

**PROTEIN-PROTEIN INTERACTIONS OF THE
FLAVONOID BIOSYNTHETIC ENZYMES IN
*ARABIDOPSIS THALIANA***

Jiale Xiang

Master's Thesis

University of Helsinki

Department of Agricultural Sciences

MScPPS / Horticulture

September 2014

Tiedekunta/Osasto - Fakultet/Sektion - Faculty Faculty of Agriculture and Forestry		Laitos – Institution - Department Department of Agricultural Sciences	
Tekijä - Författare - Author Jiale Xiang			
Työn nimi - Arbetets titel - Title Protein – protein interaction of the flavonoid biosynthetic enzymes in <i>Arabidopsis thaliana</i>			
Oppiaine - Läroämne - Subject Plant Production Science (MScPPS) / Horticulture			
Työn laji - Arbetets art - Level Master´s thesis	Aika - Datum - Month and year September 2014	Sivumäärä - Sidoantal - Number of pages 70	
Tiivistelmä - Referat – Abstract <p>Flavonoids are a group of secondary metabolites, which are not only important for plants' survival, but also have been found to have medicinal properties for human health. Several enzymes are involved in the flavonoid biosynthesis. It is thought that these enzymes work together and may form enzymatic complexes. But the way of these enzymes interact with each other is still not clear. In arabidopsis, the number of gene family members that encode these enzymes is less than in other model plants, which makes it as a suitable model to investigate the interactions of enzymes involved in the flavonoid biosynthetic pathway.</p> <p>In this study, ten full-length flavonoid pathway genes were successfully amplified from cDNA of the arabidopsis flower. They are <i>PAL1</i>, <i>C4H</i>, <i>CHS</i>, <i>CHI</i>, <i>F3H</i>, <i>F3'H</i>, <i>DFR</i>, <i>FLS1</i>, <i>ANS</i> and <i>GT</i>. These genes were cloned into different prey vectors (pPR3-N and pPR3-SUC) and bait vectors (pDHB1 and pBT3-SUC). After that, the constructs were transformed separately into yeast. The protein-protein interactions were analyzed via yeast two-hybrid system.</p>			
Avainsanat - Nyckelord - Keywords <i>Arabidopsis thaliana</i> , flavonoid, protein-protein interaction, yeast two-hybrid			
Säilytyspaikka - Förvaringsställe - Where deposited Viikki Science Library & Department of Agricultural Sciences			
Muita tietoja - Övriga uppgifter - Further information Supervised by Prof. Teemu Teeri, Dr. Milla Pietiäinen and doctoral student Hany El Sayed			

Table of Contents

ABBREVIATIONS	5
1. INTRODUCTION	6
2. LITERATURE REVIEW	8
2.1 Flavonoids	8
2.1.1 Functions of flavonoids	8
2.1.2 Anthocyanins.....	9
2.2 Flavonoid biosynthetic pathway	11
2.2.1 Flavonoid biosynthetic pathway in arabidopsis including the phenylpropanoid pathway	11
2.2.2 Branches of flavonoid biosynthetic pathway in other plants	15
2.3 Yeast two-hybrid systems	15
2.3.1 Molecular mechanism of the GAL4 yeast two-hybrid system..	15
2.3.2 Development and modification of the yeast two-hybrid system	16
2.4 Metabolons in plants	18
3. OBJECTIVES	20
4. MATERIALS AND METHODS	21
4.1 Plant material and growth conditions.....	21
4.2 DUAL hunter system and DUAL membrane pairwise system	21
4.3 The cloning vectors and yeast strains for bait and prey	21
4.4 Primer design.....	24
4.5 RNA extraction.....	24
4.6 cDNA synthesis	25
4.7 Polymerase chain reaction and electrophoresis	25
4.8 Gene isolation.....	26
4.9 Construct confirmation.....	27
4.10 Plasmid DNA extractions	28
4.11 Bait testing	28
4.12 Yeast two-hybrid assay	29

4.13 β -galactosidase assay	30
5. RESULTS.....	30
5.1 RNA extraction and cDNA synthesis	30
5.2 PCR reactions and electrophoresis	31
5.3 Construction of the bait and prey vectors.....	32
5.4 <i>Sfi</i> I digestion	33
5.5 The gene sequencing.....	34
5.6 Bait testing.....	37
5.7 Yeast two-hybrid assay.....	39
5.9 β -galactosidase assay	44
6. DISCUSSION	46
6.1 Gene isolation.....	46
6.2 Bait testing showed a weak self-activation of bait vectors.....	47
6.3 The enzymatic interactions based on Y2H assay	47
6.4 Improvements of yeast two-hybrid assay	49
6.4.1 Sources of false positives in yeast two-hybrid.....	49
6.4.2 Approaches to improve the reliability of yeast two-hybrid system	49
6.5 β -galactosidase assay confirms yeast two-hybrid results	50
7. CONCLUSIONS.....	52
8. ACKNOWLEDGMENTS.....	53
9. REFERENCES.....	54
10. APPENDICES	65

ABBREVIATIONS

PAL	Phenylalanine ammonia-lyase
C4H	cinnamate 4-hydroxylase
4CL	4-coumaroyl ligase
CHS	chalcone synthase
CHI	chalcone isomerase
F3H	flavanone 3-hydroxylase
F3'H	flavonoid3'-hydroxylase
DFR	dihydroflavonol 4-reductase
FLS	flavonol synthase
ANS	anthocyanidin synthase
GT	glucosyltransferase
DHK	dihydrokaempferol
DHQ	dihydroquercetin
Y2H	yeast two-hybrid
trp	tryptophan
leu	leucine
his	histidine
ade	adenine
DB	DNA-binding domain
AD	activation domain
3-AT	3-Amino-1,2,4-triazole
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

1. INTRODUCTION

Flavonoids are a class of secondary metabolites that are ubiquitously distributed in the plant kingdom. They protect plants from UV damage (Stapleton and Walbot 1994), act as negative regulators of auxin transport in vivo in *arabidopsis* (Brown *et al.* 2001), improve plants' resistance to biotic and abiotic stresses (reviewed by Treutter 2006), regulate the levels of reactive oxygen species (Taylor and Grotewold 2005) and attract pollinators (Iwashina 2003).

About ten enzymes are specifically involved in the flavonoid biosynthetic pathway (Figure 1). The corresponding genes of these enzymes are either single or members of small gene families, depending on plant species. For example, in *gerbera hybrida*, three family members of CHS were identified as (Helariutta *et al.* 1995, Deng *et al.* 2014). In *arabidopsis*, most of the enzymes that are involved in the flavonoid biosynthetic pathway are encoded by a single gene, except for *PAL* (Ohl *et al.* 1990), *FLS* (Owens *et al.* 2008a) and *4CL* (Ehltling *et al.* 1999, Hamberger & Hahlbrock 2004).

In the flavonoid biosynthetic pathway, enzymes are believed to work together, forming enzymatic complexes (metabolons) (Ovadi & Srere 2000). But there are not so many studies that directly reveal the interactions of these enzymes. The work from Burbulis and Shirley (1999) showed protein interactions between CHS and DFR, CHI and CHS, and DFR and CHI. Moreover, FLS1 displayed interactions with F3H and DFR in both orientations whereas it interacted with CHS only when fused to the activation domain in the yeast two-hybrid system (Owens *et al.* 2008a). But we still know very little about the interactions of other enzymes of the flavonoid biosynthetic pathway. Due to the simple gene family number and the published genome sequences, *arabidopsis* is a very good model to study how these enzymes work together.

In my study, I tried to explore the enzymatic interactions in the flavonoid

biosynthetic pathway of arabidopsis that could give evidence to enzymatic complexes (metabolons).

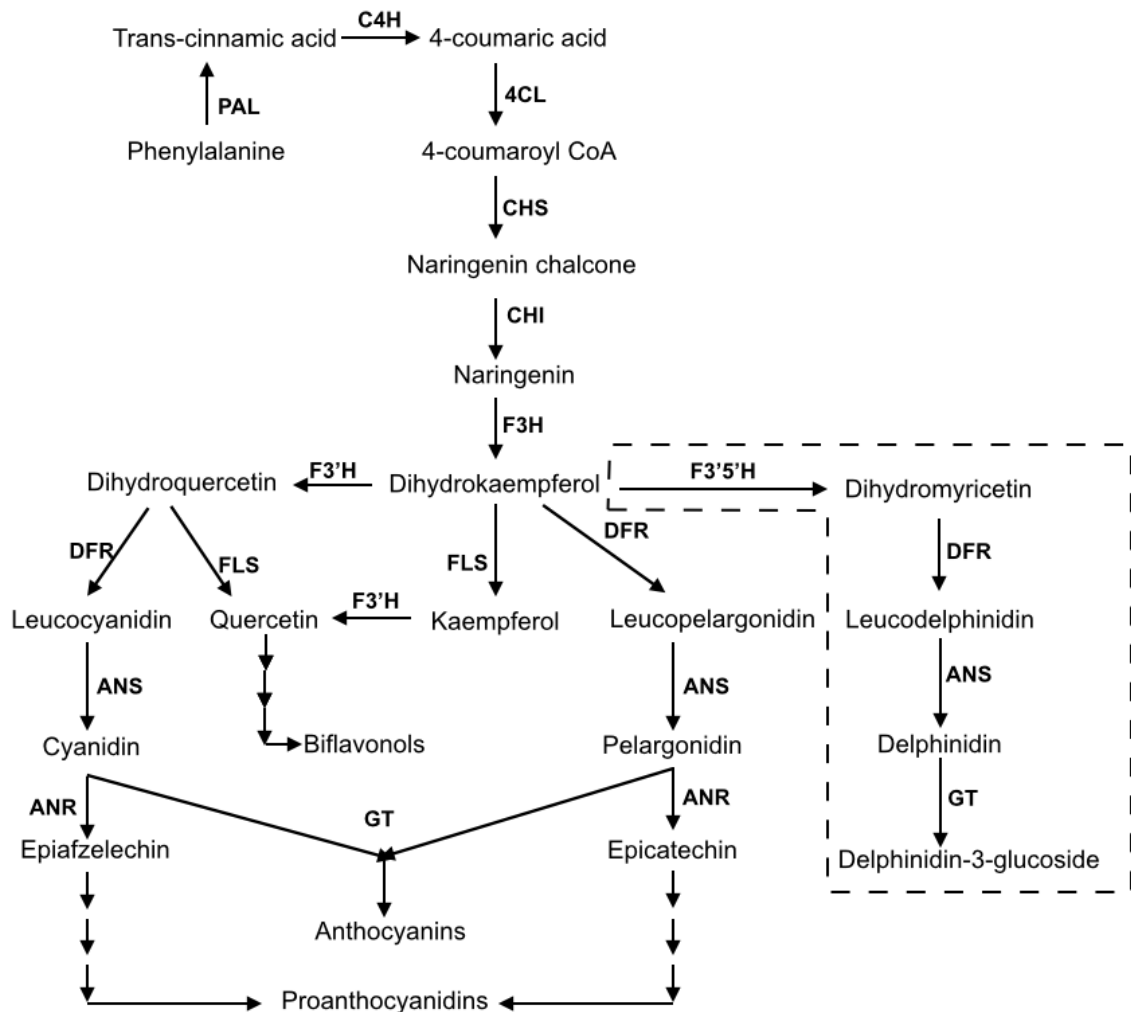


Figure 1. The schematic map of flavonoid biosynthetic pathway. PAL: Phenylalanine ammonia-lyase; C4H: cinnamate 4-hydroxylase; 4CL: 4-coumaroyl CoA-ligase; CHS: chalcone synthase; CHI: chalcone isomerase; F3H: flavonol 3-hydroxylase; F3'H: flavonoid 3'-hydroxylase; DFR: dihydroflavonol reductase; FLS: flavonol synthase; ANS: anthocyanidin synthase; GT: glycosyltransferase; ANR: anthocyanidin reductase; F3'5'H: Flavonoid-3',5'-hydroxylase. The branch via F3'5'H to delphinidin does not exist in Arabidopsis.

2. LITERATURE REVIEW

2.1 Flavonoids

Secondary metabolites are a group of chemical compounds that are not necessary for plants' living, but closely relate to their interactions with surrounding environments. Terpenoids, polyketides, phenylpropanoids and alkaloids are four important sub-groups of secondary metabolites in plants. Phenylpropanoids consist of coumarins, lignin, lignans, tannins and flavonoids. Flavonoids generally have a fifteen-carbon skeleton, consisting of two phenyl rings (A-ring and B-ring) that are connected by a three-carbon bridge (C-ring) (Iwashina 2000). Flavonoids are categorized into six most important groups: chalcones, flavones, flavonols, flavandiols, anthocyanins and proanthocyanidins (condensed tannins) (Shirley 2001). Some species are capable of synthesizing specialized flavonoids, such as isoflavonoids in legumes (Aoki *et al.* 2000). The carbon skeleton of flavonoids is decorated by glycosylation, acylation and methylation, leading to a large number of different molecules. Arabidopsis has at least 54 flavonoid molecules including 35 flavonols, 11 anthocyanins and eight proanthocyanidins (Saito *et al.* 2013).

