA with malonyl-CoA to produce *p*-hydroxyphenylbut-3-ene-2-one in *Rubus idaeus* (raspberry; Borejsza-Wysocki and Hrazdina 1996) and the C_6-C_4 skeleton of phenylbutanoids in *Rheum palmatum* (rhubarb; Abe et al. 2001). Liu et al. (2007) found a biphenyl synthase (BIS) from yeast-extract-treated cell cultures of *Sorbus aucuparia* (mountain ash). Its preferred substrate is benzoyl-CoA that undergoes iterative condensation with three molecules of malonyl-CoA to give 3,5-dihydroxybiphenyl via intramolecular aldol condensation. Abe et al. (2004) cloned a heptaketide synthase gene from rhubarb. A pentaketide chromone synthase (PCS) and an octaketide synthase (OKS) from *Aloe arborescens* (aloe) were also cloned and characterised (Abe et al. 2005a and

2005b). Other type III PKSs are bibenzyl synthases (BBSs), homoeriodictyol/eriodictyol synthase (HEDS), benzophenone synthase (BPS), phlorisovalerophenone synthase (VPS), and coumaroyl triacetic acid synthase (CTAS; Austin and Noel 2003). They all show high amino acid sequence similarity and demonstrate that the flexibility of PKS III enzymes has been utilized repeatedly during evolution of plants.

In addition to CHS, leguminous plants possess chalcone reductase (CHR). The coupled catalytic action of these two enzymes results in the formation of 6'-deoxychalcone (4,2',4'-trihydroxychalcone). 6'-deoxychalcones are precursors for phytoalexins (e.g., isoflavonoids; Austin and Noel 2003) and are also involved in the induction of symbiotic root nodulation by *Rhizobium* bacteria leading to nitrogen fixation (Graham 1991). CHR activity was first demonstrated in crude extracts of *Glycyrrhiza echinata* (licorice; Ayabe et al. 1988). The first *CHR* cDNA was isolated from *Glycine max* (soybean; Welle and Grisebach 1989) and the three-dimensional structure of alfalfa chalcone reductase, bound to the $NADP^+$ cofactor, was elucidated by Bomati et al. (2005).

1.3.2.2 Flavanones

The first central compounds in the flavonoid pathway are flavanones, which occur in most plants (Seigler 1998). They are commonly present in the Asteraceae and Fabaceae, but they have been reported in at least 60 other families. The synthesis of flavanones is catalyzed by chalcone isomerase (CHI) which catalyses the stereospecific isomerization of chalcones into their corresponding (2*S*)-flavanones (Figure 1C; Moustafa and Wong 1967). CHIs are classified into two types, and their distribution is highly plant family-specific. One class of CHIs accepts only 4,2',4',6'-tetrahydroxychalcones (naringenin chalcone; abbreviated as 6'-hydroxychalcone) and the other one 4,2',4'-trihydroxychalcones (6'-deoxychalcone) forming (*2S*)-naringenin (5,7,4'-trihydroxyflavanone), the key precursor of anthocyanins, and (*2S*)-5-deoxyflavanone (7,4'-dihydroxyflavanone), respectively (Wink 1999). The cyclization reaction can also proceed spontaneously under physiological conditions yielding a racemic mixture of (2*RS*)-flavanones, but only (2*S*)-flavanones are intermediates of the subsequent flavonoid metabolism. The structure of CHI from alfalfa was determined by Jez et al. (2000a).

1.3.2.3 Dihydroflavonols

Hydroxylation in position 3 of flavanones leads to dihydroflavonols (Figure 1D). This reaction was first detected in *Matthiola incana* (Forkmann et al. 1980) and is catalysed by flavanone 3-hydroxylase (F3H) that was classified as a 2-oxoglutarate-dependent dioxygenase. The enzyme requires 2-oxoglutarate, Fe^{2+}, ascorbate and molecular oxygen. The most common dihydroflavonols are dihydrokaempferol, dihydroquercetin, and dihydromyricetin.

1.3.2.4 Flavones and flavonols

Formation of a double bond between positions 2 and 3 of flavanones and dihydroflavonols leads to flavones and flavonols, respectively (Figure 1E and 1F). The flavones apigenin and luteolin as well as flavonols kaempferol, quercetin and myricetin are among the most common flavonoid nuclei. Two flavone synthase enzymes (FNSI and FNSII) have been described. FNSI is apparently quite restricted in its distribution while FNSII is much more widespread. FNSI activity was found from cell cultures of parsley (Britsch, et al. 1981). It was demonstrated that the enzyme requires oxygen, 2-oxoglutarate, Fe^{2+}, and ascorbate as cofactors. FNSII activity was first reported by Stotz and Forkmann (1981) using flowers of *Antirrhinum majus* (snapdragon). The enzyme has an absolute requirement for NADPH and molecular oxygen. Martens and Forkmann (1999) reported the first isolation and expression of a functional FNSII cDNA clone from *Gerbera hybrida* (gerbera). The earliest demonstrations of flavonol synthase (FLS) activity came from the study of Britsch et al. (1981). FLS has the same cofactor requirements as FNSI.

1.3.2.5 Isoflavones and isoflavonoids

Isoflavones, for example genistein and daidzein, are distinct from other flavonoid classes by having the B-ring attached to position 3 of the heterocyclic ring C, instead of position 2 (Figure 1G). They are produced almost exclusively in legumes, many of which are commercially important such as soybean, *Phaseolus vulgaris* (green beans), *Pisum sativum* (peas), and alfalfa. Isoflavone synthase (IFS) is the first enzyme in the synthesis of the isoflavones from flavanones (Figure 2). It is encoded in soybean by two genes, IFS-1 and IFS-2, which have been cloned and examined in some detail by several groups (Akashi et al. 1999; Jung et al. 2000; Steele et al. 1999). All isoflavonoids are derived from isoflavones, and the formation of isoflavone skeletons is the critical step in isoflavonoid biosynthesis. Isoflavones are involved, for example, in plant defence (Graham and Graham 1996) and root nodulation (Kosslak et al. 1987). They are also studied for their substantial health benefits for humans (Jung et al. 2000). The low incidences of osteoporosis, breast cancer and menopausal symptoms have been connected to women who eat a lot of isoflavones in their diet (Hooper 2006). Isoflavonoids play ecophysiological roles as defensive agents and as symbiotic signals to rhizobial bacteria and show estrogenic activity in animals (Akashi et al. 2006). Isoflavonoids possessing pterocarpan and isoflavan skeletons are the most commonly found phytoalexins of leguminous plants.

1.3.2.6 Leucoanthocyanidins

Reduction of the carbonyl group in position 4 of dihydroflavonols leads to flavan-3,4-diols i.e. leucoanthocyanidins, which are very unstable molecules (Figure 1H). Dihydroflavonol 4-reductase (DFR) is an enzyme catalysing the reduction of three dihydroflavonols

(dihydrokaempferol, dihydroquercetin, and dihydromyricetin) into leucoanthocyanidins, which are common precursors for anthocyanin and condensed tannin biosynthesis. DFR is encoded by as a single gene or a small gene family in several plant species, e.g., *Pseudotsuga menziesii* (Douglas fir; Stafford and Lester 1982), *Hordeum vulgare* (barley; Kristiansen and Rohde 1991), gerbera (Helariutta et al. 1993), and *Lotus japonicus* (Shimada et al. 2005).

1.3.2.7 Flavans and proanthocyanidins

Proanthocyanidins, i.e., condensed tannins are widely distributed and are particularly common in conifers. They can be formed from leucoanthocyanidins via two alternative routes (Figure 2). NADPH-dependent leucoanthocyanidin reductase (LAR) converts leucoanthocyanidins to 2,3-*trans*-flavan-3-ols such as catechin (Stafford and Lester 1984; Tanner et al. 2003). Xie et al. (2003) found an anthocyanidin reductase (ANR), a novel enzyme catalyzing the NADPH-dependent reduction of anthocyanidins to 2,3-*cis*-flavan-3-ols. The factors controlling the condensation of *trans*- and *cis*-flavan-3-ol monomers to form condensed tannin polymers are still unknown. However, it has been suggested that *TRANSPARENT TESTA10* (*TT10*) gene of *Arabidopsis* encodes a protein that exhibits strong similarity with polyphenol oxidases of the laccase type and may act at unknown steps in the formation of condensed tannin polymers (Pourcel et al. 2005).

1.3.2.8 Anthocyanidins and anthocyanins

The first coloured compounds in the flavonoid biosynthetic pathway are anthocyanidins (pelargonidin, cyanidin and delphinidin), which are derived from colourless leucoanthocyanidins. Anthocyanins are glycosides of anthocyanidins. These molecules possess a conjugated system of double bonds, which results in the typical red or blue colours of these compounds (Figure 1L). Anthocyanins and anthocyanidins are widely distributed among gymnosperms and both monocotyledonous and dicotyledonous angiosperms. Six relatively common aglycones (pelargonidin, cyanidin, delphinidin, peonidin, petunidin, and malvidin) are found in most cases; almost all others are rare. There are at least seven factors contributing to flower pigmentation: pH, the nature of the aglycone, the extent of glycosylation and acylation of the sugar moieties, the concentration of the anthocyanin, complexation with metals, the presence of pectins and sugars, and complexation with tannins and other phenols (copigments; Seigler 1998).

Anthocyanidin synthase (ANS) catalyzes the conversion of leucoanthocyanidins to anthocyanidins in the presence of ferrous ion, 2-oxoglutarate, and ascorbate, e.g., in snapdragon, *Petunia hybrida* (petunia), *Torenia fournieri* (torenia), and *Zea mays* (maize; Nakajima et al. 2001; Saito et al. 1999). Anthocyanidins with free hydroxyl groups are unstable under physiological pH and need to be glycosylated. The *in vitro* formation of coloured cyanidin 3-glucoside from leucocyanidin, via a cyanidin intermediate, was demonstrated using petunia ANS and UDP-glucose:flavonoid 3-*O*-glucosyltransferase (3-

GT; Nakajima et al. 2001). Fukuchi-Mizutani et al. (2003) purified a UDP-glucose:anthocyanin 3'-*O*-glucosyltransferase (3'GT) enzyme to homogeneity from *Gentiana triflora* (gentian) blue petals and isolated a cDNA encoding the enzyme. Characterization of the enzymatic properties revealed that, in contrast to most of flavonoid glucosyltransferases, it had strict substrate specificity. The best-studied flavonoid glucosyltransferases are 3-GT enzymes that have been characterized biochemically in many plant species. Anthocyanins from a variety of plants are also 5-glycosylated (Saito and Yamazaki 2002). Glycosylation at the 5-position is believed to allow for more stable complexes in copigmentation of anthocyanins, resulting in flowers with bright reddish-purple colour. The enzyme catalyzing this step, UDP-glucose:anthocyanin 5-*O*-glucosyltransferase (5-GT), has been partially characterized in several plant species and its cDNA was first isolated from *Perilla frutescens* (Yamazaki et al. 1999). The cDNA encoding UDP-Glc: flavonoid 7-*O*-glucosyltransferase (7-GT), with broad substrate specificity, was cloned from hairy root culture of *Scutellaria baicalensis* (Hirotani et al. 2000).

1.3.2.9 Other minor classes

Stilbenes are a relatively small, but widely distributed, group of plant SMs found, for example, as heartwood constituents (Seigler 1998). Over 200 naturally occurring stilbenes and stilbene glycosides are known from higher plants and the *para*-hydroxylated compound, resveratrol, is the most widespread stilbene in nature. Stilbene synthases (STSs) synthesize the same linear tetraketide intermediate from 4-coumaroyl-CoA and malonyl-CoA as CHSs, but release resveratrol as the end product of aldol condensation. STS enzymes belong also to plant type III polyketide synthase enzymes and show high sequence similarity with CHSs indicating common origin in evolution.

The bright yellow colour of a number of flowers (e.g., snapdragon, *Cosmos*, and *Dahlia*) is provided by aurones, which are 2-benzylidene-coumaran-3-ones with multiple hydroxyl groups in their aromatic rings. Nakayama et al. (2000) identified aureusidin synthase responsible for the synthesis of aurone from chalcones in the yellow snapdragon flower. Ono et al. (2006) reported that chalcone 4'-*O*-glucosyltransferase is essential for aurone biosynthesis and yellow colouration *in vivo*. Coexpression of the snapdragon chalcone 4'-*O*-glucosyltransferase and aureusidin synthase genes was sufficient for the accumulation of aureusidin 6-*O*-glucoside in transgenic *Torenia hybrida* flowers.

Other minor classes of flavonoids are dihydrochalcones, *C*-glycosylflavonoids, isoflavanones, isoflavans and isoflavenes, 3-arylcoumarins, tetracyclic isoflavonoids, flavonoids containing nitrogen, and neoflavonoids.

1.3.3 Regulation of the flavonoid pathway

Controlled transcription of biosynthetic genes is a major mechanism regulating secondary metabolite production in plant cells. The regulation is achieved by specific transcription

factors that are sequence-specific DNA-binding proteins that interact with the promoter region of their target genes. Regulatory genes that control the expression of the structural genes of the flavonoid biosynthetic pathway have been identified in many plants. The best-characterised model systems are snapdragon, petunia, maize, and *Arabidopsis*. The identification of a large number of genes shows that the flavonoid synthesis is regulated by an unexpectedly complex network. However, it is still largely unclear how a cell determines which class of flavonoid it synthesizes.

The flavonoid pathway has been shown to be activated by similar MYB, bHLH, and WD40 proteins in a wide variety of plant species (Koes et al. 2005). Plant MYB proteins are transcriptional regulators that are characterized by either two (R2 and R3) or three (R1, R2, and R3) imperfect repeats of the MYB DNA-binding motif that define their binding specificity to the target gene promoters (Martin and Paz-Ares 1997). All MYB proteins show considerable sequence divergence outside of the MYB motif. bHLH proteins have a basic helix-loop-helix domain and show substantial amino acid conservation throughout the protein (Ramsay and Glover 2005). The basic region of the domain is responsible for sequence-specific DNA binding to a consensus hexanucleotide E-box (CANNTG) whereas the HLH region allows homo- and heterodimer formation. WD40 repeat proteins comprise a family that is characterized by a glycine-histidine (GH) dipeptide and a tryptophan-aspartate (WD) dipeptide motif (Smith et al. 1999). This motif is tandemly repeated four to sixteen times in the same protein. A common function of WD40 repeat units is that they facilitate protein-protein interactions and have no intrinsic enzymatic function (Ramsay and Glover 2005). It has been shown that MYB and bHLH proteins can interact physically indicating that they may operate in one transcription activation pathway and may activate their target genes as a complex. MYBs can act as master regulators of phenylpropanoid biosynthesis or activate only a specific part of biosynthesis pathways. The overexpressed PAP1 MYB transcription factor in *Arabidopsis* enhanced the expression of *PAL, CHS*, *DFR*, and the gene encoding glutathione *S*-transferase, resulting in enhanced product accumulation with increases in lignin, wall-bound hydroxycinnamic acid esters, flavonols, and anthocyanins (Borevitz et al. 2000). The expression of *Vitis vinifera* (grape) *MYB5a* gene in tobacco induced accumulation of anthocyanins, flavonols, and tannins as well as altered lignin metabolism (Deluc et al. 2006). The more specific regulators are, e.g., the subgroup 7 (SG7) of the R2R3-MYB factors from *A. thaliana* that are flavonol-specific transcriptional regulators (Mehrtens et al. 2005; Stracke et al. 2007).

1.4 Secondary metabolites and genetically modified plants

Genetic modification (GM) has been used as a research and breeding tool for plants since the first genes were transferred to plants (Bevan et al. 1983; Herrera-Estrella et al. 1983). A number of GM crop plants (soybean, maize, cotton, and rapeseed) are today in large-scale cultivation, mainly in USA, Canada, Argentina, and Brazil (James 2007). While GM technologies are within reach of an increasing number of cultivated plants, crop improvement by GM remains controversial, at least in Europe. Many of the concerns relating to both human health and to ecological impacts relate to changes in the secondary

metabolites of GM crops. Random insertion of transgenes into the plant genome may lead to disruption, modification or silencing of active genes or to activation of silent genes, which may result in the formation of either new metabolites or altered levels of existing metabolites (unintended effects; Kuiper et al. 2001). The occurrence of unintended effects is not specific for GM, but also occurs in conventional breeding. In order to identify these effects, different strategies may be applied, e.g., targeted (compound-specific) approaches or non-targeted (profiling/fingerprinting) approaches.

Targeted studies include analyses of a number of key nutrients such as proteins, carbohydrates, fats, vitamins and other nutritional/anti-nutritional compounds which, if unintentionally modified, might affect nutritional value and safety (Kuiper et al. 2001). However, this approach has limitations with respect to unknown anti-nutrients and natural toxins, especially in less well-known crops. In non-targeted approaches potential changes of the modified organism at the genome, proteome, or metabolome level are screened. The method with best coverage is the DNA microarray technology that allows comparison of expression profiles of a large number of genes under different conditions at the same time (Schena et al. 1995). Proteomics refers to quantitative profiling of a large number of proteins in a given sample. Both DNA microarray technology (Baudo et al. 2006; Gregersen et al. 2005; Kristensen et al. 2005) and proteomics (Corpillo et al. 2004; Lehesranta et al. 2005; Ruebelt et al. 2006) have been used for searching differences between transgenic plants and their non-transgenic counterparts.

The metabolome is the final downstream product of the genome and is defined as the total quantitative collection of small-molecular-weight compounds (metabolites) present in a cell or organism. There are three general approaches to analyse small molecules: metabolite profiling, metabolic fingerprinting, and metabolomics (Last et al. 2007). The most common approach, metabolite profiling, is the analysis of small numbers of known metabolites in specific compound classes. Metabolic fingerprinting detects many compounds, but their structures are rarely identified. In metabolomics, high-throughput analytical strategies are used to identify and measure the concentration of as many metabolites as possible. A multi-compositional analysis of the biologically active compounds in plants may indicate whether intended and/or unintended effects have taken place as a result of genetic modification.

Many research groups have studied indirect and unexpected changes in secondary metabolism between transgenic plants and their non-transgenic counterparts. Kristensen et al. (2005) introduced an entire high-flux pathway, directing synthesis and intracellular storage of high amounts of a cyanogenic glucoside or a glucosinolate, into *Arabidopsis thaliana*, without pleiotropic side effects. When the maize transcription factors *Lc* and *C1* were simultaneously expressed in tomato, the expected increase in the content of six main flavonoid glycosides was observed (Le Gall et al. 2003). In addition, changes in at least 15 other metabolite levels were observed. However, these changes were described as relatively minor. Baker et al. (2006) compared three transgenic wheat lines expressing additional endogenous high-molecular-weight subunit genes and the corresponding parental lines using metabolite profiling of samples grown in replicate field trials on two sites for three years. The results demonstrated that the environment affects the metabolome and that any differences between the control and transgenic lines are

generally within the same range as the differences observed between the control lines grown on different sites and in different years. Defernez et al. (2004) analyzed metabolite profiles of about 40 GM potato lines and controls (derived from cv. Record or cv. Desireé and modified in primary carbon metabolism, starch synthesis, glycoprotein processing, or polyamine/ethylene metabolism). The most obvious differences seen in both NMR and HPLC-UV profiles were between the two varieties. Catchpole et al. (2005) made a comparison of total metabolites in field-grown GM and conventional potato tubers. A major finding was the large variation in metabolite profiles between the conventional cultivars. They also showed that the metabolite composition of field-grown GM potatoes were within the range of classical cultivars and were very similar to the progenitor lines with the exception of the products of the introduced genes. These results point out the importance of comparison with a range of equivalent cultivars instead of the single parental line when changes in metabolome of transgenic plants are studied.

Once differences have been identified, further safety evaluation of the observed differences may be needed by specific *in vivo* and/or *in vitro* testing. Presently, acute and sub-chronic toxicity studies of GM plants are performed on animals (Brake et al. 2004; Chen et al. 2003; Flachowsky et al. 2005; Hammond et al. 2006; Sanhoty et al. 2004). Predictive toxicology is defined as the extrapolation of *in vivo* toxic potential of a sample based on *in vitro* and/or *in silico* methodologies. Considering acute toxicity, the most widely applied alternative methods are *in vitro* cytotoxicity and cell viability assays, which are based on measurement of endpoints in cell number, morphology, and metabolic activity after exposure of mammalian cell cultures to the samples tested (Wilson 2000). The role of *in vitro* cell viability assays as tools for predictive toxicology has been established particularly in drug industry, where chemical agents with undesired safety profiles need to be excluded from drug discovery processes as early as possible (Slater 2001). Besides drug industry, the *in vitro* cytotoxicity assays are exploited in safety screening of food additives and other food-borne chemicals (Eisenbrand et al. 2002). The predictive value of *in vitro* cytotoxicity assays has been demonstrated by the good correlation of cytotoxicity EC50 values to *in vivo* acute toxicity in a diverse range of chemicals (Clemedson et al. 2000; Eisenbrand et al. 2002) and thus these assays provide means for preliminary safety assessment of large sample collections by relatively simple, inexpensive and ethical way.

Since the late 1980s, publications on GM crop safety and GM food/feed safety started to emerge in refereed journals (Vain 2007). Studies on the safety of GM crop cultivation focus on gene flow and effects on non-target organisms, whereas those on GM food/feed safety mostly deal with compositional, toxicological and nutritional analyses. In addition to publications in refereed journals, there are a great number of non-refereed reports and conference proceedings as well as confidential safety studies conducted by industry.

Despite the enourmous amount of research conducted on GM crop and food/feed safety, no unexpected or toxic metabolites that would endanger human or animal health have been found. Due to the long and thorough investigation process before GM plant is released to commercial market, GM plants are better characterized than conventional non-GM plants developed by other breeding methods, e.g., by the artificial manipulation of

chromosome number and chemical and radiation treatments to induce mutations and choromosome re-arrangements.

2 AIMS OF THE STUDY

In order to understand biosynthetic pathways fully, it is necessary to isolate and characterise the enzymes of the pathways to study the individual steps and the regulatory genes of the pathways. The aim of the first part of the study was to characterise novel chalcone synthase and chalcone synthase-like genes and the corresponding enzymes of the ornamental plant *Gerbera hybrida*, obtained from the gerbera EST database. In addition, the regulation of the flavonoid pathway in gerbera was studied. Putative target genes of GMYB10, a previously identified MYB transcription factor from gerbera, were identified using the gerbera 9K cDNA microarray.

When discussing the risks of GM plants, large-scale screens for unintended effects in transgenic lines are often called for, but rarely performed. The aim of the second part of the study was to find qualitative and quantitative differences in metabolic fingerprints and cytotoxicity between 225 *Gerbera hybrida* GM lines and 42 non-GM *Gerbera* varieties, and to test different *in vitro* methodologies in order to find an *in vitro* safety assessment method for screening a large variety of GM plants in standardized conditions.

3 MATERIALS AND METHODS

This section summarises briefly the methods used. For details, please refer to the original publications (I-IV).

Plant material

Gerbera hybrida var. Terra Regina and 41 other varieties were obtained from Terra Nigra B.W. (De Kwakel, The Netherlands) and grown under standard greenhouse conditions. Transgenic gerbera lines were produced from var. Terra Regina using *Agrobacterium*-mediated transformation as described by Elomaa and Teeri (2001).