2.1.1 Functions of flavonoids

Flavonoids play important and indispensable roles for plants' self-protection. The function of serving as a filter to prevent the DNA damage from UV-B has been identified both *in vitro* (Kootstra 1994) and *in vivo* (Stapleton and Walbot 1994). The transparent testa-4 (*tt4*, *CHS*) mutant, *tt5* (*CHI*) mutant and *tt6* (*F3H*) mutant of arabidopsis showed reduced flavonoid levels, and these mutants were highly sensitive to UV-B irradiance and damaging effects of UV-B radiation, respectively (Li *et al.* 1993). Flavonoids acting as a UV-B protective chemical were not only identified in arabidopsis, but also in other species, such as turnip (*Brassica naps*) (Olsson *et al.* 1998), cabbage (*Brassica oleracea*) (Gitz *et al.* 1998), barley (*Hordeum vulgare*) (Reuber *et al.* 1996), etc., which

were all well reviewed by Harborne and Williams (2000). Furthermore, flavonoids are one of the determinants of flower color. Carotenoids are major pigments in many yellow flowers but anthocyanin pigments contribute to most orange, red, purple and blue flowers.

Flavonoids may protect plants from microbial invasion. Resistant accessions of chickpea (*Cicer arietinum*) have a much higher maackiain (an isoflavonoid) content compared to susceptible accessions and the maackiain concentration increased hugely after plants were inoculated with the *Botrytis* grey mould (BGM). *In vitro*, the germination of BGM spores was inhibited (Stevenson and Haware 1999).

Flavonoids in plants possess medicinal properties. Knekt revealed that people with higher quercetin and kaempferol intake had a lower mortality from ischemic heart disease (Knekt *et al.* 2002). Another experiment found out that quercetin treatment had a protective effect in diabetes leading to a decreased oxidative stress and a preservation of pancreatic β -cell integrity due to their free radical scavenging properties (Coskun *et al.* 2005). Baicalin, a flavonoid isolated from the medicinal plant *Scutellaria baicalensis*, is a potential candidate in the prevention and treatment of periodontal diseases because their ability to block the matrix metalloproteinase-8 release from polymorph nuclear leukocytes (Zhu *et al.* 2007). Antiviral activity of flavonoids has been reported as well. For instance, eleven biflavones isolated from *Rhus succedanea* and *Garcinia multiflora* were tested for the anti-*HIV-1* reverse transcriptase (RT) activity, which are necessary for viral replication. Robustaflavone and hinokiflavone inhibited *HIV-1* RT. By contrast, amentoflavone, agathisflavone, morelloflabone and GB-1a achieved moderate effects against *HIV-1* RT (Lin *et al.* 1997).

2.1.2 Anthocyanins

Anthocyanins are a group of the most important flavonoids that result in visible colors in plants. Anthocyanins have many members including cyanidin, peonidin, delphinidin, petunidin, malvidin and pelargonidin. In terms of

commercial interests, the alternation of flower color, which creates new colorful flowers, is a potential market. In these processes, the knowledge of anthocyanin biosynthetic pathway is a key point. Maize (*Zea mays*), snapdragon (*Antirrhinum majus*) and petunia (*Petunia hybrida*) are particularly used as models in research of the anthocyanin biosynthetic pathway (Holton and Cornish 1995). In petunia, dihydroflavonol 4-reductase (DFR) does not catalyze reduction of dihydrokaempferol (DHK) to leucopelargonidin whereas it reduces dihydroquercetin (DHQ) to leucocyanidin and dihydromyricetin to leucodelphinidin, which eventually leads to the lack of orange (pelargonidin) (Forkmann and Ruhau 1987). Moreover, due to the absence of flavonoid 3'5'-hydroxylase (F3'5'H), the cultured rose (*Rose hybrida*) and gerbera (*Gerbera hybrida*) are unable of synthesizing blue delphinidin derived anthocyanin. However, the color modification was successfully achieved by genetic engineering, which produced delphinidin blue in carnations (Mol *et al.* 1999) and roses (Katsumoto *et al.* 2007). In gerbera, an alternation of flower pigment to cream white was achieved by the transfer of a full-length antisense of the chalcone synthase encoding cDNA in *Gerbera hybrida* (Elomaa *et al.* 1993).

Besides flower colors, anthocyanins are involved in other physiological processes in plants. Castellarin *et al.* (2007) showed that the anthocyanin contents in grapes are 37% to 57% higher when plants are suffering water deficit compared to control plants. Potato tubers with the ectopic expression of anthocyanin 5-O-glucosyltransferase display two times higher resistance to bacterial infections (Lorenc-Kukula *et al.* 2005). Anthocyanins also play important roles in cold stress (Christie *et al.* 1994) and have antioxidant activities (Li *et al.* 2012).

Anthocyanins have a lot of promising dietary applications. For example, anthocyanins can intake the insulin secretion of pancreatic β -cells, which suggests that fruits containing abundant anthocyanins or purified anthocyanins could be used for preventing type-2 diabetes (Jayaprokasam *et al.* 2005). Furthermore, as the result of expression of two transcription factors (*Del* and *Ros1*) from snapdragon, tomato has a higher anthocyanin level with purple

coloration in both peel and flesh. More interestingly, cancer-susceptible mice had a longer life span after fed with these high-anthocyanin tomatoes (Butelli *et al.* 2008). In addition, anthocyanins are promising natural colorants for food industry (Espín *et al.* 2000, He and Giusti 2010, Buchweitz *et al.* 2013).

2.2 Flavonoid biosynthetic pathway

2.2.1 Flavonoid biosynthetic pathway in arabidopsis including the phenylpropanoid pathway

Phenyl alanine is the primary metabolite of the phenylpropanoid pathway, which is catalyzed by phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H) and 4-coumaroyl ligase (4CL), leading to the production of 4-coumaroyl-CoA, the precursors of flavonoids (Figure 1), lignin and lignans. Malonyl CoA and 4-coumaroyl CoA (Figure 1) are the precursors of all flavonoids, which are catalyzed by following enzymes, including chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), flavonol synthase (FLS) and dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS) and glucosyltransferase (GT) (Holton & Cornish 1995). After several steps catalyzed by enzymes (Figure 1), they are finally transformed to anthocyanins, proanthocyanidins or other components via glycosylation, methylation and acylation.

PAL (phenylalanine ammonia-lyase) converses phenylalanine to *trans*-cinnamic acid by deamination. *PAL* was postulated to have a small gene family in arabidopsis (Ohl *et al.* 1990). Until now, four gene family members have been discovered and characterized (Wanner *et al.* 1995, Cochrane *et al.* 2004). *PAL1* and *PAL2* have similar structure and expression levels. *PAL3* always clusters with *PAL4* (Raes *et al.* 2003). However, it is still possible to have other undetected family members in arabidopsis as *pal1pal2pal3pal4* quadruple knockout mutants still contain about 10% of the wild-type PAL activity (Huang *et al.* 2010). It was revealed that *PAL1*, *PAL2* and *PAL4* relate with lignin synthesis which share the same first three steps with flavonoid biosynthetic

pathway (Raes *et al.* 2003, Rohde *et al.* 2004). In addition, *PAL1* and *PAL2* are involved in abiotic environmental-triggered flavonoid synthesis (Olsen *et al.* 2008) and the *pal1pal2* double mutant has reduced anthocyanin and proanthocyanidin production (Huang *et al.* 2010). Compared to *PAL3* with a very low expression level in arabidopsis, *PAL1*, *PAL2* and *PAL4* have high expression levels in inflorescence stem tissue (Raes *et al.* 2003).

C4H (cinnamate 4-hydroxylase) is a cytochrome P450 monooxygenase which hydroxylates *trans*-cinnamic acid to 4-coumaric acid. *C4H* is a single gene in arabidopsis (Bell-Leong *et al.* 1997). Reduced epidermal fluorescence 3 (*ref3*) mutants with decreased C4H activity have a low content of proanthocyanidins, lignin and flavonoid compared to wild type of arabidopsis (Schillmiller *et al.* 2009). Another important function of C4H is that C4H is believed to function as an anchor to endoplasmic reticulum (ER), recruiting other enzymes together to form enzyme complexes (Shirley 1999).

4CL (4-coumaroyl ligase) converses 4-coumaric acid to 4-coumaroyl CoA. Similar to *PAL*, 3 isoforms of *4CL* were isolated in 1999 (Ehlting *et al.* 1999) and a fourth family member *4CL4* was identified in 2004 in arabidopsis (Hamberger and Hahlbrock 2004). Only *4CL3* is likely to participate in the flavonoid biosynthetic pathway whereas *4CL1* and *4CL2* are involved in lignin formation and biosynthesis of other cell-wall-bound phenolic compounds (Ehlting *et al.* 1999). C4H catalyzes 4-coumaric acid to 4-coumaroyl CoA, which is the final step of pheyylpropanoid biosynthetic pathway.

CHS (chalcone synthase) and CHI (chalcone isomerase) are the first two enzymes specific to the flavonoid biosynthetic pathway. CHS condenses 4-coumaroyl CoA and 3-malonyl CoA to naringenin chalcone via cyclization reaction, and CHI converses naringenin chalcone to naringenin through isomerization. They both are frequently used to analyze the localization of enzymes, the existence of enzymatic complexes and the interactions with other enzymes in this pathway (Pelletier and Shirley 1996, Saslowsky and Shirley 2001, Saslowsky *et al.* 2005). *CHS* was first cloned in 1988 and it was found

that high light intensity hugely increases CHS activity, ultimately leading to the accumulation of anthocyanins in arabidopsis (Feinbaum and Ausubel 1988, Feinbaum *et al.* 1991). CHI is the smallest of the flavonoid biosynthetic enzymes and can move through the nuclear pore complex (Saslowsky *et al.* 2005). In addition, no anthocyanidin compounds were found in the mutants of *CHS* (*tt4*) and *CHI* (*tt5*), respectively, consistent with that they are encoded by single-copy genes (Shirley *et al.* 1995).

F3H (flavanone 3-hydroxylase) converses naringenin to dihydrokaempferol (DHK) by hydroxylation in position 3 of flavanones. DHK is the precursor of three main classes of flavonoids in arabidopsis: flavonols, anthocyanins and proanthocyanidins (Figure 1). In arabidopsis, the *F3H* gene has only one family member and is coordinately expressed with *CHS* and *CHI* in seedlings (Pelletier and Shirley 1996). The analysis of the flavonoid accumulation pattern showed that flavonols and anthocyanidins could be detected in low levels in tissues of the mutant of *F3H* (*tt6*). Further, the intermediate naringenin has never been detected in the mutant *tt6*, which should be theoretically accumulated due to the flavonoid biosynthetic defect (Shirley *et al.* 1995 and Peer *et al.* 2001). It is thought that FLS and ANS could compensate F3H activity in arabidopsis. All three enzymes are 2-oxoglutarate-dependent oxygenases with somewhat relaxed substrate specificity (Owens *et al.* 2008b). In Owens' experiments, it was shown that a part of naringenin is catalyzed to dihydrokaempferol by FLS1 (Owens *et al.* 2008a).

In arabidopsis, DHK is hydroxylated by F3'H (flavonoid 3'-hydroxylase) to produce dihydroquercetin (DHQ) which is subsequently used to produce cyanidin catalyzed by DFR and ANS (Figure 1). F3'H is another cytochrome P450-depend monooxygenase involved in the flavonoid biosynthetic pathway. It was firstly isolated from petunia (Brugliera *et al.* 1999) and was identified in 2000 to be encoded by *tt7* in arabidopsis (Schoenbohm *et al.* 2000).