Method	Used and described in study
cDNA microarray analyses	II
Cloning and protein expression in *E. coli*	I
Determination of total anthocyanin content	II
Enzyme assays	I
Preparation of *Gerbera hybrida* methanol extracts	III and IV
GT1-7 cell culture	IV
HPLC analysis of anthocyanidins	II
Lung, liver and intestinal cell cultures	III
Particle bombardment and luciferase assay	II
Propidium iodide fluorescence cytoxicity screening assay	IV
Quantitative RT-PCR	I and II
RT-PCR and PCR	I
Test for allelism	I
Thin-layer chromatography analysis (TLC)	III
WST-1 cytotoxicity assay	III

4 RESULTS AND DISCUSSION

4.1 Characterisation of new *chs* and *chs*-like genes and enzymes in *Gerbera hybrida* (I)

Previously, Helariutta et al. (1995) analysed the sequences and expression patterns of three gerbera genes homologous to *CHS* (*GCHS1*, *GCHS2* and *GCHS3*), as well as the catalytic properties of the corresponding enzymes. The comparative analysis of the amino acid sequences and analysis of the enzymatic activity of proteins produced in *E. coli* showed that GCHS1 and GCHS3 are typical CHS enzymes that produce naringenin from 4-coumaroyl-CoA and malonyl-CoA-derived acetate units. According to RNA blot analysis, *GCHS1* was expressed in pappus hairs (the specialised calyx of Asteraceae) and corolla, *GCHS3* only in pappus, but *GCHS2* expression was detected in both vegetative and reproductive organs. Helariutta et al. (1996) showed also that at least three *GCHS2*-like genes are encoded in the gerbera genome. *GCHS17* was a truncated genomic clone missing the first 56 codons, while *GCHS26* encoded a full-length protein. Eckermann et al. (1998) showed that GCHS2 is a triketide synthase that synthesises TAL from the starter substrate acetyl-CoA and 6-phenyl-4-hydroxy-2-pyrone from benzoyl-CoA. The enzyme was named to 2-pyrone synthase (2PS). Parsimony analysis of gerbera sequences together with *CHS*-like genes isolated from other taxa suggested that *2PS* has evolved from *CHS* via a gene duplication event that occurred before the diversification of Asteraceae.

4.1.1 GCHS4

Data mining for *CHS*-like sequences in the gerbera EST database (Laitinen et al. 2005) brought up *GCHS4* three times from the late petal cDNA library. The plasmid G0000100008B02 contained the full open-reading frame (ORF) for *GCHS4*. The gene encodes for a 42.9 kDa protein with 389 amino acids and a calculated pI of 6.19 (Table 1). The full-length sequence of *GCHS4* was cloned into the pQE-60 expression vector (Qiagen). The gene was functionally expressed in *E. coli* and the enzymatic activity of the recombinant protein was tested using 4-coumaroyl-CoA, benzoyl-CoA and acetyl-CoA as substrates. The GCHS4 did not accept acetyl-CoA as a substrate but used one coumaroyl-CoA and three malonyl-CoAs, yielding naringenin chalcone, via a tetraketide intermediate as did the previously identified GCHS1 enzyme used as a reference for normal chalcone synthase activity. Identity between GCHS4 amino acid sequence and GCHS1/GCHS3 is 80-88%, and between GCHS4 and gerbera 2PS/GCHS26b/GCHS17b 68-70% (Table 2). The enzyme activity studies in combination with the results of sequence analysis show that GCHS4 is a new CHS enzyme in gerbera. According to quantitative RT-PCR analysis, *GCHS4* gene is expressed mainly in corolla, though weak expression can be seen in leaf petiole, stamen, and carpel. In transgenic gerbera calli expressing the anthocyanin regulator *GMYB10* (Elomaa et al. 2003), *GCHS4* shows strong transcriptional induction based on quantitative RT-PCR analysis (II). Helariutta et al. (1995) demonstrated that

GCHS1 expression in corolla is temporally related to both anthocyanin and flavonol accumulation during corolla differentiation. In anti-*GCHS1* gerbera transformant lines (Elomaa et al. 1993) that show reduction in anthocyanin pigmentation and CHS enzymatic activity, expression of *GCHS3* and *2PS* is also suppressed (Elomaa et al. 1996). Since amino acid sequences of GCHS1, GCHS3 and GCHS4 are all similar, it can be assumed that also the *GCHS4* gene was down-regulated in anti-*GCHS1* transformants. GCHS1 and GCHS4 are both expressed in corolla and have been related to the anthocyanin biosynthesis in gerbera. One of the enzymes or both together contribute probably mainly to flower colour of gerbera. The exact roles of the gerbera CHS enzymes in flavonoid biosynthesis remain unclear using the antisense approach mentioned above due to the similarity of the sequences. Burbulis and Winkel-Shirley (1999) demonstrated specific interactions between CHS, CHI and F3H (flavonol-3-hydroxylase) enzymes in *Arabidopsis* using a yeast-two-hybrid system, affinity chromatography and co-immunoprecipitation. They suggested that the enzymes of flavonoid biosynthetic pathway may assemble as macromolecular complexes with contact between multiple proteins. It would be interesting to study the possible interactions between gerbera CHS enzymes and other flavonoid biosynthetic enzymes in gerbera. Since the CHS reaction is the first step of flavonoid biosynthesis, this could reveal specific roles for different GCHS enzymes and explain, e.g., why GCHS3 is mainly expressed in the pappus but not in the corolla.

Table 1. *Open reading frames (ORFs), the number of amino acids, molecular weights and isoelectric points (pIs) of the PKS gene family in Gerbera*

Enzyme	ORF (bp)	Amino acids	Molecular weight (kDa)	pI
GCHS1	1197	398	43,5	6,8
GCHS3	1212	403	44,1	6,6
GCHS4	1170	389	42,9	6,2
2PS	1209	402	43,7	6,8
GCHS17b	1209	402	44,6	6,2
GCHS26	1212	403	44,2	6,5

Table 2. *Identities between Gerbera PKS family members based on amino acid sequences (%)*

	GCHS1	GCHS3	GCHS4	2PS	GCHS17	GCHS17b	GCHS26	GCHS26b
GCHS1	100	87.8	81.9	73.1	65.8	74.9	76.2	76.6
GCHS3		100	79.9	73.2	63.2	72.8	74.7	74.4
GCHS4			100	67.7	61.7	68.0	69.2	70.5
2PS				100	71.6	81.4	83.6	83.1
GCHS17					100	85.3	72.2	71.1
GCHS17b						100	83.6	82.9
GCHS26							100	97.8
GCHS26b								100

4.1.2 GCHS17b

GCHS17b was recovered twice (but not full-length) from the gerbera early petal cDNA library. The full-length gene was amplified by RT-PCR with gene-specific primers from early petal mRNA. *GCHS17b* gene has a 1209 bp ORF and encodes a 44.6 kDa protein with 402 amino acids and a calculated pI of 6.19 (Table 1). Amino acid identity between GCHS17b and gerbera 2PS/GCHS26b is 81-83% (Table 2). *GCHS17b* is a *PKS* gene that is expressed weakly in corolla, carpel, ovary, and capitulum. The enzymatic activity of GCHS17b was also explored using recombinant proteins produced in *E. coli*. Recombinant gerbera 2PS enzyme was used as a reference for PKS activity. Even though 2PS was able to produce the expected products in enzyme assays, GCHS17b did not accept any of the substrates applied (4-coumaroyl-CoA, benzoyl-CoA and acetyl-CoA). When the GCHS17b amino acid sequence is compared to the previously published genomic GCHS17 sequence that misses the first 56 codons (Helariutta et al. 1996), 16 nucleotide changes leading to 4 amino acid changes can be detected between these two sequences. One of the changes is a conserved cysteine (Cys84, all amino acid numbering as in *M. sativa* CHS2) that has been replaced with glycine. Lanz et al. (1991) identified the role of cysteines in resveratrol and chalcone synthases. They made site-directed mutagenesis of all strictly conserved cysteines and investigated the activity of the unchanged and mutant proteins. Only the mutation at Cys164 abolished enzyme activity completely. Interestingly, the change of cysteine to alanine also reduced the activity in most other cases, including the cysteine of GCHS17b that has been changed to Gly. Dana et al. (2006) used molecular dynamics simulations of protein structures to study the local and long-range effects of single amino acid substitutions on CHS enzyme function. They suggested that single changes of highly conserved amino acids that are not located at or near the residues of known functional significance may still alter the architecture and dynamic movement of the enzyme with dramatic effects on enzyme function. The same phenomenon may have caused the failure of the enzyme assays with GCHS17b, i.e., the protein may have deficient enzymatic activity and be encoded by a pseudogene. Alternatively, the correct substrate for GCHS17b still remains unknown or, for some reason, the recombinant enzyme was inactive.

4.1.3 GCHS26b

An internal fragment identical to sequence from the chromosomal *GCHS26* gene was observed when PKS-encoding sequences were amplified with the degenerate primers described in Helariutta et al. (1995) using gerbera root cDNA as a template (Hannu Hotti, personal communication). Using primers designed from the previously isolated genomic sequence (Helariutta et al. 1996), a full-length sequence (*GCHS26b*) was amplified from root mRNA with RT-PCR. Fourteen nucleotide changes leading to four amino acid changes were detected between the genomic fragment GCHS26 and transcribed fragment GCHS26b. GCHS26b should be cloned into an expression vector and the enzymatic activity of the recombinant GCHS6b should be analysed.

4.1.4 *GCHS26/GCHS26b* and *GCHS17/GCHS17b* represent separate loci

The genomic and the transcribed sequences of *GCHS26* vs. *GCHS26b* and *GCHS17* vs. *GCHS17b* are similar and they could represent allelic forms. In order to study whether the genomic and the transcribed sequences represent alleles of a single locus or of separate loci, we made use of a selfed progeny of the gerbera cultivar Terra Regina, from which all sequences described here were isolated. Among twenty individuals in the selfed progeny, we could detect both versions of *GCHS17* and *GCHS17b* as well as *GCHS26* and *GCHS26b* sequences in all cases. Based on this analysis, it is highly unlikely ($p<0.001$; χ^2 test for segregation) that the two pairs of similar sequences represent alleles of the same locus.

4.1.5 Comparison of CHS/PKS amino acid sequences and enzyme assays

Since Reimold et al. (1983) isolated the first *CHS* gene, a great number of *CHS* and other plant type III *PKS* genes have been identified from several gymnosperms and angiosperms. Amino acid sequences of these genes are exceedingly similar so enzyme assays and structural data are needed to make more defined predictions of the functions of these enzymes. The first insights into the structure of CHS were brought to light, when Ferrer et al. (1999) published the three-dimensional structure of CHS2 from alfalfa. They found three conserved residues (Cys164, His303, and Asn336) likely to participate in the multiple decarboxylation and condensation reactions that take place during the CHS-catalysed reaction. In addition, 15 inert active site residues were proposed to provide steric guidance to the linear polyketide. Jez et al. (2000b) determined the structure of gerbera 2PS in complex with the reaction intermediate acetoacetyl-CoA. 2PS and CHS share a common three-dimensional fold, a set of conserved catalytic residues, and similar CoA binding sites, but the active site cavity in 2PS was smaller, in concordance with the preference for a smaller substrate in 2PS.

In recent years, many functions for the amino acids in PKS enzymes have been described. The most interesting residues are the conserved Thr197, Gly256, and Ser338 in CHS. Jez et al. (2000b) showed that in 2PS, these amino acids were replaced with Leu, Leu, and Ile, respectively. They also showed that a CHS triple mutant (Thr197Leu, Gly256Leu, and Ser338Ile) yielded an enzyme that was functionally identical to 2PS. These three residues are sterically altered in a number of functionally divergent type III PKSs including aloe PCS (Thr197Met, Gly256Leu, and Ser338Val; Abe et al. 2005b), aloe OKS (Thr197Gly, Gly256Leu, and Ser338Val; Abe et al. 2005a), and rhubarb ALS (Thr197Ala, Gly256Leu, and Ser338Thr; Abe et al. 2006). Abe et al. (2005a and 2005b) have suggested that the active site position 197 determines the polyketide chain length and product specificity, and position 256 determines the substrate specificity of the enzymes (Abe et al. 2006). Gerbera CHS4 is a typical CHS enzyme having the conserved Thr197, Gly256, and Ser338 residues in the amino acid sequence (Table 3). 2PS, PCS, ALS, and OKS all use acetyl-CoA as a substrate and they all have leucine in the place of Gly256 found in CHS. Gly256 has been replaced with leucine also in gerbera CHS26/26b and

CHS17/17b enzymes. According to previous information, CHS26/26b may use acetyl-CoA as the substrate. Comparison between amino acid residues at position 197, which has been suggested to determine the polyketide chain length, reveals that the more malonyl-CoA units the enzyme adds, the smaller the amino acid residue at position 197 tends to be. In gerbera CHS26/26b and CHS17/17b enzymes, there is isoleucine at position 197, whereas in 2PS, there is leucine at that position. Isoleucine is an isomer of leucine and both are similar hydrophobic amino acids. According to previous information, it can be hypothesed that CHS26/26b and CHS17/17b are either triketide synthases similar to 2PS or at least do not make much longer polyketide chains than does 2PS.

Determination of the true functions of these gerbera CHS/PKS enzymes still needs data from enzymatic assays. Amino acid sequence of the GCHS4 contains the conserved residues found in all other CHS enzymes. The enzyme does not accept acetyl-CoA as a substrate but synthesizes naringenin from 4-coumaroyl-CoA and three malonyl-CoA units. The precise functions of GCHS17/17b and GCHS26/26b remain unclear. The recombinant GCHS26 protein does not accept 4-coumaroyl-CoA as a substrate but makes the same end products as the recombinant 2PS using benzoyl-CoA and acetyl-CoA as substrates. GCHS26b should also be cloned into an expression vector and the enzymatic activity of the recombinant GCHS6b should be analysed in detail and compared to GCHS26 and 2PS as well as GCHS17/17b activities.

Table 3. *Active sites of chalcone synthase enzymes*

Alfalfa CHS2[1]	GCHS1	GCHS3	GCHS4	2PS	GCHS17b	GCHS26
Cys 164	Cys	Cys	Cys	Cys	Cys	Cys
His 303	His	His	His	His	His	His
Asn 336	Asn	Asn	Asn	Asn	Asn	Asn
Val 98	Val	Val	Val	Val	Val	Ile
Thr 132	Thr	Thr	Thr	Thr	Thr	Thr
Ser 133	Ser	Ser	Ser	Ala	Ala	Ala
Met 137	Met	Met	Met	Met	Met	Met
Thr 194	Thr	Thr	Thr	Thr	Thr	Thr
Val 196	Val	Val	Val	Ile	Ile	Ile
Thr 197[2]	**Thr**	**Thr**	**Thr**	**Leu**	**Ile**	**Ile**
Gly 211	Gly	Gly	Gly	Ala	Gly	Gly
Phe 215	Phe	Phe	Phe	Phe	Phe	Phe
Gly 216	Gly	Gly	Gly	Gly	Gly	Gly
Ile 254	Ile	Ile	Ile	Met	Met	Met
Gly 256[3]	**Gly**	**Gly**	**Gly**	**Leu**	**Leu**	**Leu**
Phe 265	Phe	Phe	Phe	Phe	Phe	Phe
Ser 338	**Ser**	**Ser**	**Ser**	**Ile**	**Thr**	**Thr**
Pro 375	Pro	Pro	Pro	Pro	Pro	Pro

[1] Structure determined by Ferrer et al. 1999

[2] Determines the polyketide chain length and product specificity (Abe et al. 2005a and 2005b).

[3] Determines substrate specificity of the enzymes (Abe et al. 2006)

4.1.6 Tissue-specific expression of the *PKS* genes

Spatial expressions of the gerbera *PKS* genes were analysed with quantitative RT-PCR in 12 tissues from several developmental stages and flower types. These included root, leaf blade, leaf petiole, scape, bracts, receptacle, pappus, corolla, stamen, carpel, ovary, and young developing capitulum. For each *PKS* gene, two biological replicate samples were analysed in triplicate. The traditional way of analysing gene expression profiles is RNA blot analysis (Helariutta et al. 1995; Kotilainen et al. 1999). The gerbera *PKS* genes are so similar that it was not possible to design probes that would have distinguished the genes in RNA blot analysis. As designing gene-specific primers for quantitative RT-PCR was possible, we decided to test the suitability of real-time RT-PCR for investigating the gene expression profiles. To start with, we tested several different available gerbera genes in order to find a good reference gene. The best reference gene was ubiquitin (EST G0000100001G09; Laitinen et al. 2005).

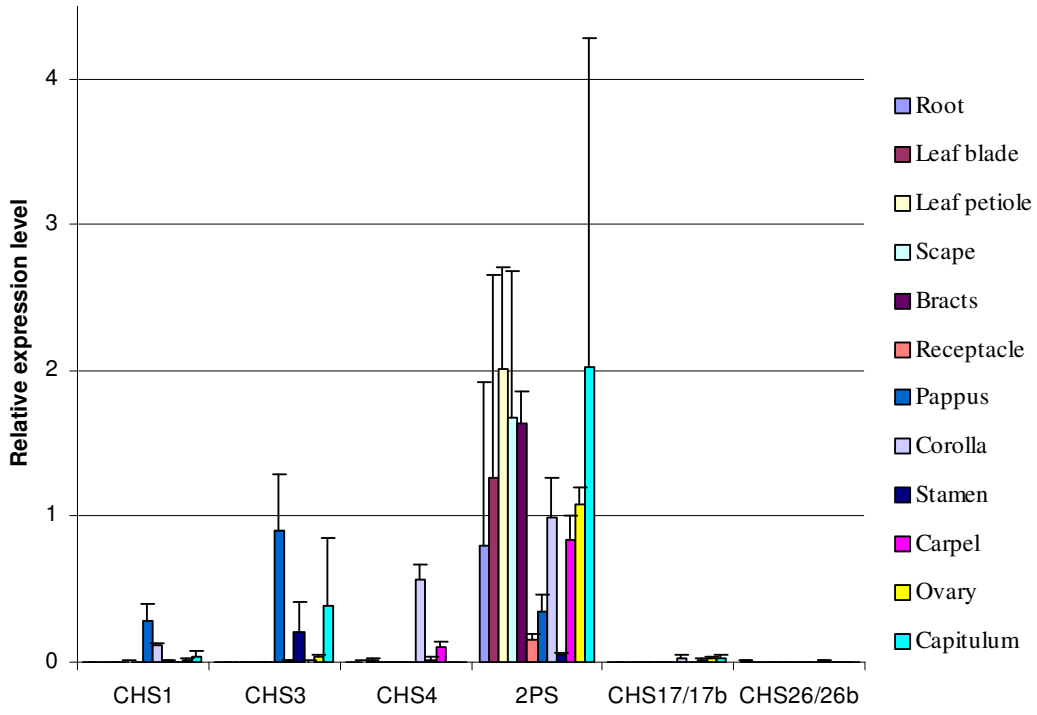

Figure 3 *Expression of the PKS gene family in gerbera*

Figure 3 shows the expression patterns of the gerbera *PKS* genes in 12 tissues. These expression data are more detailed than the previously published expression profiles of *GCHS1*, *GCHS3*, and *2PS*, obtained with RNA blot analysis (Helariutta et al. 1995). Both methods showed that *GCHS1* was expressed in pappus and corolla and *GCHS3* in pappus.

In addition, *GCHS1* was expressed in stamen, ovary, and capitulum and *GCHS3* in corolla, stamen, carpel, ovary, and capitulum using quantitative RT-PCR. Quantitative RT-PCR showed that *2PS* was expressed in all vegetative and reproductive tissue samples, whereas RNA blot analysis revealed no expression in root, bract, and stamen samples. The *GCHS4* gene was expressed in leaf petiole, corolla, stamen, and carpel. The primers used in quantitative RT-PCR did not distinguish between *GCHS17* and *GCHS17b*, or between *GCHS26* and *GCHS26b*. The *GCHS17/17b* genes were expressed weakly in corolla, carpel, ovary, and capitulum whereas *GCHS26/26b* expression could be detected only in root and stamen.

The *PKS* gene family in gerbera can be divided into two groups based on the similarities in the amino acid sequences of the enzymes and the known functions of CHS1, CHS3, CHS4 and 2PS. The CHS-encoding genes *CHS1*, *CHS3* and *CHS4* form the first group. The second group consists of the known *2PS* gene and the other four genes (*CHS17/17b* and *CHS26/26b*) that differ from the strict CHS consensus sequence and are similar to 2PS in their amino acid sequence. The true CHS group is expressed only in reproductive tissues except for one example, i.e., *GCHS4* expression in leaf petiole. This specific expression can be explained by the roles of gerbera CHSs in anthocyanin biosynthesis that determines flower colour formation. From the second group, 2PS is the only functionally characterised enzyme. It has been shown to produce triacetic acid lactone (TAL; Eckermann et al. 1998), which is a putative precursor for two glucosides of gerbera, gerberin and parasorboside. The enzymatic reactions from TAL to these glucosides still need to be characterised. There are also indications that gerberin or parasorboside, or their aglycones, are involved in the defence against biotic stress factors. 2PS-antisense gerberas that lack gerberin and parasorboside (Eckermann et al. 1998) are more susceptible to infestations by fungi than the wild type Terra Regina varieties (Satu Koskela, personal communication). 2PS expression was detected in all tissues studied. The relative expression of the *2PS* gene in root, leaf blade and petiole samples also varied between two biological replicates, which suggests that the gene is inducible. These characteristics, the general expression pattern and the inducibility, support the hypothesis that 2PS is involved in defence. The function of *CHS17/17b* and *CHS26/26b* is an interesting issue. Their expressions are very weak. GCHS17b did not accept three offered substrates used whereas GCHS26 showed enzymatic activity. Are all four enzymes really active? Do the enzymes also produce TAL or make longer polyketide chains? Only studies using all enzymes and many different substrates may solve the mystery.

Since quantitative RT-PCR is a more sensitive method than RNA blot analysis, it detects weaker expressions that would otherwise be missed, but there are also many difficulties in using it reliably. When an accurate absolute or relative quantification is the goal and many tissues are investigated, the most difficult task is to find an appropriate reference gene. All commonly used universal reference genes are regulated to some extent and none are constitutively expressed in all cell types and under all conditions, independently of experimental design (Bustin et al. 2005). An alternative is to use multiple reference genes (Vandesompele et al. 2002). The drawback is that the method requires extensive practical validation to identify a combination of reference genes appropriate for individual experiment. Normalisation against total RNA concentration is also an

alternative (Bustin 2002; Tricarico et al. 2002), although it is not ideal for quantifying gene expression between tissues in which transcriptional activity may vary. In addition, it is not always possible to quantify total RNA.

4.2 Regulatory genes of the anthocyanin pathway in *Gerbera hybrida* (II)

4.2.1 R2R3-type MYB transcription factor GMYB10

Elomaa et al. (2003) identified an R2R3-type MYB transcription factor from gerbera, *GMYB10*, which shared high sequence similarity to and was phylogenetically grouped together with many previously characterised regulators of anthocyanin pigmentation in petunia and *Arabidopsis*. The gene was able to induce anthocyanin pigmentation in *Nicotiana tabacum* (tobacco) leaves, stems, sepals, anthers, and ovaries and in undifferentiated calli of gerbera (Elomaa et al. 2003). Using yeast two-hybrid assay, it was also shown that GMYB10 interacts with gerbera bHLH transcription factor GMYC1. In this study, phenotypes of mature transgenic gerbera plants overexpressing *GMYB10* gene (Lines Tr1, Tr8 and Tr14) were investigated. These lines have altered pigmentation pattern and colour intensity compared to the control (II, Figure 1). The overexpression of *GMYB10* is able to induce anthocyanin pigmentation in both vegetative and floral tissues in contrast to the previous hypothesis according to which GMYB10 needs a bHLH partner to induce the accumulation of gerbera anthocyanins (Elomaa et al. 2003). Studies with healthy green leaves of gerbera and red non-transgenic leaves of gerbera showed that *GMYB10* expression correlated with the accumulation of cyanidin pigments (data not shown). It can be concluded that abiotic or biotic stress might be able to induce *GMYB10* expression, which might lead to accumulation of anthocyanins in otherwise acyanic parts of the plant. Activation tagging in *Arabidopsis* and tomato has also identified MYB-type regulatory genes, *PAP1* and *ANT1*, respectively, which both share high similarity with *GMYB10*. Constitutive overexpression of these genes in activation-tagged lines of *Arabidopis* and tomato also lead to the formation of intense purple colouration in vegetative organs (Borevitz et al. 2000; Mathews et al. 2003).