In arabidopsis, FLS (flavonol synthase) is a member of the 2-oxoglutarate-dependent dioxygenases, which converses either dihydroquercetin to quercetin

or dihydrokaempferol to kaempferol. *FLS* has more than one gene family member, similar to *PAL* and *4CL* (Ohl *et al.* 1990, Ehltng *et al.* 1999, Owens *et al.* 2008a and Hamberger and Hahlbrock 2004). The arabidopsis *FLS* gene family has six members (*FLS1*, *FLS2*, *FLS3*, *FLS4*, *FLS5* and *FLS6*). However, according to biochemical and genetic analysis, only *FLS1* is thought to code for a catalytically competent protein (Owens *et al.* 2008a, Stracke *et al.* 2009). One year later, *FLS3* was identified as a second functional flavonol synthase (Preuß *et al.* 2009)

DFR (dihydroflavonol 4-reductase) is a key enzyme in the anthocyanin branch of the flavonoid biosynthetic pathway for flower colors due to its substrate specificity. It reduces the carbonyl group in position 4 of *DHK* to leucopelargonidin and *DHQ* to leucocyanidin. *Petunia* does not have orange flowers due to the inefficiency of *DFR* in this species to reduce *DHK*, but this substrate specificity could be altered (Johnson *et al.* 2001). In arabidopsis, *DFR* gene was first isolated in 1992 (Shirley *et al.* 1992) and completely no *DFR* mRNA was detected in the *DFR* mutant (*tt3*) (Shirley *et al.* 1995), consistent with that *DFR* has only one gene family member.

ANS (anthocyanidin synthase) is a 2-oxoglutarate iron-dependent oxygenase, converting leucoanthocyanidins to anthocyanidins (Saito *et al.* 1999). *ANS* was first isolated in 1997 as a “late” gene in flavonoid biosynthetic pathway in arabidopsis (Pelletier *et al.* 1997). Amongst three 2-oxoglutarate dependent oxygenases of flavonoid biosynthetic pathway, *ANS* is more closely related to *FLS* compared to *F3H* based on structural level (Wilmouth *et al.* 2002).

Glycosyltransferases (GTs) are a big chemical family. Anthocyanin 3-O-glucosyltransferase has been identified and the mutant of which showed decreased anthocyanin contents (Tohge *et al.* 2005). The enzymes that are involved in glycosylation of cyanidin, pelargonidin and delphinidin, transfer UDP-glucose to low molecular weight substrates in plants, occurring at the C-3, C-5 and C-7 positions of flavonoid aglycones, which increases the stability of aromatic nucleus. Finally, cyanidin-3-glucoside and delphinidin-3-glucoside are

further substituted by 5-glucosyltransferases (5-GTs), rhamnosyl transferases (RTs), acyltransferases (ATs) and methyltransferases (MTs) (Koes *et al.* 2005). Genes encoding anthocyanin methyltransferases have not been characterized in arabidopsis until now, but they are found in petunia and grape (Huguene *et al.* 2009, Provenzano *et al.* 2014).

2.2.2 Branches of flavonoid biosynthetic pathway in other plants

Flavonoids are ubiquitous in other plant species as well, such as potato, maize, gerbera and rice, etc. In some species, there are some specific branches of flavonoid biosynthetic pathway that do not exist in arabidopsis. For example, arabidopsis and many other plants lack isoflavone synthase (IFS), whereas legumes (soybeans, green beans and peas) have it. Therefore, legumes are capable of synthesizing isoflavones because IFS catalyzes the first committed step of isoflavone biosynthesis (reviewed by Aoki *et al.* 2000). Isoflavone synthase was identified and the expression of soybean *IFS* in arabidopsis leads to the production of isoflavone genistein that is not naturally synthesized in arabidopsis (Jung *et al.* 2000). The production of isoflavone genistein was achieved in tobacco plants as well. At the same time, when CHR that can produce the substrate for isoflavone synthesis was introduced to tobacco, isoflavone daidzein was produced (Yu *et al.* 2000). Further, there are a few known species that can synthesize 3-deoxyanthocyanins such as sorghum (Dykes *et al.* 2009), gloxinia (*Sinningia cardinalis*) (Winefield *et al.* 2005) and maize (*Zea mays*) (Halbwirth *et al.* 2003).

2.3 Yeast two-hybrid systems

2.3.1 Molecular mechanism of the GAL4 yeast two-hybrid system

In 1986, a transcriptional factor GAL4 (881 amino acids) that binds with a specific DNA sequence in yeast was discovered. GAL4 activates transcription in the presence of galactose. When the GAL4 was split into two parts (N-terminal fragment and C-terminal fragment), only N-terminal fragment bound to DNA

sequence but without the function of activating transcription in the presence of galactose. However, when these two fragments combined together, they could again activate the transcription. Therefore, two domains, DNA-binding domain (DB) and activation domain (AD), were identified (Keegan *et al.* 1986). Based on this finding, yeast two-hybrid (Y2H) system was proposed to detect protein-protein interactions in living yeast cells of *Saccharomyces cerevisiae*. Proteins of interest that were fused to DB domain were named as baits and proteins that were fused to AD domain were named as preys. If the proteins on bait and prey constructs could interact, they complete the transcriptional factor followed by the activation of the reporter genes, containing the GAL4 binding site in their promoters (Fields & Song 1989). The GAL4 based Y2H revolutionized protein-protein interaction experiments but still has several limitations (reviewed by Brückner *et al.* 2009).

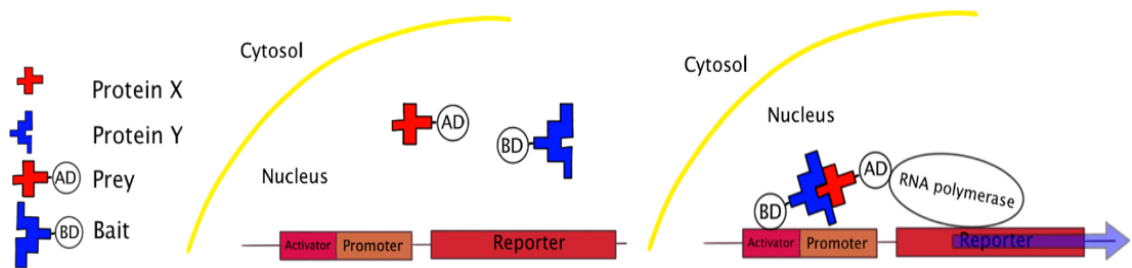


Figure 2. The classic yeast two hybrid system. In nucleus, when prey and bait interact, RNA polymerase is recruited and activates the transcription of reporter gene.

2.3.2 Development and modification of the yeast two-hybrid system

The traditional GAL4 Y2H system has been widely adopted. However, it requires the protein interactions to occur in the nucleus. Johnson and Varshavsky (1994) proposed an improved system based on split-ubiquitin. It allows the detection of protein reactions in the cytosol. In this system, one protein is fused to the C-terminal half of ubiquitin (Cub) with a transcription factor (TF), whereas another protein is fused to a mutated N-terminal half of ubiquitin (NubG). The mutation in NubG prevents spontaneous linking of the Cub and Nub halves. Once the two proteins interact, the ubiquitin molecule becomes functional and the TF is cleaved off by ubiquitin specific proteases

(UBPs). Being free to diffuse into the nucleus, the TF activates the transcription of reporter genes (Johnsson and Varshavsky 1994).

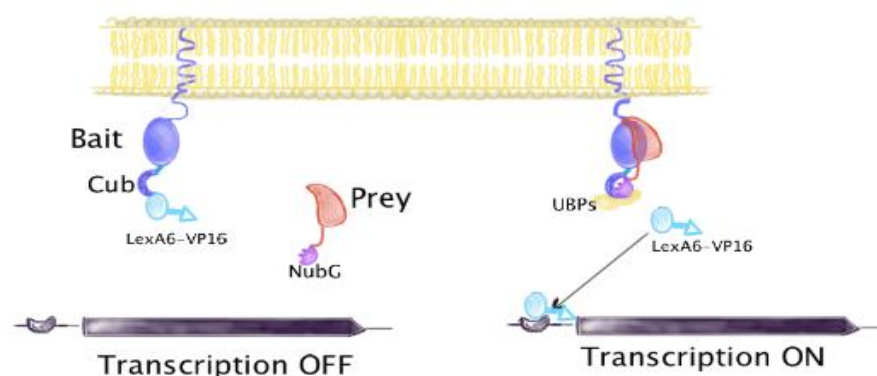


Figure 3. The DUAL hunter system (Modified from the User Manual of DUAL hunter starter Kit)

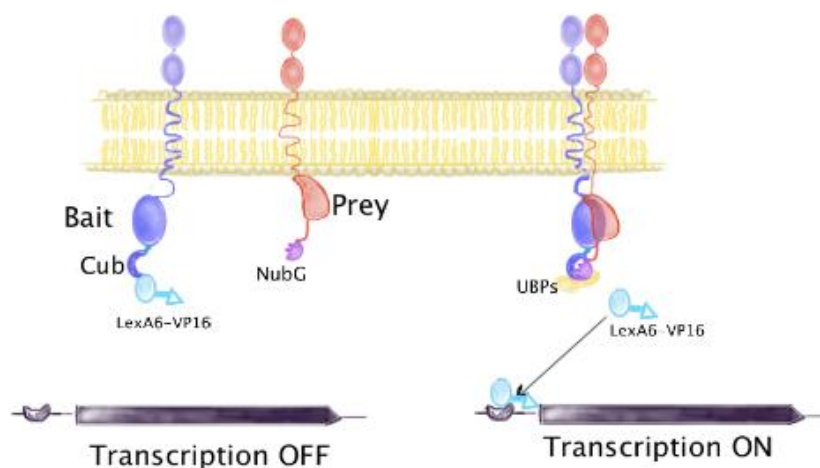


Figure 4. The DUAL membrane pairwise system (Modified from the User Manual of DUAL membrane pairwise interaction Kit)

Later on, Stagljar and his colleagues took advantage of this split-ubiquitin system and developed it to the membrane yeast two-hybrid system that could be used to detect interactions of integral membrane proteins (Stagljar *et al.* 1998). One protein is fused to Cub with transcriptional factor LexA-VP16 and anchored to a membrane. Another protein is fused to mutated NubG. While two proteins are interacting, the Cub and NuG reconstituted. The polypeptide between Cub and LexA is cleaved by UBPs (Figure 3). Similarly, for membrane proteins, they are separately fused to vectors that have peptide signaling, which leads them to membrane (Figure 4). Once two proteins interact, the Cub and

NuG reconstituted and the reporter genes are activated. Because proteins are anchored to membranes, they are incapable of locating in nucleus, which avoids false positives deriving from proteins that act as activators.

In addition to that, Y2H system was extended to many directions. For example, a one-hybrid system was proposed in 1993 to identify genes that recognize a specific DNA sequence (Li and Herskowitz 1993). Based on that, a Gateway-compatible yeast one-hybrid (Y1H) system was designed to rapid, large-scale protein-DNA identification (Deplancke *et al.* 2004) and a bacterial one-hybrid system was proposed in 2005 to determine the DNA-binding specificity of a transcription factors (Meng *et al.* 2005). In 1996, a new system, named three-hybrid was developed to analyze specific RNA-protein interactions relying on physical, rather than biological properties, of the RNA (SenGupta *et al.* 1996). This system was later applied on detecting ligand–receptor interactions *in vivo* (Licitra and Liu 1996), detecting biocatalysts *in vivo* (Firestine *et al.* 2000) and analyzing mRNA-protein complexes (Bernstein *et al.* 2002)

2.4 Metabolons in plants

An amount of evidence suggests that metabolons play important roles in metabolic pathways (reviewed by Ovadi & Srere 2000). Metabolons lead to substrate channeling, meaning that reaction intermediates do not freely diffuse out of the metabolons. Metabolons are thought to be tightly linked multi-enzyme complexes. The formation of metabolons improves catalytic efficiency of biosynthesis and offers a “safe” place for plants to synthesize natural products without or with reduced metabolic interference and supply a way for avoiding toxic intermediates (Shirley 1999). Besides, metabolons also co-ordinate metabolic cross-talk and provide a possibility for a swift re-direction of metabolism (Jørgensen *et al.* 2005)

Stafford (1974) first suggested that enzymes involved in phenylpropanoid and flavonoid biosynthetic pathway are functioning as multi-enzyme complexes. Burbulis and Shirley (1999) elucidated that some of the enzymes involved in the

flavonoid biosynthetic pathway interact with each other in the GAL4 yeast two-hybrid system. More specifically, CHS, CHI and DFR showed interactions with DFR, CHS and CHI, respectively, and the interaction occurred in a defined orientation (Burbulis and Shirley 1999). Similarly, FLS1 displayed interactions with F3H and DFR in both orientations whereas it only interacted with CHS when FLS1 was fused to the activation domain in the yeast two-hybrid system (Owens *et al.* 2008a). Many enzymes in phenylpropanoid and flavonoid biosynthetic pathways are encoded by a single gene in arabidopsis, which simplifies the investigation of protein interactions *in vivo*. One of the exceptions is that *FLS* has six gene family members but only FLS1 was found to have catalytic function in flavonoid biosynthetic pathway. Still, FLS5 displays interactions with DFR and CHS in yeast two-hybrid system (Owens *et al.* 2008a). The results demonstrate that non-enzymatic proteins of flavonoid biosynthetic pathway could be involved in forming metabolons, although probably only at a structural level.

Evidence supports that endoplasmic reticulum (ER) is a site of phenylpropanoid and flavonoid metabolism (Wagner & Hrazdina 1987) and that metabolons exist in the phenylpropanoid and flavonoid biosynthetic pathways (Reviewed by Shirley 1999). The immunolocalization assay showed an association between CHS and endoplasmic reticulum membranes in buckwheat (Hrazdina *et al.* 1987). Later on, CHS was discovered to co-localize with CHI at the endoplasmic reticulum and tonoplast in arabidopsis (Saslowsky and Shirley 2001). It was believed that enzymes are recruited to endoplasmic reticulum via a weak interaction in a linear arrangement by membrane proteins (C4H and F3'H), which act as anchors to the membrane (Hrazdina and Wagner 1985b, Saslowsky and Shirley 2001). However, it was also suggested that multiple contacts exist in proteins, indicating a globular complex rather than the linear array (Burbulis and Shirley 1999). On the other hand, Saslowsky and his colleagues found that at least two enzymes of flavonoid biosynthesis pathway, CHI and CHS, are located both in the cytoplasm and in the nuclei of some cells (Saslowsky *et al.* 2005). This dual cytoplasmic/nuclear localization gives rise to a new way how these enzymes function in the flavonoid biosynthesis pathway.

However, the relationships of other enzymes in phenylpropanoid and flavonoid biosynthetic pathway are still unknown, particularly interactions between the membrane bound P450 enzymes of the pathway (C4H and F3'H) have not been demonstrated. In my experiments, the interactions between enzymes of these two pathways were investigated by using the Dual hunter and membrane yeast two-hybrid system.

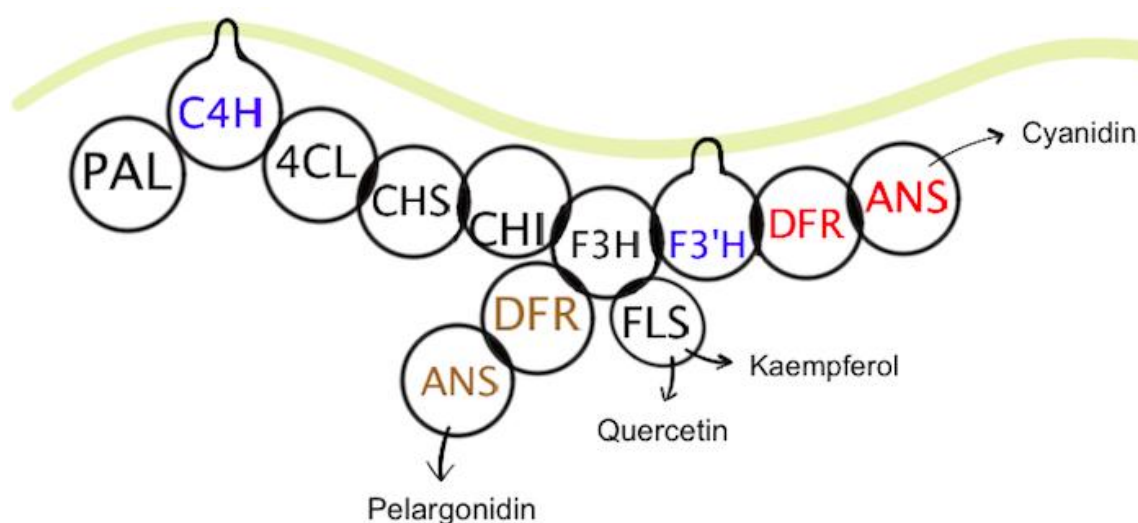


Figure 5. Models for the organization of flavonoid biosynthetic pathway enzymes as complexes at the endoplasmic reticulum (Modified from Winkel-Shirley 1999).

3. OBJECTIVES

The aim of this master's thesis was to analyze the interactions of proteins that are involved in flavonoid biosynthetic pathway. More specifically, the objectives were 1). to amplify 12 genes (*F3H*, *DFR*, *CHS*, *FLS1*, *FLS3*, *4CL1*, *GT*, *PAL*, *ANS*, *CHI*, *F3'H* and *C4H*) from cDNA, deriving from arabidopsis flower tissue. 2). to construct yeast bait and prey vectors and transform them into yeast strains. 3). to assay protein-protein interactions by the Dual hunter yeast two-hybrid system and the DUAL membrane pairwise interaction system.

4. MATERIALS AND METHODS

4.1 Plant material and growth conditions

Arabidopsis was grown in peat-vermiculite (volume ratio of 1:1) in the greenhouse at around 24 °C. The day length was 18 hours and the humidity was controlled by using water spray from time to time in the greenhouse.

4.2 DUAL hunter system and DUAL membrane pairwise system

In these two systems, a protein of interest (the bait) is fused to the C-terminal half of ubiquitin (Cub) and the transcriptional factor LexA-VP16, whereas the prey is fused to the mutated N-terminal half of ubiquitin (NubG). When the bait and prey interact, LexA6-VP16 will be released and turn on the transcription (Figure 3 and Figure 4). Consequently, due to the expression of two auxotrophic growth markers (*HIS3* and *ADE2*), the yeast strains could successfully grow on selection plates (Dualsystems Biotech). In the experiments, a DUAL membrane pairwise system was applied to F3'H and C4H because they are membrane proteins (Figure 4).

4.3 The cloning vectors and yeast strains for bait and prey

In the experiments, six cloning vectors were used for transformation of bait and prey: pPR3-N, pPR3-SUC, pBT3-SUC, pDHB1, pTSU2-APP and pOst1-Nubl (Figure 6). pPR3-N and pPR3-SUC are the vectors for prey cloning, whereas pDHB1 and pBT3-SUC are vectors for bait cloning. At the same time, control vectors were used to check the function of bait and prey. Specifically, pOst1-Nubl is the positive control and pPR3-N is the negative control for bait. pTSU2-APP is the positive control for prey. The expression of pOst1-Nubl led to the production of the wild type Nub portion of yeast ubiquitin which has strong affinity with Cub expressed from bait constructs. The combination of Cub and Nub ubiquitin liberate the LexA-VP16 transcription factor, activating the expression of two auxotrophic growth markers *HIS3* and *ADE2*, which leads to

the growth of yeast on SD -trp -leu -his -ade selection medium. By contrast, NubG expressed from pPR3-N prey vector has no affinity with Cub of bait vectors. Then split-ubiquitin cannot be formed and reporter genes are not activated. Consequently, NMY51 cannot survive on SD -trp -leu -his -ade selection plates. The reason for using pPR3-SUC and pBT3-SUC for F3'H and C4H is that these two proteins are membrane enzymes of the flavonoid biosynthesis pathway. The vectors for preys always consist of ampicillin resistance and vectors for baits have kanamycin resistance.

NMY51 (α) and NMY61 (α) are two yeast strains constructed for the Y2H systems used. In yeast, two strains respectively have an ' α loci' and ' α loci', which enables them to mate together. Here, NMY61 is the mating partner for NMY51. NMY51 is a standard reporter strain carrying the *HIS3* and *ADE2* reporter gene. *HIS3* and *ADE2* are two auxotrophic growth markers, the expressions of which make yeast survive on defined minimal medium without histidine and adenine. Besides, *HIS3* is a selection marker on medium with addition of 3-Amino-1,2,4-triazole (3-AT). When *ADE2* report gene is not transcribed, the adenine synthesis pathway is blocked and a red color appears. However, activation of *ADE2* report gene will unblock the adenine metabolic pathway, which finally leads to the a presence of faint pink to white color, depending on the strength of the interaction (DUAL membrane pairwise interaction kit, Dualsystems Biotech).

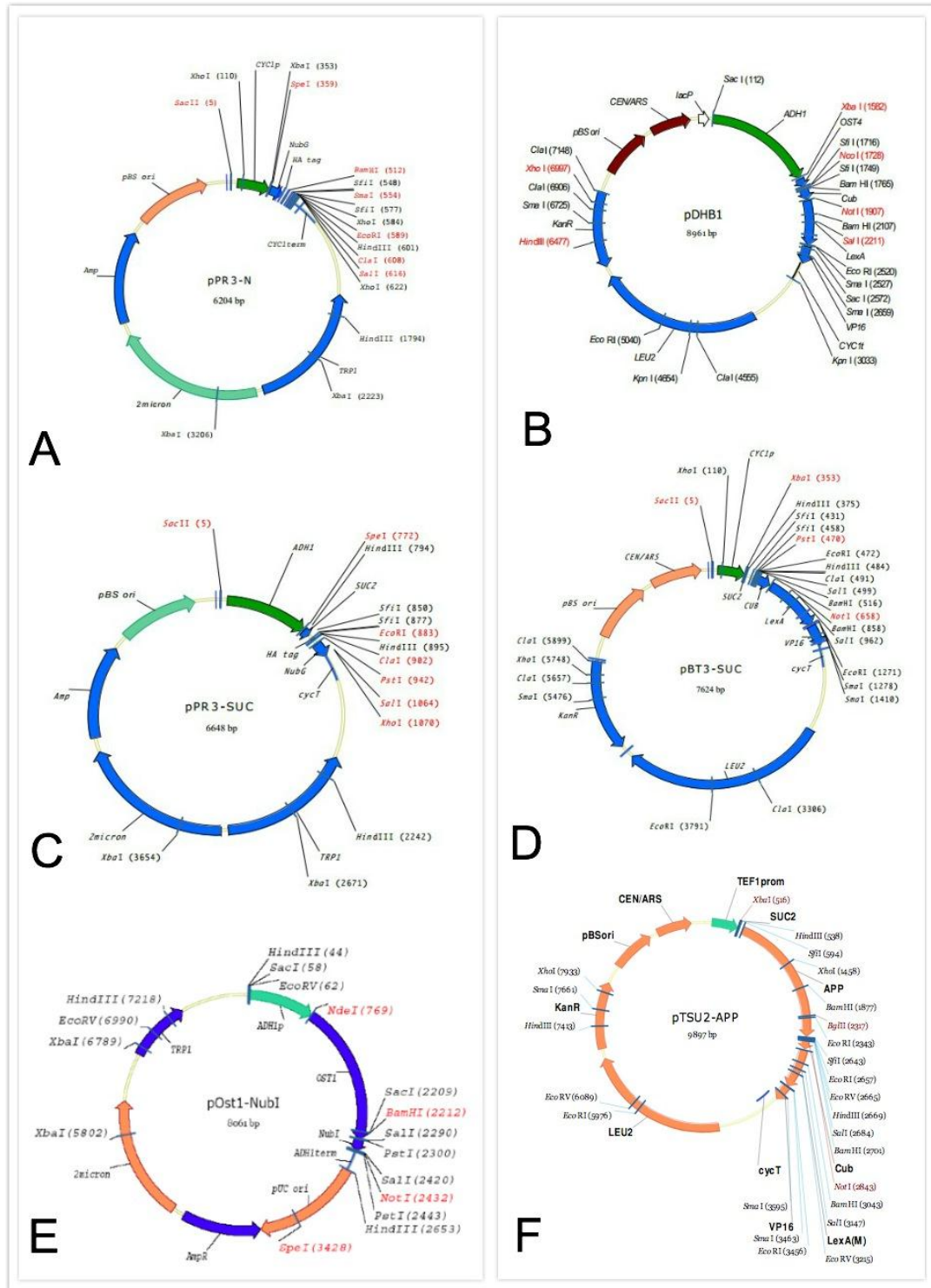


Figure 6. Maps and features of vectors for cloning. A=pPR3-N, B=pDHB1, C=pPR3-SUC, D=pBT3-SUC, E=pOst1-NubI and F=pTSU2-APP. pPR3-N and pPR3-SUC are vectors for prey with ampicillin resistance (Amp^R) and Trp selection marker. pDHB1 and pBT3-SUC are vectors for bait with kanamycin resistance (Kan^R) and Leu selection marker. pOst1-NubI is the control vector for baits with Amp^R and Trp selection marker. pTSU2-APP is the control vector for preys with Kan^R and Leu selection marker.

4.4 Primer design

For cloning of each gene insert, the bait and prey vectors share the same forward primer. The forward primer consists of three parts. The buffer component is followed by the *SfiI* sites (GGCC ATTAC GGCC), increasing the efficiency of *SfiI*. The reason for using *SfiI* sites is that this site is quite rare in eukaryotic genomes. More importantly, all *SfiI* sites are not identical, which ensure the right orientation of inserts to vectors. The third part is the gene-specific sequence, starting with the ATG codon. For the reverse primers of baits and preys, it consists of buffer sequence AACTGATT from 5' to 3', *SfiI* site (GGCC CAGGC GGCC) and the gene specific sequence. But the difference is that the reverse primers of prey end with the stop codon of the ORF (Opening read frame). By contrast, the stop codons of the reverse primer of baits are removed, which allows the continuous translation from N-terminus to C-terminus, resulting in the expression of *Cub-LexA-VP16* (Appendix 4).