The total anthocyanin content and anthocyanidin profiles of control and transgenic samples were compared spectrophotometrically and with HPLC, respectively. The samples included green callus tissues transformed with an empty vector (without the transgene insert) and purple calli overexpressing *GMYB10*. The ray flower petals were separated into two samples containing either the ligular part of the petals or the basal tube area. The disc flower samples contained either the unpigmented (control) or fully red/purple (transgenic) stamens, respectively. Both quantitative and qualitative changes were observed in anthocyanin pigment accumulation in the transgenic lines (II, Table S1). Non-pigmented control calli appeared completely devoid of anthocyanin pigments (<0,072 mg/g DW) while the red calli accumulated on average 4,73 mg/g DW. Thus, at

least 66-fold increase in pigment levels was observed. In floral tissues, Tr14 showed the most severe phenotype with at least 75-fold increase in total anthocyanin concentration in petal tubes, at least 4-fold increase in ligules of the petals and at least 27-fold increase in whole disc flowers (which contain the differentially pigmented stamens). In Tr1 and Tr8 lines, the changes were more modest although clearly enhanced pigmentation could be measured. HPLC was used to analyze the amounts of anthocyanidins (delphinidin, cyanidin, and pelargonidin). The compounds were identified based on their retention times and UV-visible absorption spectra in comparison with commercially available reference compounds and quantified using external standards. In green tissues such as callus and leaves, overexpression of *GMYB10* led to specific activation of cyanidin biosynthesis (II, Figure 2). Tohge et al. (2005) have also shown that overexpression of *PAP1* in *Arabidopsis* leads to specific accumulation of cyanidin-type anthocyanins and quercetin-type flavonols. In non-transgenic gerbera plants, petal ligules and disc flowers (developmental stage 7) accumulated only pelargonidin while in petal tubes, no anthocyanidin accumulation was detected. In Tr14, cyanidin was the major anthocyanidin in all these floral tissues but also enhanced pelargonidin levels were detected. In ligules, pelargonidin concentrations were more than 2-fold increased in comparison to the control sample (II, Figure 2). In short, the overexpression of *GMYB10* caused specific changes in the accumulation of anthocyanins. Cyanidin and pelargonidin started to accumulate in acyanic tissues of gerbera (stamens and petal tubes). Thus, the lack of anthocyanin pigments in these tissues may be due to the lack of functional MYB factor. Another possibility is that *GMYB10* overexpression leads to induction of some unknown downstream regulators that then directly activate the biosynthetic genes. Also the amount of pelargonidin was increased in tissues where pelargonidin was also normally accumulating. This is probably caused by an indirect effect. The early flavonoid pathway produces more precursors to both pelargonidin and cyanidin routes.

The gerbera 9K cDNA microarray (Laitinen et al. 2005) was used to compare gene expression profiles of transgenic tissues against the corresponding control tissues to reveal putative target genes for GMYB10. Callus, stamen and petal samples (containing both ligule and tube tissues) were included in the analysis. The expressions of seventeen genes, identified as upregulated in microarrays, were verified using quantitative RT-PCR. GMYB10 affected the expression of both early and late anthocyanin biosynthetic genes in anthocyanin-accumulating transgenic tissues. In the cyanidin-accumulating calli and stamens, cDNA homologs of phenylalanine ammonia-lyase, cinnamate-4-hydroxylase (*C4H*), chalcone synthase, flavanone 3-hydroxylase and flavonoid 3' hydroxylase (*F3'H*) were upregulated. Chalcone isomerase, flavanone 3-hydroxylase and anthocyanidin synthase were upregulated in transgenic petal tissues. Very similar results have been shown for *ANT1* in tomato and *PAP1* in *Arabidopsis* (Borevitz et al. 2000; Mathews et al. 2003; Tohge et al. 2005). The expression of *F3'H* gene was upregulated in all analysed transgenic tissues (callus, stamens, petals), suggesting that it is the key target gene for GMYB10 needed for opening the cyanidin pathway in gerbera tissues. In addition, a new member of the chalcone synthase gene family (*GCHS4*) was extremely highly upregulated in all tissues. GCHS4 is a typical CHS enzyme using 4-coumaroyl-CoA as a substrate (I). Its expression also correlates with the accumulation of anthocyanins. It has been

previously shown that the expression of *GCHS1* correlates temporally with the accumulation of anthocyanins while the expression of *GCHS3* correlates with the accumulation of flavonols (Helariutta et al. 1995). These genes were, however, not upregulated by the overexpression of *GMYB10*. This supports the hypothesis of Burbulis and Winkel-Shirley (1999) that the enzymes of flavonoid pathway may assemble as multienzyme complexes that direct the biosynthesis of different end products of the flavonoid biosynthesis. GCHS4 could be involved in a multienzyme complex that is specifically needed for cyanidin production.

4.2.2 GMYB11 and GMYB12

Two new MYB domain factors, named as GMYB11 and GMYB12, were also upregulated in microarray analysis by *GMYB10* overexpression. *GMYB11* (G0000100029D05) was highly upregulated in all pigmented tissues overexpressing *GMYB10* while *GMYB12* (G0000300014D01) was specifically upregulated in transgenic petal tissues. Northern analysis of wild type Terra Regina revealed that *GMYB11* is highly petal-specific and not detected in other organs (II, Figure 4). Its expression starts at stage 5 of petal development, correlating with the accumulation of anthocyanin pigments, and continues until stage 8 when petals are fully expanded and the inflorescences open. *GMYB12* is more ubiquitously expressed. Highest expression of *GMYB12* is detected in leaves and petioles, as well as in all floral organs. Its expression shows the opposite pattern to *GMYB11* during petal development, as it is most highly expressed at early stages before visible anthocyanin accumulation.

The functionality of the GMYB11 and GMYB12 as MYB transcriptional factors regulating the biosynthesis of anthocyanins was studied using particle bombardment for transient expression of these genes alone or together with the GMYC1, a bHLH factor from gerbera, and their putative target promoter, *Pgdfr2*, which was fused to the firefly luciferase reporter gene. Elomaa et al. (1998) showed with the same method that GMYC1, together with a heterologous MYB regulator from petunia, was able to activate a gerbera *dfr* promoter which suggested that GMYC1 has a major role in regulating anthocyanin biosynthesis in gerbera. The GMYB10 gene together with GMYC1 was used as a positive control, and 35S-*ruc* was used as an internal control and was co-bombarded in each case. GMYB10 alone was not able to activate the gerbera *dfr* promoter but together with GMYC1, the promoter was activated. Neither GMYB11 nor GMYB12 alone or together with GMYC1 were able to activate the gerbera *dfr* promoter (data not shown), which suggests that these genes are not involved in dihydroflavonolreductase regulation.

4.3 Safety assessment of genetically modified gerbera lines and non-GM gerbera varieties (III and IV)

The use of GM plants in agriculture and food production is the object of an intense and divisive debate. There is a gap between attitudes towards GM crops in Europe and the rest

of the world. In Europe, people are deeply suspicious about GM plants regardless of the massive amount of research conducted on GM crop and food/feed safety for the last two decades, whereas in the other parts of the world GM plants are gaining more and more acceptance. In many publications, only a few transgenic lines and their safety have been studied by using various methods. It is difficult to find a collection of transgenic lines of the same plant species for safety research purposes.

The Gerbera laboratory (Department of Applied Biology, University of Helsinki, Finland) has studied a large number of genes related to metabolism and development in the ornamental plant *Gerbera hybrida* using reverse genetics (e.g., Kotilainen et al. 2000; Teeri et al. 2006). There is a large collection of transgenic gerbera lines where different gerbera genes (encoding transcription factors, enzymes and cell wall proteins) have been cloned under the 35S promoter in sense, antisense or hairpin loop (RNAi) orientation. These transgenic gerbera plants have been essential in providing information on the function of the transferred genes (Elomaa et al. 1993; Elomaa et al. 1996; Elomaa et al. 2003; Kotilainen et al. 1999; Uimari et al. 2004).

This large collection of transgenic gerberas (III, Table 1) was utilised in investigating indirect and unexpected changes in secondary metabolism using cytotoxicity assays and metabolic fingerprinting analysis. Altogether 225 GM lines were analysed, representing 1-20 independent lines per construct. As a control group, 42 gerbera varieties (III, Table 2) developed by traditional methods of breeding (crossing and selection), were used in assessing whether the changes in GM plants are qualitatively or quantitatively different from variation among non-GM cultivars.

4.3.1 Cell viability assays

For risk assessment of genetic modifications, the effects of methanol extracts of GM and non-GM gerbera were evaluated with a validated cytotoxicity screening assay WST-1 using three continuous human cell lines derived from lung, liver, and intestinal tissues that represent the potential targets to be exposed to plant metabolites. In addition, a high-throughput cytoxicity screening assay based on the use of mouse GT1-7 immortalized neuronal cells stained with propidium iodide was developed.

The WST-1 assay represents the second generation of tetrazolium salt tests allowing simplified assay protocol and thus higher throughput capacity (Ishiyama et al. 1995). Prior to safety screening, the assay was automated and validated to ensure the assay robustness and uniform data quality (Pohjala et al. 2007). The pre-set limit values for the assay quality parameters signal-to-background S/B, signal-to-noise S/N, and assay quality coefficient Z' (Zhang et al. 1999) as well as the thresholds set for the positive control (polymyxin B sulphate) were met throughout the screening process. In the assay, the cells were seeded into 96-well plates and the plates were incubated for 20-24 h prior to exposure to the extracts. Exposure was performed at extract concentrations 40 µg/ml and 100 µg/ml. The higher concentration (100 µg/ml) is a practical upper limit for solubility of the sample, and was used to test toxicity at highest possible concentration. The lower concentration (40 µg/ml) was included to record possible dose effects. Each sample had

three replicates. After 24 h exposure, cultures were washed and WST-1 cell proliferation reagent (Roche Diagnostics) was added. The tetrazolium salt was reduced to coloured formazan during 1 h incubation at 37°C with shaking. After incubation, colorimetric readout at 440 nm was performed. Cell viabilities were expressed as percentages of formazan absorbance in each sample compared to the untreated control. A threshold of >20% difference in cell viability (<80% or >120%) was set to indicate a significant effect.

Hit extracts, i.e., extracts scoring below the lower or above the upper pre-set threshold, were found in both GM and control group when screened against each cell line (III, Table 3). However, the statistical frequency analysis of these hit extracts showed that the GM lines and traditional varieties did not differ in cytotoxicity. Quantitatively, all observed effects on cell viability were minor in nature, as no decrease on cellular viability greater than 30% was observed, and for increased viability, 1.5-fold changes were the largest ones found. No difference between GM and control groups was found in this respect either. The complex nature of the crude inflorescence extracts may contribute to the fluctuation in the data, and thus offer explanations for certain threshold breakers. For instance, the likelihood of the presence of different carbohydrates may have a positive effect on cell proliferation, giving rise to increased metabolic activity.

The GT1-7 based cytotoxicity assay was also proved to be reliable, with good quality (Z'= 0.51) and low plate-to-plate and day-to-day variations. In the assay, all gerbera extracts produced cell viability values that fell within the established hit limits for cell viability values and thus no hit extracts were found. The average cytotoxicity registered for GM samples was also found statistically similar to the one registered in non-GM extracts. In general, these results indicate that the safety profile of the extracts from the GM gerbera lines is similar to the viability pattern caused by the extracts from the non-GM gerbera varieties with four different cell lines (intestinal epithelial, bronchial epithelial, hepatocellular and hypothalamic cells) and with at least two different cell viability assays.

4.3.2 Metabolic fingerprinting analysis

Metabolic fingerprinting analysis was chosen to compare the compounds in methanol extracts of GM and non-GM gerberas. The main aim was to screen indirect changes in the secondary metabolites of these plants. 42 different traditional non-GM gerbera cultivars and 2 non-transgenic regenerated lines (from var. Terra Regina) were used as a control group to 206 GM gerbera lines that were transformed with altogether 45 different gene constructs. All transformants were derived from the variety Terra Regina that was included in the control group. The inflorescence samples were extracted with methanol using ultrasonic extraction and analysed with TLC using two different mobile phases. Mobile phase A (methanol:dichloromethane, 15:85) separates glucosides and aglycones, whereas mobile phase B (water:formic acid:acetic acid:ethyl acetate, 23:10:10:91) separates different glucosides. The absorbance of compounds separated using mobile phase A was measured under UV light. The TLC plates run with mobile phase B were

dipped in sulfuric acid:methanol (15:85), heated at 120ºC for 10 min and then visualized under white light.

Altogether 33 different zones were detected and quantified as their integrated area values. Mobile phase A separated 15 different zones under UV light and mobile phase B separated 18 different zones under white light. Quantified area values from two different TLC runs were combined and the data were normalized. The 33 variables were analysed by principal component analysis (PCA), the nearest neighbour classifier, and Fligner-Killeen test (R Development Core Team 2005).

PCA is a technique that can be used to simplify and reduce the dimensionality in a dataset while retaining those characteristics of the dataset that contribute most to its variance. It is a linear transformation that chooses a new coordinate system for the dataset so that the greatest variance by any projection of the data set comes to lie on the first axis (called the first principal component), the second greatest variance on the second axis, and so on. It has been used widely in assessing the differences between transgenic plants and their non-transgenic counterparts both at the proteome and at the metabolome level (Baker et al. 2006; Catchpole et al. 2005; Defernez et al. 2004; Le Gall et al. 2003; Lehesranta et al. 2005; Shepherd et al. 2006). Lehesranta et al. (2005) used PCA to explore whether one or more transgenic potato lines or varieties would separate from others and to identify groups of proteins that, in combination, had different expression levels among the 2-DE gels. They showed at the proteome level that there was much less variation between GM lines and their non-GM controls compared with that found between different varieties and landraces. When GM gerbera lines were compared to the control group by PCA, no different groups were formed that could separate GM and non-GM plants (III, Figure 1). Still, it was observed that independent lines for several transferred gene constructs clustered near each other in the PCA projection. For example, lines where the chalcone synthase-like gene *g2ps* is down-regulated leading to the dramatic reduction of the two main metabolites in gerbera, gerberin and parasorboside (Eckermann et al. 1998), are clustered together far away from the other GM lines and varieties where gerberin and parasorboside can be found. Interestingly, also the two non-transgenic regenerated lines from variety Terra Regina (TAT751b Tr6 and MAT102 Tr8) are in the same spot in the PCA projection.

In the nearest neighbour classifier, a set of data points in d-dimensional space are pre-processed into a data structure, so that given any query point q, the nearest points to q are reported. 206 transgenic gerbera lines were divided into 18 groups according to the supposed function of the transgene (III, Table 1). Using the 33 area values gained with TLC, also the nearest neighbour classifier showed that the quantified and normalised area values of lines with identical or similar gene constructs clustered. 54% of the GM lines were clustered into the expected group and 66% of the GM lines found at least the second best neighbour within the expected group. In fact, clustering was observed in the nearest neighbour analysis of transgenic lines from constructs where previous analysis had concluded that they do not express a phenotype. Obviously, a more careful chemical analysis is necessary in these cases.

As shown previously by Catchpole et al. (2005), it was expected that metabolic fingerprints in the non-GM gerbera variety group would vary, qualitatively and

quantitatively, more than in the GM gerbera group, which was derived from a single variety. The same compounds (TLC zones) were observed in both groups, and no new compounds were found from the transgenic group compared to the control. The Fligner-Killeen nonparametric test was used to test the homogeneity of variances of the 33 TLC variables between transgenic gerbera lines and varieties, i.e., to query if the amounts of compounds vary more in GM gerberas than in the non-GM varieties. p values from 33 tests were corrected with Bonferroni correction and False Discovery Rate. In 29 of the 33 normalized area value groups, no significant differences in variance were found. In 3 groups, the variances of non-GM varieties were significantly greater than in transgenic lines as expected. In one group, the variance of transgenic gerbera lines was significantly greater than in varieties with both corrections, i.e., the amounts of one compound varied more in GM gerberas than in non-GM ones. It would be interesting to identify this compound in order to explain the reason for the variation.

The safety study using a large collection of GM gerberas did not show anything deviant about the safety of GM plants compared to the previous publications. Higher number of GM lines were screened than in previous studies but the main result, i.e. that there were no differences in cytotoxicity or metabolic fingerprints between GM gerbera lines and varieties, remained the same. Plants have been bred for thousands of years. Using traditional breeding methods, breeders can produce a desired trait into a plant, but at the same time, they mix thousands of genes, requiring many years of back-crossing to remove the unwanted traits that occur. Modern plant breeding using GM is a more precise and faster method, because a breeder can give a specific, new characteristic to a plant without transferring undesirable traits, just by adding a single gene (or a few genes) into a plant. When GM was started to use as a breeding method, people focused their attention on the method of breeding for the first time and questioned it. The new gene technology had scared the people and the main point, characteristics of the bred line, was forgotten. Using oligonucleotide arrays, Batista et al. (2008) analysed the gene expression of two GM rice plants and two traditionally bred lines, γ-irradiated mutants, as well as the respective controls. It was shown that the improvement of a plant variety through the acquisition of a new desired trait, using either mutagenesis or GM, may lead to a stress-related altered expression of untargeted genes. Interestingly, the observed alteration was more extensive in mutagenized than in transgenic rice plants in all the cases studied. These results point out that the method used in breeding should not be the main focus but the bred line and its characteristics as such. Every new plant line regardless of the way it has been bred should be evaluated in the same way and with the same principles.

5 CONCLUSIONS

The roles of flavonoids are diverse from flower pigmentation to nutritional and pharmaceutical value to humans. Due to this importance, the genetic modification of the pathway is an attractive target. Metabolic engineering may be targeted to produce novel flower colours and colour patterns as well as to increase or otherwise alter the levels of compounds with health benefits. For example, blue rose (*Rosa hybrida*) has been just a dream for many years but now it is reality, at least in a research laboratory (Katsumoto et al. 2007). In order to modify a biosynthesis pathway, it is necessary to understand the whole pathway from single genes and enzymatic steps to the regulation of the pathway. In the first part of this study, the first enzymatic step of the flavonoid biosynthetic pathway in gerbera was studied. It was shown that, in addition to the previously characterised GCHS1 and GCHS3, a new typical CHS enzyme in gerbera, GCHS4, is involved in anthocyanin biosynthesis. The effects of the overexpression of *GMYB10*, a known regulator of anthocyanin pathway in gerbera, were also investigated. *GCHS4* gene was highly upregulated, opposite to other *GCHS* genes, in tissues where *GMYB10* was overexpressed. The overexpression of *GMYB10* caused the accumulation of cyanidin in otherwise acyanic tissues of the gerbera. In addition, the amounts of pelargonidin were increased. It can be suggested that the *GCHS4* is specifically involved in cyanidin production whereas other GCHS genes have roles in the pelargonidin biosynthesis.

Depending on the goals of modification, it should be decided if single-enzyme modifications are used or if control of multiple steps is needed. Generally, decreasing the accumulation of a compound has been easier to achieve than increasing flux by modifying the activity of single enzymes (Broun and Somerville 2001). Although single enzymes can enhance flux, control of multiple steps involving the simultaneous overexpression and/or suppression of multiple genes may be necessary to achieve net gains in product accumulation. The use of regulatory genes, e.g., transcription factors to control the abundance or activity of several enzymes is also becoming more widespread (Capell and Christou 2004). Elomaa et al. (1993) inhibited the formation of flower pigmentation by transferring the single *GCHS1* gene into the gerbera in antisense orientation. The wild type red gerberas became white. Thus, the reduction of pigmentation was achieved relatively easily. No single biosynthetic gene that has been transferred into gerbera has been able to increase the inflorescence pigmentation. The increase in gerbera inflorescence pigmentation was achieved when the known regulator of the anthocyanin pathway, *GMYB10*, was overexpressed. It can be concluded that the regulatory genes can be used to modify secondary metabolite pathways in a highly predicted manner.

In spite of the increasing knowledge on biosynthetic pathways and their regulation, as well as improved gene technology, the major obstacle to using GM horticultural and crop plants is the public perception of these plants, at least in Europe. The negative European attitude can also affect the developing countries in Africa and in Asia. A good example is the Golden Rice (Ye et al. 2000), which accumulates provitamin A (β-carotene) in the grain and was supposed to contribute to the reduction of clinical and sub-clinical vitamin A deficiency-related diseases in developing countries. However, the present strict testing of the invention prevents the use of it at the moment. Thus, studying the

unintended effects on metabolism in transgenic lines and developing high-throughput screening assays of analysis of adverse effects of GM crops is extremely important. To conclude from safety assessment of GM gerbera lines and non-GM gerbera varieties, there was no evidence for any major qualitative or quantitative changes between GM lines and varieties. In addition, the *in vitro* safety profiling methods of plant extracts applied in this research provide valuable information for GM gerbera risk assessment purposes and offer a model scheme for cell-based cytotoxicity screening of a large variety of GM plants in standardized conditions.

6 ACKNOWLEDGEMENTS

This study was carried out at the Department of Applied Biology in the University of Helsinki. Funding was provided by Helsinki Graduate School of Biotechnology and Molecular Biology and by the ESGEMO (Environmental, Societal and Health Effects of Genetically Modified Organisms) programme of the Academy of Finland (grant no 207410).

I wish to thank the three people who have contributed most to this work: I am grateful to my supervisor Professor Teemu Teeri for your guidance and for sharing your deep knowledge in plant science and molecular biology as well as the continuous enthusiasm for making science. Your inspiring attitude creates a true scientific atmosphere in the laboratory. The famous crayfish parties have also been an annual enjoyment. PhD Sanna Koutaniemi, I wish to thank you from my heart. During these years you have always been there for me when I have had problems to solve. You feel like a supervisor to me. It has been a great pleasure to share a room with you. I will never get a friendlier and more cheerful colleague than you have been. I would also like to warmly thank Marja Huovila, the best technician I can ever imagine. Without your open-minded attitude towards molecular biology and your hard work at the lab, this thesis would not exist.

I am thankful to my follow-up group members Professor Ykä Helariutta and Docent Günter Brader for giving valuable scientific advice and guidance. Professor Sirpa Kärenlampi and Docent Günter Brader are acknowledged for kindly reviewing this thesis. Special thanks are due to the co-authors Leena Pohjala, PhD Roosa Laitinen, PhD Adyary Fallarero, Docent Päivi Tammela, PhD Panu Somervuo, Professor Pia Vuorela, Malena Sandberg, Suvi Broholm and Professor Paula Elomaa for their contribution to the articles.

All the former and present members of the Gerbera laboratory as well as the staff at the Department of Applied Biology are acknowledged for creating a pleasant working atmosphere. I want to express my sincere gratitude to Professor Paula Elomaa and Docent Mika Kotilainen for practical and theoretical advice and sharing ideas. Special thanks to my roommate Tino for many interesting conversations. The gardeners Sanna Peltola and Paula Junikka as well as Kaj-Roger Hurme and Arto Nieminen from Instrument Centre are thanked for generous help.

My friends PhD Tarja Malm and PhD Shea Beasley are thanked for long discussions and understanding during the desperate moments of this project. All my friends are warmly thanked for friendship and all the good moments.

My sister-in-law Anja, parents Marjatta and Reijo, my brother Sami and Lotta are thanked for their love and support. Finally, I owe my deepest gratitude to my husband Antti and my rascal Jesse. Antti, you have pushed me into finishing this project. Jesse, you bring endless joy into my life and make my life perfect!