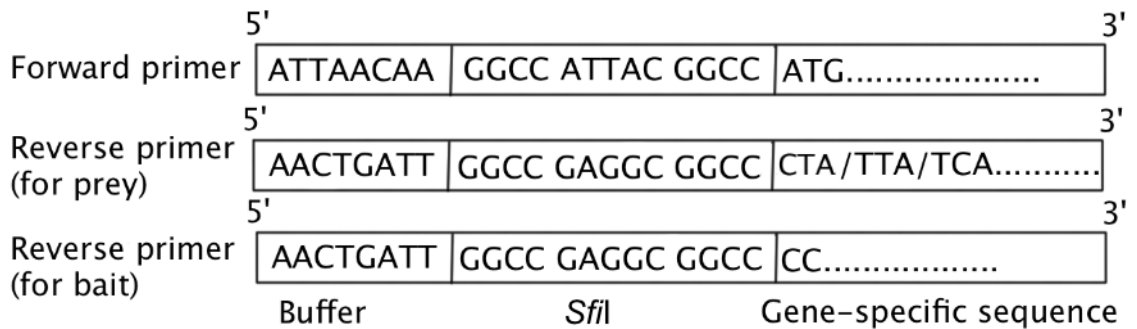


Figure 7. The primer design of forward primers and reverse primers. Primers consist of buffer, *SfiI* site and gene-specific sequence. The gene-specific sequence of forward primer always starts from ATG. The reverse primer for prey vectors ends up with one of the stop codons (TAG, TAA and TGA). The 'CC' in the reverse primer for bait vectors is to ensure in-frame fusion with downstream *Cub-LexA-VP16*.

4.5 RNA extraction

Total RNA was extracted by TRIzol Reagent (Invitrogen). Arabidopsis petals were collected and ground in a 1.5 ml Eppendorf tube into power under liquid nitrogen. The processed samples (around 100 mg) were incubated in 1 ml TRIzol reagent and homogenized at RT for 5 min. Afterward, 200 μ l of

chloroform was added followed by vigorously shaking for 15 seconds and incubation at room temperature (RT) for 3 min. Tubes were centrifuged at 12,000 x g at RT for 15 min. The aqueous phase was moved to a new Eppendorf tube, followed by addition of 500 μ l of 100% isopropanol and then centrifuged at 12,000 x g in 4 min after they had been incubated at RT for 10 min. The RNA pellet was washed with 1ml of 75% ethanol, and then centrifuged at 7,500 x g at RT for 5 min. Finally, the RNA pellet was re-suspended in 30 μ l RNase-free water after the ethanol was totally evaporated and incubated in 55°C for 10 min. 1 μ l of RNA with 1 x loading buffer was loaded on 1% 0.5xTBE agarose gel (~7 cm) at 20 min at 100 V. The RNA was stored at -80°C.

4.6 cDNA synthesis

SuperScript™ III Reverse Transcriptase Kit (Invitrogen) was used for cDNA synthesis. Briefly, 1 μ l of RNA was mixed with 0.5 μ l of 10 μ M Oligo (dT)₁₅ and 1 μ l of 10 mM dNTP mix and then adjusted to 13 μ l by sterile, distilled water. The mixture was heated at 65°C for 5 min and incubated on ice for 2 min, followed by the addition of 4 μ l of 5xFirst-strand buffer, 1 μ l of 0.1M DTT, 1 μ l of “RNase out” RNase inhibitor and 1 μ l of SuperScript™ III Reverse Transcriptase. The mixture was incubated at 50°C for 1 hour and then the reaction was terminated by heating at 70°C for 15 min and stored at -20°C.

4.7 Polymerase chain reaction and electrophoresis

The Polymerase chain reactions (PCR) were conducted with Phusion High-Fidelity DNA Polymerase (Thermo scientific) (Table 1). After the PCR, the amplified products were run on 1% 0.5xTBE agarose gel. The correct sized fragments were cut and collected under UV light and then purified according to the Gel Extraction Kit (Omega). Lambda/PstI (Figure 8) was used as the marker for electrophoresis.

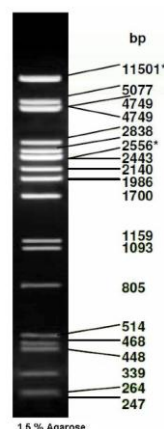


Figure 8. The Lambda/PstI marker (Adopted from GeneOn: <http://www.taq-dna.com>)

Table 1. The PCR reaction mix and PCR program

PCR reaction system		PCR programme		
Reagents	Amount		Temperature	Time
cDNA	1 μ l	1	98°C	30 s
5xPhusion HF Buffer	10 μ l	2	98°C	10 s
dNTP(10mM each)	1 μ l	3	X °C	30 s
Forward primer* (10mM)	2.5 μ l	4	72 °C	30s-2 min
Reverse primer (10mM)	2.5 μ l	5	go back to step 2	29 cycles
Phusion DNA Polymerase	0.5 μ l	6	72 °C	10 min
Distilled & sterile water	to 50 μ l		4 °C	Forever

* Primers are shown in Appendix 4. The annealing temperature varies according to each primer pair, and is presented in Appendix 4.

4.8 Gene isolation

Both the target PCR products and the four vector DNAs (pPR3-N, pDHB1, pPR3-SUC and pBT3-SUC) were digested with *Sfi*I Fast digestion enzyme (Thermo Scientific) at 50°C for 1 hour to produce cohesive ends. Then, the digested PCR products and vectors were separated in a 1% 0.5xTBE gel and purified with the Gel Extraction Kit (Omega). 2 μ l of purified PCR products were ligated to vectors by mixing with 10 μ l of 10xReaction buffer, 1.5 μ l of cut plasmid DNA and 16.5 μ l of nuclease-free water and 1 μ l of T4-DNA ligase (5 U

/ μl), and incubated at RT for 15 min.

5 μl of each ligation reaction was pipetted into a tube with 200 μl competent cells (DH5 α) and then incubated on ice for 30 min. Subsequently, heat shock was conducted at 42°C for 1 min. Then, 700 μl Luria Broth (LB) was added into the tubes and they were incubated at 37°C for 1 hour with shaking. The supernatants were removed after centrifugation at 2500 rpm for 3 min. Finally, the pellets were re-suspended in 150 μl LB and transferred to plates with either kanamycin or ampicillin antibiotics (25 μg / ml Kanamycin or 100 μg / ml ampicillin). The plates were incubated at 37°C for overnight.

4.9 Construct confirmation

Single colonies from the transformation plate were picked and purified on a new selection plate and then checked by colony PCR to confirm the right insert.

Table 2. The PCR reaction mix and PCR program for Colony PCR

PCR reaction system			PCR programme	
Reagents	Amount		Temperature	Time
10x Taq buffer	2 μl	1	95°C	3 min
25 mM MgCl ₂	1.2 μl	2	94°C	30 s
dNTP(10mM each)	2 μl	3	67 °C	30 s
Forward primer*(10mM)	0.4 μl	4	72 °C	30 s
Reverse primer*(10mM)	0.4 μl	5	go back to 2	29 cycles
Phusion DNA Polymerase	0.5 μl	6	72 °C	10 min
distilled &sterile water	to 20 μl		4 °C	Forever

* Primer lists are presented in Appendix 5.

A single colony was transferred by a sterilized tip from plates to a PCR tube. At the same time, the tips with remaining bacterial cells were streaked on new antibiotic plates for culture further use (Miniprep) once the insertions were confirmed. 20 μl of mixture (Table 2) was added into the PCR tubes, followed by

a gentle vortex. Then, the PCR was conducted under the program (Table 2).

4.10 Plasmid DNA extractions

The plasmid DNAs were isolated by GeneElute™ HP Plasmid Miniprep Kit (Sigma). Generally, one single colony from selection plate was cultured in 5 ml LB supplemented with antibiotics (25 µg / ml Kanamycin or 100 µg / ml ampicillin) at 37 °C overnight with shaking. The plasmids were isolated according to the manufacturer's instructions. The plasmid DNA was digested by fast digest Enzymes *Sfi*I (at 50°C for 10 min (Thermo scientific) (Table 3). Once the insert DNA was confirmed, plasmids containing right sized inserts were sent for sequencing. The sequences were aligned by using ClustalW2 (EMBI-EBI). Complete alignment was required in the experiments as any nucleotide change would lead to unexpected protein products.

Table3. The components of Plasmid Fast Digestion

Component	Amount
Plasmid DNA	1 µl
Enzyme	1 µl
10xFast Digest Green Buffer	2 µl
water	to 20 µl

4.11 Bait testing

A single colony of yeast strains NMY51 from fresh plate (not older than one week) was inoculated into 10 ml YPAD growth medium and grown at 28 °C with shaking for overnight. The overnight culture was used to adjust 50 ml fresh YPAD so that the OD 600 was around 0.2-0.3. The 50 ml yeast cultures were incubated at 28 °C with shaking (200 RPMs) until the OD600 reached 0.6-0.7. The cultures were centrifuged at 1,000 x g for 5 min, and then the cells were re-suspended in 40 ml of 1xTE. Afterwards, the cells were centrifuged at 1,000 x g for 5 min and re-suspended in 2 ml of 0.1 M LiAc/0.5xTE. Finally, the cells were incubated at RT for 10 min.

For each transformation, 1.5 µg both plasmid DNAs (baits and control preys) and 100 µg denatured (boiled before use) salmon sperm DNA with 100 µl of the yeast cell suspension from previous step were well mixed. 700 µl of 0.1 M LiAc/40% PEG-3350/1xTE were added, following incubation at 28°C for 30 min without shaking. 88 µl of DMSO was added and mixed well with heat shock at 42 °C for 15 min. Subsequently, the tubes were kept on ice for 2 min. Supernatants were removed after centrifuging 10 seconds with full speed and the pellet was re-suspended in 1 ml of 1x TE and re-pelleted. After that, the pellets were re-suspended in 200 µl of 1x TE and 50 µl of each were plated onto three different plates: YPAD (A complex medium to inhibit reversion of the *ade1* and *ade2* mutations), SD -leu -trp (synthetically defined medium lacking leucine and tryptophan), SD -trp -leu -his -ade selection medium (SD medium lacking tryptophan, leucine, histidine, and adenine). The plates were incubated at 28°C for five days and then the growing results were recorded.

4.12 Yeast two-hybrid assay

The prey vectors and bait vectors with correct inserts were individually transformed into both NMY51 and NMY61 as described above. One colony of each construct from fresh plate was inoculated into 5 ml YPAD medium for overnight growing at 28°C with shaking (200 RPMs). 3 µl drops of the bait cultures were pipetted on the YPAD plate and then 3 µl of the prey culture were set on the top of bait droplets. After drying, plates were incubated at 28°C for overnight. Short strokes were drawn from each spot with wooden stick onto SD -leu -trp plate for 3 to 4 days growth at 28°C to select for presence of both plasmids. Afterwards, strokes were made again onto YPAD growth medium, SD -leu -trp -his -ade growth medium, and SD -leu -trp growth medium plates in this particular order, using the same stick. YPAD plates were placed at 28°C for overnight growth for the β-galactosidase assay. Other plates were placed at room temperature for 5-7 days.

In my study, prey of CHS, bait of CHS and prey of F3'H obtained from Gerbera Laboratory, University of Helsinki, which explains that they were absent from

colony PCR and Sfil digest (Figure 12, 13, 14 and 15)

4.13 β -galactosidase assay

When *LacZ* reporter gene is activated by the transcription factor, β -galactosidase was encoded. The X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside) will yield insoluble blue compounds after it is hydrolyzed by β -galactosidase. In this way, the expression of *LacZ* reporter could be analyzed via the presence of blue color.

Two round-shaped 50 mm Whatman 541 filter papers, saturated with 3.5 ml of 2% X-gal solution, were placed on a 10 cm petri dish. Any air bubbles were removed. The yeasts from the surface of the YPAD plates were obtained by a 50 mm Whatman 541 filter paper and then completely immersed in liquid nitrogen for 15 seconds and then set them on the top of the soaked Whatman filters. Plates were sealed with a parafilm and incubated at 37°C for overnight. The results were recorded after 24 hours by photography.

5. RESULTS

5.1 RNA extraction and cDNA synthesis

Total RNA from arabidopsis flowers was isolated and RNA quality was checked on agarose gel. Bands of 28s and 18s ribosomal RNAs were expected to be clear on the gel. The result showed 28s and 18s ribosomal RNAs as well as some other bands (Figure 9). They probably came from the minor DNA contamination and the degradation of RNA. The RNAs were used as templates for cDNA synthesis.

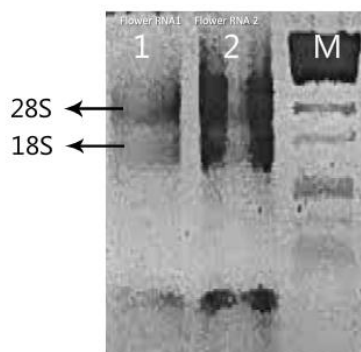


Figure 9. The RNAs of arabidopsis flowers on 1% 0.5xTBE agarose gel. M=Lambda/PstI marker. 1 and 2=RNAs of arabidopsis flower, comprising 28S and 18S rRNA

5.2 PCR reactions and electrophoresis

Twelve genes involved in flavonoid biosynthesis were planned to be amplified from cDNA of arabidopsis flowers. They were *PAL*, *C4H*, *4CL1*, *CHS*, *CHI*, *F3H*, *F3'H*, *DFR*, *FLS1*, *FLS3*, *ANS* and *GT*. The RT-PCR reactions were conducted for 30 cycles (Table 2). Amplified products were examined on agarose gel. 10 out of 12 target candidates were successfully amplified, exceptions being *FLS3* and *4CL1*. The sizes of PCR products matched with predictions (Figure 10 and Figure 11).

For each gene, two pairs of primers were used for PCR amplification. Forward primers were shared but the sequences of reverse primers for prey vectors had a stop codon whereas those for bait vectors did not (Appendix 4). However, the primers of membrane genes, *F3'H* and *C4H*, for prey and bait vectors (pBR3-SUC and pBT3-SUC) share the same sequences without stop codons. The reason of not including stop codons from primers for vectors pDHB1, pBR3-SUC and pBT3-SUC is allow the continuous translation from N-terminus to C-terminus, resulting in the expression of *Cub-LexA-VP16* (Figure 4B).

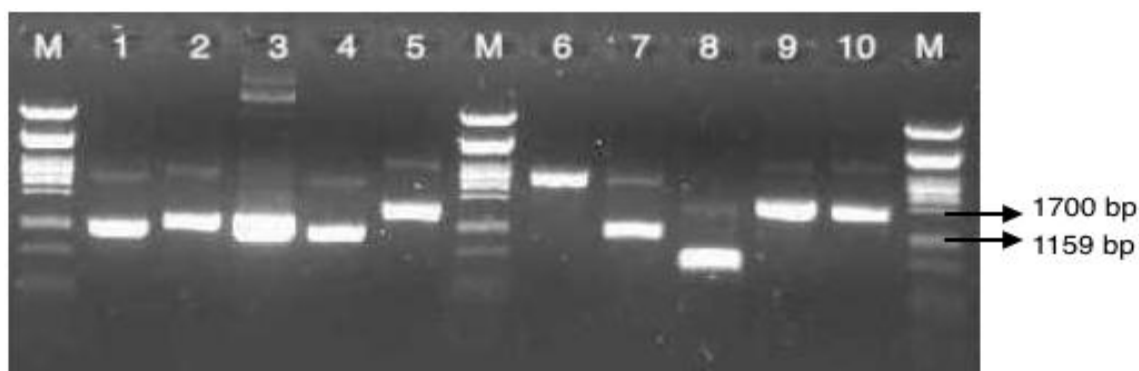


Figure 10. PCR amplified bait genes on 1% 0.5xTBE agarose gel. M=Lambda/PstI marker (bands of 1700 bp and 1159 bp marker). 1=*F3H* (1077 bp), 2=*DFR* (1149 bp), 3=*CHS* (1188 bp), 4=*FLS1* (1011 bp), 5=*GT* (1350 bp), 6=*PAL* (2178 bp), 7=*ANS* (1071 bp), 8=*CHI* (741 bp), 9=*F3'H* (1542 bp) and 10=*C4H* (1518 bp).

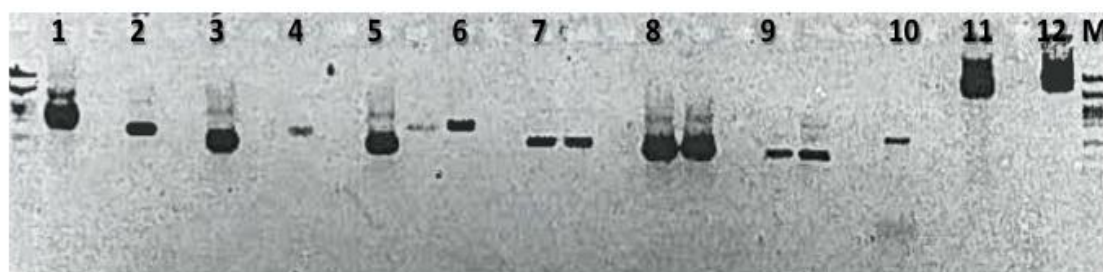


Figure 11. PCR amplified prey genes on 1% 0.5xTBE agarose gel. M=Lambda/PstI marker. 1=*PAL* (2178 bp), 2=*C4H* (1518 bp), 3=*CHS* (1188 bp), 4=*GT* (1350 bp), 5=*F3H* (1077 bp), 6=*F3'H* (1542 bp), 7=*ANS* (1071 bp), 8=*FLS1* (1011 bp), 9=*CHI* (741 bp), 10=*DFR* (1149 bp), 11=pPR3-N (digested), 12=pPR-SUC (digested).

5.3 Construction of the bait and prey vectors

The PCR products were purified from the gel and then digested with the restriction enzyme *Sfi*I. All vectors were digested in the same conditions to produce cohesive ends. Each gene was ligated to appropriate vectors and then transformed into *E.coli* for propagation. More specifically, genes were separately ligated to prey vector pPR3-N and bait vector pDHB1 and the two membrane genes (*F3'H* and *C4H*) were cloned into prey vector pPR3-SUC and bait vector pBT3-SUC. The single colonies growing on selection plates were examined by colony PCR to confirm that vectors contained correct inserts (Figure 12 for preys and Figure13 for baits). The verified colonies were propagated and then used for plasmid DNA extraction.

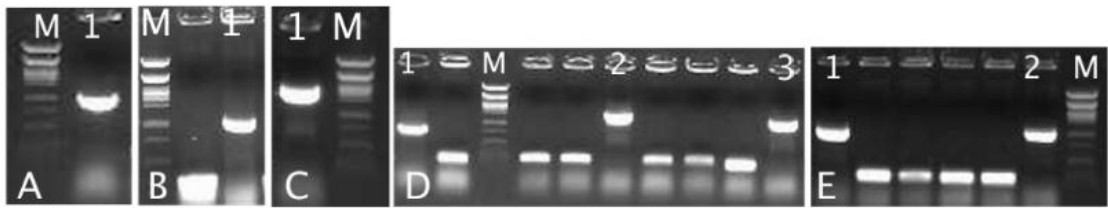


Figure 12. Colony PCR of prey clones on 1% 0.5xTBE agarose gel. M= Lambda/*Pst*I marker. A1= *ANS*, B1= *DFR*; C1=*PAL*, D1=*CHI*, D2=*C4H*, D3=*GT*, E1=*FLS1* and E2=*F3H*. The smaller bands presents in panel D and E are primer dimers.

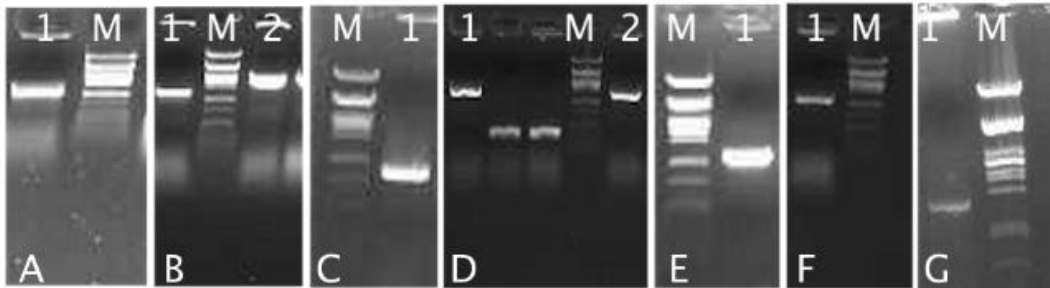


Figure 13. Colony PCR of bait clones on 1% 0.5xTBE agarose gel. M= Lambda/*Pst*I marker. A1=*ANS*, B1=*DFR*, B2=*PAL*, C1=*CHI*, D1=*C4H*, D2=*GT*, E1=*FLS1*, F1=*F3H* and G1= *F3'H*.

5.4 *Sfi*I digestion

All the plasmids were digested by the enzyme *Sfi*I to further confirm correct sizes of inserts. The plasmids pPR3-N, pDHB1, pPR3-SUC, and pBT3-SUC and candidate genes were well separated (Figure 14 and Figure 15). The inserts (Appendix 4) were correct as well, compared to predictions and the marker.

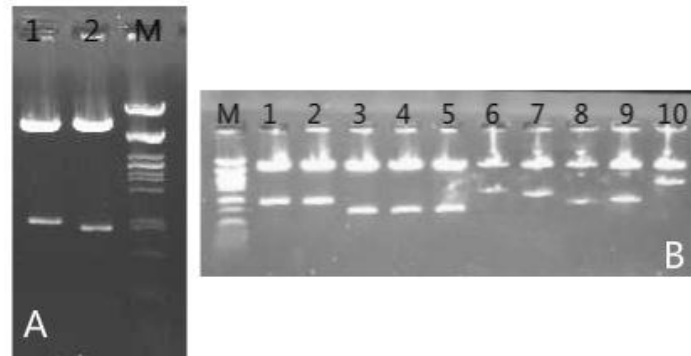


Figure 14. The digestion of prey plasmid DNA containing target genes with *SfiI* on 1% 0.5×TBE agarose gel. M=Lambda/*PstI* marker. A1=*DFR*, A2=*ANS*, B1 and B2=*ANS*, B3, B4 and B5=*CHI*, B6=*C4H*, B7=*GT*, B8=*FLS1*, B9=*F3H*, B10=*PAL*.

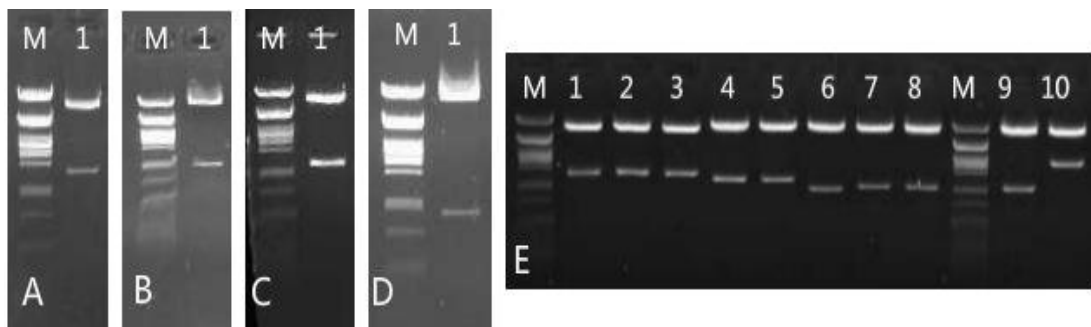


Figure 15. The digestion of plasmid bait DNA containing target genes with *SfiI* on 1% 0.5×TBE agarose gel. M=Lambda/*PstI* marker. A1=*F3H*, B1=*ANS*, C1=*CHI*, D1=*FLS1*, E1, E2 and E3=*C4H*, E5 and E6=*GT*, G7, G8 and G9=*DFR*, G10=*PAL*.