7 REFERENCES

Abe I, Oguro S, Utsumi Y, Sano Y, Noguchi H (2005a) Engineered biosynthesis of plant polyketides: Chain length control in an octaketide-producing plant type III polyketide synthase. J Am Chem Soc 127:12709-12716

Abe I, Takahashi Y, Morita H, Noguchi H (2001) Benzalacetone synthase: A novel polyketide synthase that plays a crucial role in the biosynthesis of phenylbutanones in *Rheum palmatum*. Eur J Biochem 268:3354-3359

Abe I, Utsumi Y, Oguro S, Morita H, Sano Y, Noguchi H (2005b) A plant type III polyketide synthase that produces pentaketide chromone. J Am Chem Soc 127:1362-1363

Abe I, Utsumi Y, Oguro S, Noguchi H (2004) The first plant type III polyketide synthase that catalyzes formation of aromatic heptaketide. FEBS Letters 562:171-176

Abe I, Watanabe T, Lou W, Noguchi H (2006) Active site residues governing substrate selectivity and polyketide chain length in aloesone synthase. FEBS J 273:208-218

Akashi T, Aoki T, Ayabe S (1999) Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. Plant Physiol 121:821-828

Akashi T, Koshimizu S, Aoki T, Ayabe S (2006) Identification of cDNAs encoding pterocarpan reductase involved in isoflavan phytoalexin biosynthesis in *Lotus japonicus* by EST mining. FEBS Letters 580:5666-5670

Austin MB, Noel JP (2003) The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep 20:79-110

Ayabe S, Udagawa A, Furuya T (1988) NAD(P)H-dependent 6-deoxychalcone synthase activity in *Glycyrrhiza echinata* cells induced by yeast extract. Arch Biochem Biophys 261:458-462

Baker JM, Hawkins ND, Ward JL, Lovegrove A, Napier JA, Shewry PR, Beale MH (2006) A metabolomic study of substantial equivalence of field-grown genetically modified wheat. Plant Biotechnol J 4:381-392

Bartley GE, Scolnik PA (1995) Plant carotenoids: Pigments for photoprotection, visual attraction, and human health. Plant Cell 7:1027-1038

Batista R, Saibo N, Lourenco T, Oliveira MM (2008) Microarray analyses reveal that plant mutagenesis may induce more transcriptomic changes than transgene insertion. Proc Natl Acad Sci USA 105:3640-3645

Baudo MM, Lyons R, Powers S, Pastori GM, Edwards KJ, Holdsworth MJ, Shewry PR (2006) Transgenesis has less impact on the transcriptome of wheat grain than conventional breeding. Plant Biotech J 4:369-380

Baumert A, Maier W, Groger D, Deutzmann R (1994) Purification and properties of acridone synthase from cell suspension cultures of *Ruta graveolens.* L Z Naturforsch 49:26-32

Bevan M, Barnes WM, Chilton MD (1983) Structure and transcription of the nopaline synthase gene region of T-DNA. Nucleic Acids Res 11:369-385

Bohm AB. 1998. Introduction to flavonoids. Singapore: Harwood academic publishers

Bomati E, Austin M, Bowman M, Dixon R, Noel J (2005) Structural elucidation of chalcone reductase and implications for deoxychalcone biosynthesis. J Biol Chem 280:30496-30503

Borejsza-Wysocki W, Hrazdina G (1996) Aromatic polyketide synthases: purification, characterization and antibody development to benzalacetone synthase from

raspberry fruits. Plant Physiol 110:791-799

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

Brake DG, Thaler R, Evenson DP (2004) Evaluation of bt (*Bacillus thuringiensis*) corn on mouse testicular development by dual parameter flow cytometry. J Agric Food Chem 52:2097-2102

Britsch L, Heller W, Grisebach H (1981) Conversion of flavanone to flavone, dihydroflavonol and flavonol with an enzyme system from cell cultures of parley. Z Naturforsch 36c:742-750

Broun P, Somerville C (2001) Progress in plant metabolic engineering. Proc Natl Acad Sci USA 98:8925-8927

Burbulis IE, Winkel-Shirley B (1999) Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. Proc Natl Acad Sci USA 96:12929-12934

Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real-time RT-PCR - a perspective. J Mol Endocrinol 34:597-601

Bustin S (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. J Mol Endocrinol 29:23-39

Capell T, Christou P (2004) Progress in plant metabolic engineering. Curr Op Biotech 15:148-154

Catchpole GS, Beckmann M, Enot DP, Mondhe M, Zywicki B, Taylor J, Hardy N, Smith A, King RD, Kell DB, Fiehn O, Draper J (2005) Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. Proc Natl Acad Sci USA 102:14458-14462

Chappell J, Nable R (1987) Induction of sesquiterpenoid biosynthesis in tobacco cell suspension cultures by fungal elicitor. Plant Physiol 85:469-473

Chen ZL, Gu H, Li Y, Su Y, Wu P, Jiang Z, Ming X, Tian J, Pan N, Qu LJ (2003) Safety assessment for genetically modified sweet pepper and tomato. Toxicol 188:297-307

Cheng AX, Lou YG, Mao YB, Lu S, Wang LJ, Chen XY (2007) Plant terpenoids: Biosynthesis and ecological functions. J Integr Plant Biol 49:179-186

Clemedson C, Barile FA, Chesne C, Cottin M, Curren R, Ekwall B, Ferro M, Gomez-lechon MJ, Imai K, Janus J, Kemp RB, Kemp RB, Kerszman G, Kjellstrand P, Lavrijsen K, Logemann P, McFarlane-abdula E, Roguet R, Segner H, Thuvander A, Walum E, Ekwall B (2000) MEIC evaluation of acute systemic toxicity: Part VII: Prediction of human toxicity by results from testing the first 30 reference chemicals with 27 further in vitro assays. ATLA 28:161-200

Corpillo D, Gardini G, Vaira AM, Basso M, Aime S, Accotto GP, Fasano M (2004) Proteomics as a tool to improve investigation of substantial equivalence in genetically modified organisms: The case of a virus-resistant tomato. Proteomics 4:193-200

Croteau R, Kutchan TM, Lewis NG. 2000. Natural products (Secondary metabolites). In: Buchanan BB, Gruissem W, Jones RL, editors. Biochemistry & molecular biology of plants. USA: Courier companies, Inc. p. 1250-1318

Dana C, Bevan D, Winkel B (2006) Molecular modeling of the effects of mutant alleles on chalcone synthase protein structure. J Mol Model 12:905-914

Defernez M, Gunning YM, Parr AJ, Shepherd LV, Davies HV, Colquhoun IJ (2004) NMR and HPLC-UV profiling of potatoes with genetic modifications to metabolic pathways. J Agric Food Chem 52:6075-6085

Deluc L, Barrieu F, Marchive C, Lauvergeat V, Decendit A, Richard T, Carde J, Merillon J, Hamdi S (2006) Characterization of a grapevine R2R3-MYB transcription

factor that regulates the phenylpropanoid pathway. Plant Physiol 140:499-511

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

Eisenbrand G, Pool-Zobel B, Baker V, Balls M, Blaauboer BJ, Boobis A, Carere A, Kevekordes S, Lhuguenot JC, Pieters R, Kleiner J (2002) Methods of *in vitro* toxicology. Food Chem Toxicol 40:193-236

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. Mol Breed 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. Bio/technology 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). Plant J 16:93-99

Elomaa P, Teeri TH. 2001. Transgenic *Gerbera*. In: Bajaj YPS, editor. Biotechnology in Agriculture and Forestry. Berlin Heidelberg: Springer-Verlag. p. 139-154

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

Ferrer J, Jez JM, Bowman ME, Dixon RA, Noel JP (1999) Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat Struct Biol 6:775-784

Flachowsky G, Chesson A, Aulrich K (2005) Animal nutrition with feeds from genetically modified plants. Arch Anim Nutr 59:1-40

Forkmann G, Heller W, Grisebach H (1980) Anthocyanin biosynthesis in flowers of *Matthiola incana*: Flavanone 3- and flavonoid 3'-hydroxylase. Z Naturforsch 35c:691-695

Fukuchi-Mizutani M, Okuhara H, Fukui Y, Nakao M, Katsumoto Y, Yonekura-Sakakibara K, Kusumi T, Hase T, Tanaka Y (2003) Biochemical and molecular characterization of a novel UDP-glucose:Anthocyanin 3'-O-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. Plant Physiol 132:1652-1663

Gabriac B, Werck-Reichhart D, Teutsch H, Durst F (1991) Purification and immunocharacterization of a plant cytochrome P450: The cinnamic acid 4-hydroxylase. Arch Biochem Biophys 288:302-309

Gershenzon J, Maffei M, Croteau R (1989) Biochemical and histochemical localization of monoterpene biosynthesis in the glandular trichomes of spearmint (*Mentha spicata*). Plant Physiol 89:1351-1357

Goodwin TW (1979) Biosynthesis of terpenoids. Annu Rev Plant Physiol 30:369-404

Gould KS, Lister C. 2006. Flavonoid functions in plants. In: Andersen OM, Markham KR, editors. Flavonoids: chemistry, biochemistry and applications. CRC Press. p. 397-442

Graham T (1991) Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates. Plant Physiol 95:594-603

Graham TL, Graham MY (1996) Signaling in soybean

phenylpropanoid responses (dissection of primary, secondary, and conditioning effects of light, wounding, and elicitor treatments). Plant Physiol 110:1123-1133

Gregersen P, Brinch-Pedersen H, Holm P (2005) A microarray-based comparative analysis of gene expression profiles during grain development in transgenic and wild type wheat. Transgenic Res 14:887-905

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. Proc Natl Acad Sci USA 101:2209-2214

Hammond B, Lemen J, Dudek R, Ward D, Jiang C, Nemeth M, Burns J (2006) Results of a 90-day safety assurance study with rats fed grain from corn rootworm-protected corn. Food Chem Toxicol 44:147-160

Helariutta Y, Elomaa P, Kotilainen M, Griesbach RJ, Schröder 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 Mol Biol 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 Mol Biol 22:183-193

Helariutta Y, Kotilainen M, Elomaa P, Kalkkinen N, Bremer K, Teeri TH, Albert VA (1996) Duplication and functional divergence in the chalcone synthase gene family of asteraceae: Evolution with substrate change and catalytic simplification. Proc Natl Acad Sci USA 93:9033-9038

Heller W, Hahlbrock K (1980) Highly purified "flavanone synthase" from parsley catalyzes the formation of naringenin chalcone. Arch Biochem Biophys 200:617-619

Herrera-Estrella L, Depicker A, Van Montagu M, Schell J (1983) Expression of chimaeric genes transferred into plant cells using a ti-plasmid-derived vector. Nature 303:209-213

Hirotani M, Kuroda R, Suzuki H, Yoshizawa T (2000) Cloning and expression of UDP-glucose:Flavonoid 7-*O*-glucosyltransferase from hairy root cultures of *Scutellaria baicalensis*. Planta 210:1006-1013

Hooper L, Cassidy A (2006) A review of the health care potential of bioactive compounds. J Sci Food Agric 86:1805-1813

Hopwood DA, Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu Rev Genet 24:37-66

Ishiyama M, Tominaga H, Shiga M, Sasamoto K, Ohkura Y, Ueno K, Watanabe M (1995) Novel cell proliferation and cytotoxicity assays using a tetrazolium salt that produces a water-soluble formazan dye. In Vitro Toxicol 8:187-190

James C (2007) Global status of commercialized biotech/GM crops. ISAAA Briefs 37

Jez JM, Austin MB, Ferrer J, Bowman ME, Schröder J, Noel JP (2000b) Structural control of polyketide formation in plant-specific polyketide synthases. Chemistry & Biology 7:919-930

Jez JM, Bowman ME, Dixon RA, Noel JP (2000a) Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. Nat Struct Biol 7:786-791

Jung W, Yu O, Lau SC, O'Keefe DP, Odell J, Fader G, McGonigle B (2000) Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. Nature Biotech 18:208-212

Katsumoto Y, Fukuchi-Mizutani M, Fukui Y, Brugliera F, Holton T, Karan M, Nakamura N, Yonekura-

Sakakibara K, Togami J, Pigeaire A, Tao G-Q, 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 Cell Physiol DOI 10.1093/pcp/pcm131