5.5 The gene sequencing

The verified plasmid constructs were sent for sequencing and the sequences were aligned with published sequences from computer database. The alignment results showed that all the nucleotide sequences completely matched with the published sequences (data are not shown here), except that one nucleotide difference (G→A) was observed in *CHI* (Figure 16). However, the protein alignment analysis, by using EMBOSS Transeq and ClustalW2 (EMBI-EBI), showed no difference between the two translational protein products (Figure 17). Therefore, all the plasmids were further used on yeast transformation.

```

PatCHI      TATACGGACTGTGAAGCTAAAGCTGTGGAGAAATTCTTGGAGATCTTCAAGGAAGAAACA 462
PatCHIF2    TATACGGACTGTGAAGCTAAAGCTGTGGAGAAATTCTTGGAGATCTTCAAGGAAGAAACA 540
*****

```

Figure 16. The alignment between sequencing result and published arabidopsis *CHI*. P=Prey; at=arabidopsis; F2=Forward sequencing.

```

AtCHI      -----MSSSNACASPSPPFAVTKLHVDSVTFVPSVKSPA 34
PatCHIF    WGSYPYDVPDYAGSKQWYQRRVAITAMSSSNACASPSPPFAVTKLHVDSVTFVPSVKSPA 60
*****

AtCHI      SSNPLFLGGAGVRGLDIQGFVIFTVIGVYLEGNAVPSLSVKWKGKTTEELTESIPFFRE 94
PatCHIF    SSNPLFLGGAGVRGLDIQGFVIFTVIGVYLEGNAVPSLSVKWKGKTTEELTESIPFFRE 120
*****

AtCHI      IVTGAFEKFIKVTMKLPLTGQQYSEKVTENCVAIWKQLGLYTDCEAKAVEKFLEIFKEET 154
PatCHIF    IVTGAFEKFIKVTMKLPLTGQQYSEKVTENCVAIWKQLGLYTDCEAKAVEKFLEIFKEET 180
*****

AtCHI      FPPGSSILFALSPTGSLTVAFSKDDSI PETGIAVIENKLLAEAVLESIIGKNGVSPGTRL 214
PatCHIF    FPPGSSILFALSPTGSLTVAFSKDDSI PETGIAVIENKLLAEAVLESIIGKNGVSPGTRL 240
*****

AtCHI      SVAERLSQLMMKNKDEKEVSDHSVEEKLAKEN----- 246
PatCHIF    SVAERLSQLMMKNKDEKEVSDHSVEEKLAKENGRGLSLRIRYQAYRYRRPRVMLVMSRLH 300
*****

```

Figure 17. The alignment of translation between published and sequenced CHI. AtCHI is the protein translated from the published gene sequence. PatCHIF is the protein from the translation of the gene with one nucleotide error (Figure 16).

Table 4. Summary of the prey and bait constructs

Name of plasmids	Description *	Reference
pAX1	Prey-at PAL: arabidopsis specific <i>PAL</i> cloned into pPR3-N; without stop codon	This work
pAX2	Bait-at PAL: arabidopsis specific <i>PAL</i> cloned into pDHB1; without stop codon	This work
pAX3	Prey-at C4H: arabidopsis specific <i>C4H</i> cloned into pPR3-SUC; without stop codon	This work
pAX4	Bait-at C4H: arabidopsis specific <i>C4H</i> cloned into pBT3-SUC; without stop codon	This work
pAX5	Prey-at CHS: arabidopsis specific <i>CHS</i> cloned into pPR3-N; with stop codon	Bashandy et al.
pAX6	Bait-at CHS: arabidopsis specific <i>CHS</i> cloned into pDHB1; without stop codon	Bashandy et al.
pAX7	Prey-at CHI: arabidopsis specific <i>CHI</i> cloned into pPR3-N; with stop codon	This work
pAX8	Bait-at CHI: arabidopsis specific <i>CHI</i> cloned into pDHB1; without stop codon	This work
pAX9	Prey-at F3H: arabidopsis specific <i>F3H</i> cloned into pPR3-N; with stop codon	This work
pAX10	Bait-at F3H: arabidopsis specific <i>F3H</i> cloned into pDHB1; without stop codon	This work
pAX11	Prey-at F3'H: arabidopsis specific <i>F3'H</i> cloned into pPR3-SUC; without stop codon	Bashandy et al.
pAX12	Bait-at F3'H: arabidopsis specific <i>F3'H</i> cloned into pBT3-SUC; without stop codon	This work
pAX13	Prey-at DFR: arabidopsis specific <i>DFR</i> cloned into pPR3-N; with stop codon	This work
pAX14	Bait-at DFR: arabidopsis specific <i>DFR</i> cloned into pDHB1; without stop codon	This work
pAX15	Prey-at FLS1: arabidopsis specific <i>FLS1</i> cloned into pPR3-N; with stop codon	This work
pAX16	Bait-at FLS1: arabidopsis specific <i>FLS1</i> cloned into pDHB1; without stop codon	This work
pAX17	Prey-at ANS: arabidopsis specific <i>ANS</i> cloned into pPR3-N; with stop codon	This work
pAX18	Bait-at ANS: arabidopsis specific <i>ANS</i> cloned into pDHB1; without stop codon	This work
pAX19	Prey-at GT: arabidopsis specific <i>GT</i> cloned into pPR3-N; with stop codon	This work
pAX20	Bait-at GT: arabidopsis specific <i>GT</i> cloned into pDHB1; without stop codon	This work

*All inserts were cloned in SfiI restriction enzyme site.

pPR3-N and pPR3-SUC have ampicillin resistance and Trp selection marker.

pDHB1 and pBT3-SUC have kanamycin resistance Leu selection marker.

5.6 Bait testing

In order to test if baits were functional or not, bait testing was conducted. Prey testing was not done because there were no positive control vectors (baits) for it. Bait vectors were co-transformed with either the positive control prey construct pOst1-Nubl or negative control prey construct pPR3-N into NMY51 by heat-shock. In bait testing, there should be robust growth of yeast on SD -trp -leu medium when the transformation was successful because only yeasts that harbor both bait and prey constructs can synthesize tryptophan and leucine. In addition, yeast growth should be observed on SD -trp -leu -his -ade selection medium when baits were co-transformed with pOst1-Nubl as the non-mutated Nubl fragment spontaneously links with the Cub fragment, and no yeast growth were displayed when baits were co-transformed with pPR3-N.

After co-transformation, yeast strain NMY51 was incubated on SD -trp -leu medium and SD -trp -leu -his -ade selection medium for 3 days at 28 °C. The growth of yeast was quantified by comparing the number of colonies on SD -trp -leu -his -ade plates to the number of colonies on SD -trp -leu plates. Bait testing showed that NMY51 had a robust growth under selection plates (Table 5), especially for yeast with pAX18 and pAX16, after they were co-transformed with pOst1-Nubl. On the other hand, NMY51 harboring bait constructs and negative construct pPR3-N, presented no or few yeast colonies on SD -trp -leu -his -ade selection plates. Particularly, no yeast growth was detected on NMY51 with pAX10, pAX16, pAX18 and pAX4. By contrast, around 20 colonies were observed on SD -trp -leu -his -ade selection plates, which account for around 6 %. The growth percentage was still lower compared to a strongly robust yeast growth compared on SD -trp -leu plates.

Table 5. The number of single colonies on SD -trp -leu and SD -trp -leu -his -ade plates. %growth under selection=number of colonies on SD -trp -leu -his -ade plates / number of colonies on SD -trp -leu plates.

Plasmid names	Genes	+ pOst1-Nubl Number of colonies			+pPR3-N Number of colonies		
		SD -trp -leu	SD -trp -leu -his -ade	%growth under selection	SD -trp -leu	SD -trp -leu -his -ade	%growth under selection
pAX2	<i>PAL</i>	70	20	29 %	60	1	2 %
pAX4	<i>C4H</i>	90	30	33 %	80	0	0 %
pAX8	<i>CHI</i>	120	50	42 %	140	5	3 %
pAX10	<i>F3H</i>	140	35	25 %	90	0	0.0%
pAX12	<i>F3'H</i>	200	90	45 %	250	7	3 %
pAX14	<i>DFR</i>	400	200	50 %	330	20	6 %
pAX16	<i>FLS1</i>	80	50	62 %	180	0	0 %
pAX18	<i>ANS</i>	80	60	75 %	200	0	0 %
pAX20	<i>GT</i>	80	25	31 %	200	1	1 %

5.7 Yeast two-hybrid assay

All the sequencing-verified bait and prey vectors were transformed into two yeast lines: NMY51 (a) and NMY61 (α). In yeast, two strains respectively have an 'a loci' and ' α loci', which enables them to mate together. Here, NMY61 is the mating partner for NMY51. The diploid cells formed by mating between NMY51 and NMY61, which ultimately led to prey/bait combination in every yeast cell. If baits and preys interacted, the reporter genes would be activated, resulting in yeast growth on selection plates. Due to the *ADE2* reporter gene, the yeast strain NMY51 would display colors, ranging from red (weak interaction) to white (strong interaction), which depends on the strength of the interactions of protein pairs.

The Dual hunter yeast two-hybrid system was used to analyze protein-protein interactions for most of the flavonoid biosynthetic enzymes. In Dual Hunter, the bait is attached to the ER membrane by fusing it to the membrane protein. Because F3'H and C4H are membrane proteins, a DUAL membrane pairwise interaction system was applied. For each assay, NMY51 with baits and NMY61 with preys were used for mating and vice versa.

All proteins interacted with the positive control preys (pOst1-Nubl), indicating the functionality of the baits (Figure 18, Table 6 and Table 7). However, certain proteins interacted with the negative control preys (pPR3-N). They were F3H, DFR, FLS1 and CHI, which led to false positive results. Membrane proteins (F3'H and C4H) interacted with any other proteins. Simultaneously, preys interacted with negative control baits (pTSU2-APP), except for pAX13 (DFR), pAX11 (F3'H) and pAX3 (C4H). Based on interactions with controls, it was conclusive that PAL only interacted with itself (Figure 18B, 18E, 18H and 18I). ANS had a strong interaction with CHI, showing robust yeast growth compared to faint yeast growth with negative control pPR3-N. CHS and GT had no interactions with other enzymes when they are fused on bait vectors. However, CHS interacted with CHI and DFR when CHS was fused on prey vector.

For the two membrane proteins, yeast was survival and robust growth was

observed when pAX12 (F3'H) and pAX4 (C4H) interacted with other preys, including pOst1-Nubl and pPR3-N (Figure 18). However, F3'H and C4H did not interact with any of the other nine proteins except with DFR, CHI and with themselves.

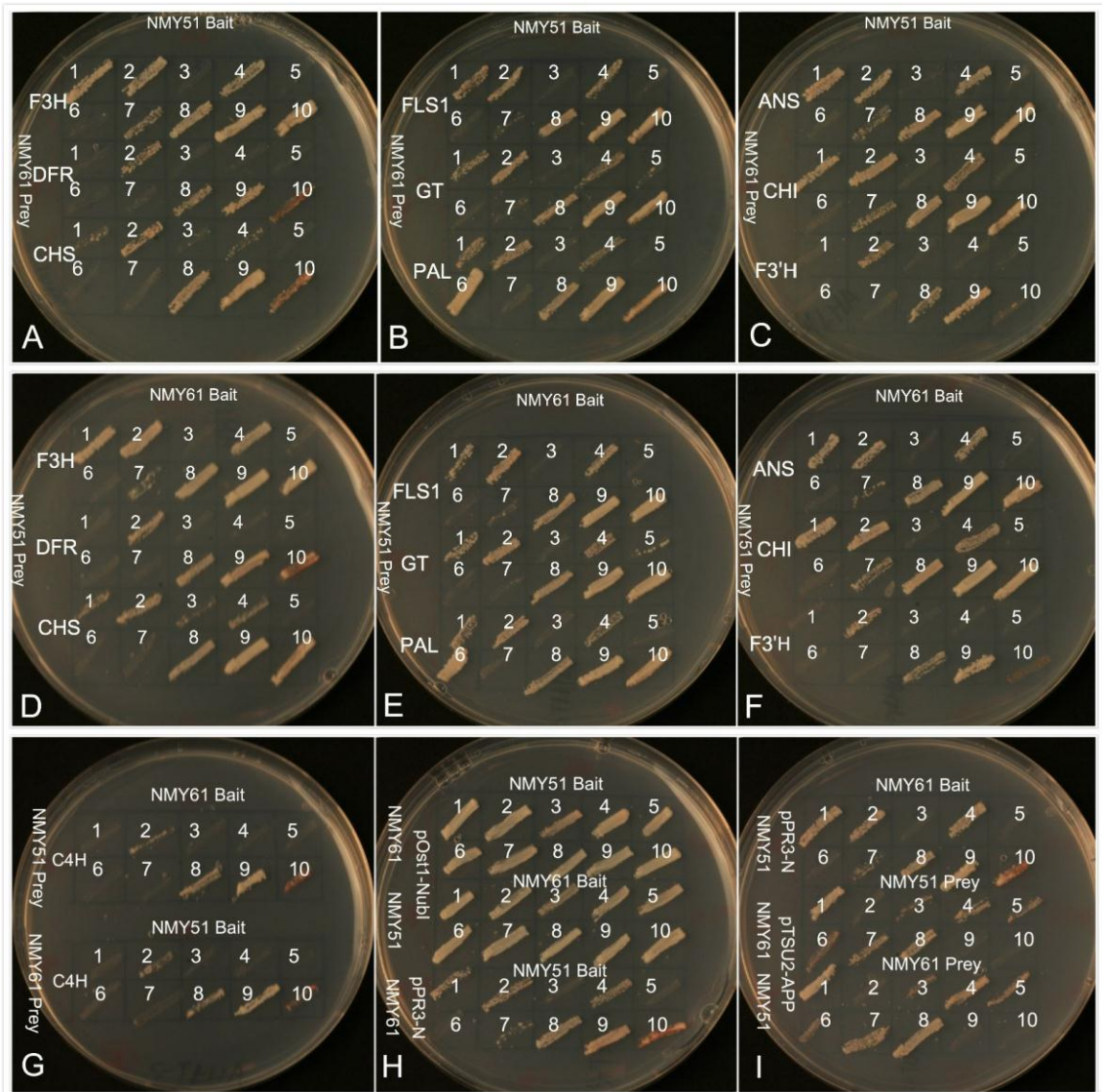


Figure 18. The yeast two-hybrid assay on SD -trp -leu -his -ade selection medium. Panel A, B, C and G demonstrate the mating between yeast strain NMY61 containing the prey constructs with strain NMY51 containing the bait constructs. Correspondingly, panel D, E, F and G demonstrate the mating between preys in NMY51 with baits in NMY61. Numbers present different genes cloned in bait vectors (pDHB1 or pBT3-SUC): 1=*F3H*, 2=*DFR*, 3=*CHS*, 4=*FLS1*, 5=*GT*, 6=*PAL*, 7=*ANS*, 8=*CHI*, 9=*F3'H* and 10=*C4H*. Panel H and I shows the mating of target baits and preys with control preys and baits. pOst1-Nubl and pPR3-N are the positive and negative control preys, respectively; pTSU2-APP is the negative control bait for prey construct.