Kleinig H (1989) The role of plastids in isoprenoid biosynthesis. Annu Rev Plant Physiol Plant Mol Biol 40:39-59

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

Kosslak R, Bookland R, Barkei J, Paaren H, Appelbaum E (1987) Induction of *Bradyrhizobium japonicum* common nod genes by isoflavones isolated from *Glycine max*. Proc Natl Acad Sci USA 84:7428-7432

Kotilainen M, Elomaa P, Uimari A, Albert VA, Yu D, Teeri TH (2000) GRCD1, an AGL2-like MADS box gene, participates in the C function during stamen development in *Gerbera hybrida*. Plant Cell 12:1893-1902

Kotilainen M, Helariutta Y, Mehto M, Pöllänen E, Albert VA, Elomaa P, Teeri TH (1999) GEG participates in the regulation of cell and organ shape during corolla and carpel development in *Gerbera hybrida*. Plant Cell 11:1093-1104

Koukol J, Conn EE (1961) The metabolism of aromatic compounds in higher plants. IV. purification and properties of the phenylalanine deaminase of *Hordeum vulgare*. J Biol Chem 236:2692-2698

Kreuzaler F, Hahlbrock K (1972) Enzymatic synthesis of aromatic compounds in higher plants: Formation of naringenin (5,7,4'-trihydroxyflavanone) from p-coumaroyl coenzyme A and malonyl coenzyme A. FEBS Lett 28:69-72

Kristensen C, Morant M, Olsen CE, Ekstrom CT, Galbraith DW, Moller BL, Bak S (2005) Metabolic engineering of dhurrin in transgenic *Arabidopsis* plants with marginal inadvertent effects on the metabolome and transcriptome. Proc Natl Acad Sci USA 102:1779-1784

Kristiansen KN, Rohde W (1991) Structure of the *hordeum vulgare* gene encoding dihydroflavonol-4-reductase and molecular analysis of ant18 mutants blocked in flavonoid synthesis. Mol Gen Genet 230:49-59

Kuiper HA, Kleter GA, Noteborn HPJM, Kok EJ (2001) Assessment of the food safety issues related to genetically modified foods. Plant J 27:503-528

Laitinen RAE, Immanen J, Auvinen P, Rudd S, Alatalo E, Paulin L, Ainasoja M, Kotilainen M, Koskela S, Teeri TH, Elomaa P (2005) Analysis of the floral transcriptome uncovers new regulators of organ determination and gene families related to flower organ differentiation in *Gerbera hybrida* (asteraceae). Genome Res 15:475-486

Lanz T, Tropf S, Marner F, Schröder J, Schröder G (1991) The role of cysteines in polyketide synthases. J Biol Chem 266:9971-9976

Last RL, Jones AD, Shachar-Hill Y (2007) Towards the plant metabolome and beyond. Nat Rev Mol Cell Biol 8:167-174

Le Gall G, Colquhoun IJ, Davis AL, Collins GJ, Verhoeyen ME (2003) Metabolite profiling of tomato (*Lycopersicon esculentum*) using 1H NMR spectroscopy as a tool to detect potential unintended effects following a genetic modification. J Agric Food Chem 51:2447-2456

Lehesranta SJ, Davies HV, Shepherd LVT, Nunan N, McNicol JW, Auriola S, Koistinen KM, Suomalainen S, Kokko HI, Kärenlampi SO (2005) Comparison of tuber proteomes of potato varieties, landraces, and genetically modified lines. Plant Physiol 138:1690-1699

Lewinsohn E, Gijzen M, Croteau R (1991) Defense mechanisms of conifers: Differences in constitutive

and wound-induced monoterpene biosynthesis among species. Plant Physiol 96:44-49

Lewinsohn E, Gijzen M, Savage TJ, Croteau R (1991) Defense mechanisms of conifers : Relationship of monoterpene cyclase activity to anatomical specialization and oleoresin monoterpene content. Plant Physiol 96:38-43

Lichtenthaler HK, Rohmer M, Schwender J (1997) Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. Physiol Plantarum 101:643-652

Liu B, Raeth T, Beuerle T, Beerhues L (2007) Biphenyl synthase, a novel type III polyketide synthase. Planta 225:1495-1503

Lütke-Brinkhaus F, Liedvogel B, Kleinig H (1984) On the biosynthesis of ubiquinones in plant mitochondria. Eur J Biochem 141:537-541

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

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

Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nat Biotechnol 21:796-802

Mathews H, Clendennen S, Caldwell C, Liu X, Connors K, Matheis N, Schuster D, Menasco DJ, Wagoner W, Lightner J, Wagner DR (2003) Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport. Plant Cell 15:1689-1703

McGarvey DJ, Croteau R (1995) Terpenoid metabolism. Plant Cell 7:1015-1026

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

Moustafa E, Wong E (1967) Purification and properties of chalcone-flavanone isomerase from soy bean seed. Phytochemistry 6:625-632

Mumm R, Posthumus MA, Dicke M (2008) Significance of terpenoids in induced indirect plant defence against herbivorous arthropods. Plant Cell Environ 31:575-585

Nakajima J, Tanaka Y, Yamazaki M, Saito K (2001) Reaction mechanism from leucoanthocyanidin to anthocyanidin 3-glucoside, a key reaction for coloring in anthocyanin biosynthesis. J Biol Chem 276:25797-25803

Nakayama T, Yonekura-Sakakibara K, Sato T, Kikuchi S, Fukui Y, Fukuchi-Mizutani M, Ueda T, Nakao M, Tanaka Y, Kusumi T, Tanaka Y (2000) Aureusidin synthase: A polyphenol oxidase homolog responsible for flower coloration. Science 290:1163-1166

Ono E, Fukuchi-Mizutani M, Nakamura N, Fukui Y, Yonekura-Sakakibara K, Yamaguchi M, Nakayama T, Tanaka T, Kusumi T, Tanaka Y (2006) Yellow flowers generated by expression of the aurone biosynthetic pathway. Proc Natl Acad Sci USA 103:11075-11080

Pichersky E, Raguso RA, Lewinsohn E, Croteau R (1994) Floral scent production in *Clarkia* (onagraceae) (I. localization and developmental modulation of monoterpene emission and linalool synthase activity). Plant Physiol 106:1533-1540

Pohjala L, Tammela P, Samanta SK, Yli-Kauhaluoma J, Vuorela P (2007) Assessing the data quality in predictive toxicology using a panel of cell lines and cytotoxicity assays. Anal Biochem 362:221-228

Pourcel L, Routaboul J, Kerhoas L, Caboche M, Lepiniec L, Debeaujon I (2005) *TRANSPARENT TESTA10* encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in *Arabidopsis* seed

coat. Plant Cell DOI 10.1105/tpc.105.035154

R Development Core Team. 2005. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing

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

Reimold U, Kröger M, Kreuzaler F, Hahlbrock K (1983) Coding and 3' non-coding nucleotide sequence of chalcone synthase mRNA and assignment of amino acid sequence of the enzyme. EMBO J 2:1801-1805

Ruebelt MC, Lipp M, Reynolds TL, Schmuke JJ, Astwood JD, DellaPenna D, Engel K, Jany K (2006) Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 3. assessing unintended effects. J Agric Food Chem 54:2169-2177

Saito K, Kobayashi M, Gong Z, Tanaka, Y. and 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*. Plant J 17:181-189

Saito K, Yamazaki M (2002) Biochemistry and molecular biology of the late-stage of biosynthesis of anthocyanin: Lessons from *Perilla frutescens* as a model plant. New Phytol 155:9-23

Sanhoty RE, El-Rahman AAA, Bögl KW (2004) Quality and safety evaluation of genetically modified potatoes spunta with *cry V* gene: Compositional analysis, determination of some toxins, antinutrients compounds and feeding study in rats. Nahrung/Food 48:13-18

Schena M, Shalon D, Davis R, Brown P (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467-470

Schröder J (1997) A family of plant-specific polyketide synthases: Facts and predictions. Trends Plant Sci 2:373-378

Seigler DS. 1998. Plant secondary metabolism. Dordrecht, The Netherlands: Kluwer Academic Publishers

Shepherd LV, McNicol JW, Razzo R, Taylor MA, Davies HV (2006) Assessing the potential for unintended effects in genetically modified potatoes perturbed in metabolic and developmental processes. Targeted analysis of key nutrients and anti-nutrients. Transgenic Res 15:409-425

Shimada N, Sasaki R, Sato S, Kaneko T, Tabata S, Aoki T, Ayabe S (2005) A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. J Exp Bot 56:2573-2585

Slater K (2001) Cytotoxicity tests for high-throughput drug discovery. Curr Opin Biotechnol 12:70-74

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

Stafford H, Lester H (1984) Flavan-3-ol biosynthesis: The conversion of (+)-dihydroquercetin and flavan-3,4-cis-diol (leucocyanidin) to (+)-catechin by reductases extracted from cell suspension cultures of douglas fir. Plant Physiol 76:184-186

Stafford H, Lester H (1982) Enzymic and nonenzymic reduction of (+)-dihydroquercetin to its 3,4,-diol. Plant Physiol 70:695-698

Steele CL, Gijzen M, Qutob D, Dixon RA (1999) Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. Arch Biochem Biophys 367:146-150

Stotz G, Forkmann G (1981) Hydroxylation of flavanones to flavones with flower extracts of

Antirrhinum majus (snapdragon). Z Naturforsch 36c:737-741

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

Tanner G, Francki K, Abrahams S, Watson J, Larkin P, Ashton A (2003) Proanthocyanidin biosynthesis in plants: Purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. J Biol Chem 278:31647-31656

Teeri TH, Kotilainen M, Uimari A, Ruokolainen S, Ng YP, Malm U, Pöllänen E, Broholm, S., Laitinen, R.A.E., Elomaa P, Albert VA (2006) Floral developmental genetics of *Gerbera* (asteraceae). ABR 44:323-351

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

Tricarico C, Pinzani P, Bianchi S, Paglierani M, Distante V, Pazzagli M, Bustin SA, Orlando C (2002) Quantitative real-time reverse transcription polymerase chain reaction: Normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. Anal Biochem 309:293-300

Uimari A, Kotilainen M, Elomaa P, Yu D, Albert VA, Teeri TH (2004) Integration of reproductive meristem fates by a SEPALLATA-like MADS-box gene. Proc Natl Acad Sci USA 101:15817-15822

Vain P (2007) Trends in GM crop, food and feed safety literature. Nature Biotech 25:624-626

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3:research0034.1-research0034.11

Verpoorte R, Alfermann AW, editors. 2000. Metabolic engineering of plant secondary metabolism. Dordrecht, The Netherlands: Kluwer Academic Publishers

Welle R, Grisebach H (1989) Phytoalexin synthesis in soybean cells: Elicitor induction of reductase involved in biosynthesis of 6'-deoxychalcone. Arch Biochem Biophys 272:97-102

Wilson AP. 2000. Cytotoxicity and viability assays. In: Masters J, editor. Animal cell culture. A practical approach. New York: Oxford University Press. p. 175-218

Wink M, editor. 1999. Biochemistry of plant secondary metabolism annual plant reviews. Sheffield, England: Sheffield Academic Press Ltd

Xie D, Sharma S, Paiva N, Ferreira D, Dixon R (2003) Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. Science 299:396-399

Yamazaki M, Gong Z, Fukuchi-Mizutani M, Fukui Y, Tanaka Y, Kusumi T, Saito K (1999) Molecular cloning and biochemical characterization of a novel anthocyanin 5-O-glucosyltransferase by mRNA differential display for plant forms regarding anthocyanin. J Biol Chem 274:7405-7411

Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, Beyer P, Potrykus I (2000) Engineering the provitamin A (-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 287:303-305

Yu D, Kotilainen M, Pöllänen E, Mehto M, Elomaa P, Helariutta Y, Albert VA, Teeri TH (1999) Organ identity genes and modified patterns of flower

development in *Gerbera hybrida* (asteraceae). Plant J 17:51-62

Zhang JH, Chun TDY, Oldenburg KR (1999) A simple statistical parameter for use in evaluation and validation of high-throughput screening assays. J Biomol Screen 4:67-73