Table 6. Summary of the yeast two-hybrid assay based on Figure 18A, 18B 18C, 18G, 18H and 18I (Preys in yeast strain NMY61 and baits in yeast strain NMY51)

		Preys (NMY 61)												
		F3H	DFR	CHS	FLS1	GT	PAL	ANS	CHI	F3'H	C4H	pOst1-Nubl (Positive control)	pPR3-N (negative control)	
Baits (NMY51)	F3H	++ ¹	-	-	+	+	+	+	+	-	-	++	+	
	DFR	++	+ ²	+	+	++	+	+	++	+	+	++	+	
	CHS	-	-	-	-	-	-	-	-	-	-	++	-	
	FLS1	++	-	-	+	+	+	+	+	-	-	++	+	
	GT	-	-	-	-	-	-	-	-	-	-	++	-	
	PAL	-	-	-	-	-	-	++	-	-	-	-	++	-
	ANS	+	-	-	-	-	-	-	+	+	-	-	++	-
	CHI	++	+	++	++	++	++	++	++	++	+	+	++	++
	F3'H	++	++	++	++	++	++	++	++	++	+	++	++	++
	C4H	++	+	+	++	++	++	++	++	++	-	+	++	++
pTSU-APP (negative control)	++	-	+	++	+	+	+	+	+	-	-	-	-	

1. ++ presents robust yeast growth on SD -trp -leu -his -ade selection plates.
2. + presents weak yeast growth on SD -trp -leu -his -ade selection plates.
3. Red presents red color was observed on yeast.
4. - in table means no yeast growth was observed, which means weak interaction.

Table 7. Summary of the yeast two-hybrid assay based on Figure 18D, 18E 18F, 18G, 18H and 18I (Preys in yeast strain NMY51 and baits in yeast strain NMY61)

		Preys (NMY51)												
		F3H	DFR	CHS	FLS1	GT	PAL	ANS	CHI	F3'H	C4H	pOst1-Nubl (Positive control)	pPR3-N (negative control)	
Baits (NMY61)	F3H	++ ¹	-	+ ²	+	+	+	+	++	-	-	++	+	
	DFR	++	++	+	+	++	+	+	++	+	+	++	+	
	CHS	-	-	-	-	-	-	-	-	-	-	++	-	
	FLS1	++	-	+	+	+	+	+	+	-	-	++	++	
	GT	-	-	-	-	-	-	-	-	-	-	++	-	
	PAL	-	-	-	-	-	++	-	-	-	-	++	-	
	ANS	+	-	-	-	-	-	-	+	-	-	++	-	
	CHI	++	++	++	++	++	++	++	++	++	+	+	++	++
	F3'H	++	++	++	++	++	++	++	++	++	+	++	++	++
	C4H	++	+	++	++	++	++	++	++	++	-	+	++	++
pTSU-APP (negative control)	++	-	+	++	+	+	+	+	+	-	-	-	-	

1. ++ presents robust yeast growth on SD -trp -leu -his -ade selection plates.
2. + presents weak yeast growth on SD -trp -leu -his -ade selection plates.
3. Red presents red color was observed on yeast.
4. - in table means no yeast growth was observed, which means weak interaction.

5.9 β -galactosidase assay

In addition to the expression of two reporter genes (*HIS3* and *ADE2*), another report gene *lacZ* in NMY51 was used to assess the strength of interactions of the protein individual transformants. The *E.coli* gene *LacZ* encodes a β -galactosidase that can be assayed with the chromogenic substrate X-gal. Filter papers were used to obtain yeast cells on YPAD plates and then were immersed in liquid nitrogen. The filter paper was next transferred on top of new filter papers soaked with X-gal solution in a clean plate. If two candidate proteins interact, they will activate the expression of *lacZ*, resulting in the production of β -galactosidase. Consequently, yeast cells will display in blue in the presence of X-gal.

Obvious blue color was detected on places where baits and positive prey pOst1-Nubl were streaked, which confirms the functionality of baits. However, blue prints were displayed also when pAX4 (C4H), pAX8 (CHI), pAX12 (F3'H) and pAX14 (DFR) were streaked together with negative control pPR3-N, consistent with results of Y2H assay (Figure 18, Table 6 and Table7). However, pAX10 (F3H) and pAX16 (FLS1) did not display blue with pPR3-N (Figure 19H and 19I) whereas yeast growth was observed on SD -trp -leu -his -ade selection plates (Figure 18H and 18I). Further, while pAX1 and pAX2 were incubated together, strong blue was displayed (Figure 19B and Figure 19E), indicating a strong self-interaction of PAL itself. Besides, it is clear that blue color was presented when F3'H and C4H were streaked with all others, even including the negative prey control pPR3-N.

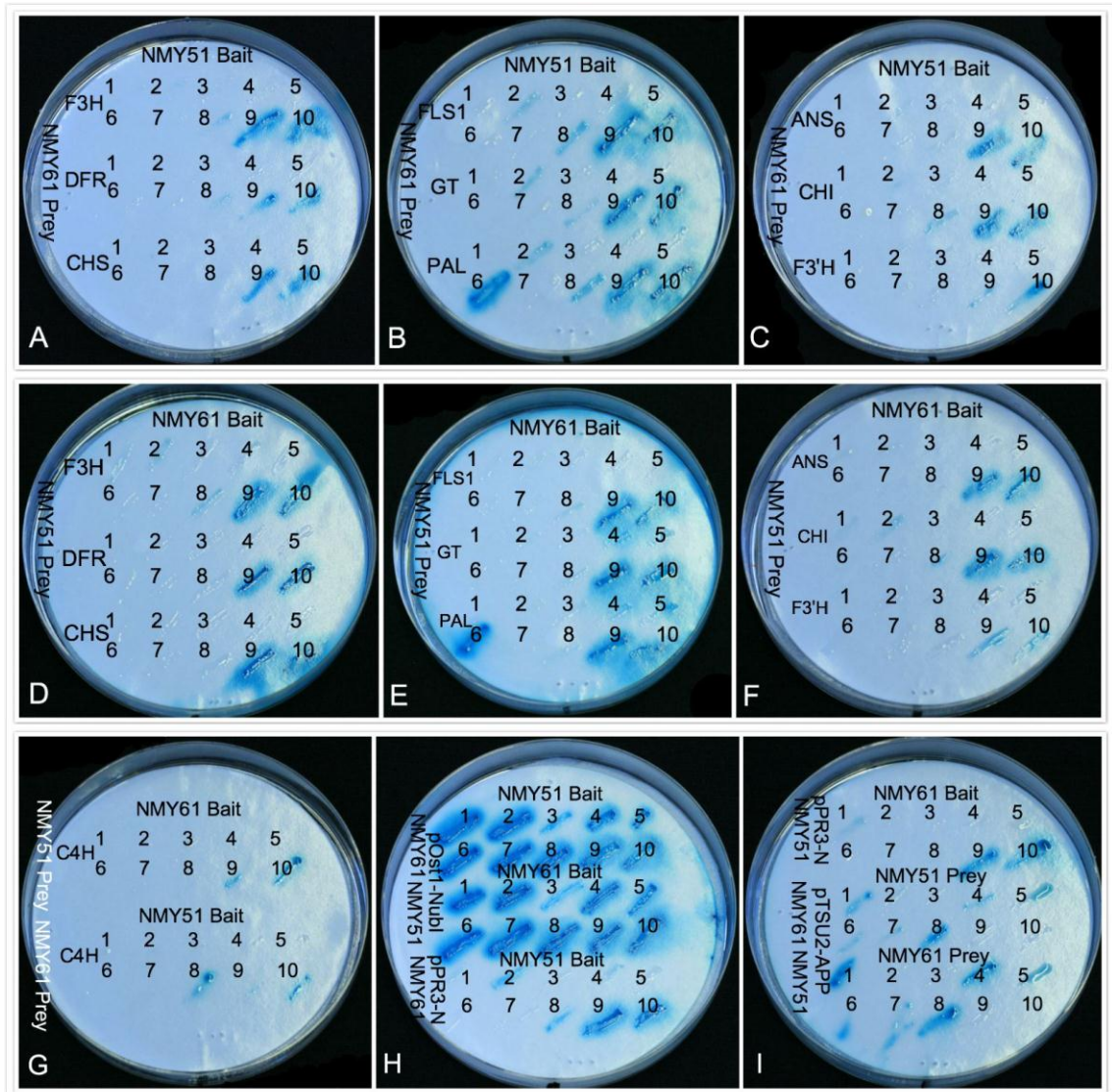


Figure 19. The X-gal assay of yeast two-hybrid assay. Panel A, B, C and G demonstrate the mating between yeast strain NMY61 containing the prey constructs with strain NMY51 containing the bait constructs. Correspondingly, panel D, E, F and G demonstrate the mating between preys in NMY51 with baits in NMY61. Numbers present different genes ligated on the bait vectors (pDHB1 or pBT3-SUC): 1=*F3H*, 2=*DFR*, 3=*CHS*, 4=*FLS1*, 5=*GT*, 6=*PAL*, 7=*ANS*, 8=*CHI*, 9=*F3'H*, and 10=*C4H*. Panel H and I present the mating of target baits and preys with control preys and baits. pOst1-Nubl and pPR3-N are the positive and negative control preys, respectively; pTSU2-APP is the negative control bait for prey construct.

6. DISCUSSION

6.1 Gene isolation

In the study, ten out of twelve full-length genes were successfully amplified from cDNA of the arabidopsis flower. Two genes failed (*FLS3* and *4CL1*). In order to amplify these two, several annealing temperatures were applied but still no products were amplified. FLS is a member of the 2-oxoglutarate-dependent dioxygenases, which has six family members (Pelletier *et al.* 1997). *FLS3* was secondly identified to involve in flavonols synthesis (Preuß *et al.* 2009), but it was either undetectable or had a low expression level in flowers, checked by using both Semi quantitative RT-PCR and microarray method (Owens *et al.* 2008a).

4CL converts 4-coumaric acid to 4-coumaroyl CoA in arabidopsis. Four family members of *4CL* were identified in arabidopsis (Ehlting *et al.* 1999, Hamberger and Hahlbrock 2004). *4CL3* may participate in flavonoid biosynthetic pathway whereas *4CL1* and *4CL2* are likely to participate in lignin biosynthesis (Ehlting *et al.* 1999). *4CL1* is mainly expressed in roots and stems (Lee *et al.* 1995) and has low expression levels in mature leaves and flowers (Ehlting *et al.* 1999).

In my study, samples were collected from arabidopsis flowers. However, *FLS3* and *4CL1* are both expressed in flowers at a low level (Ehlting *et al.* 1999, Owens *et al.* 2008a), which explains why they were not amplified from cDNA made from arabidopsis flowers.

The plasmids were all sequence-verified. All genes completely aligned with published arabidopsis sequences, except *CHI* had one nucleotide difference (G->A) (Figure 16). The protein alignment confirmed the same translation products (proteins) (Figure17).

6.2 Bait testing showed a weak self-activation of bait vectors

In the bait testing, baits of interest were co-transformed to yeast strains NMY51 with either positive control prey pOst1-Nubl or negative control prey pPR3-N. Yeast growth was displayed when baits and pOst1-Nubl were simultaneously streaked on SD -trp -leu -his -ade selection plates, which indicated the functionality of baits. At the same time, no yeast growth should be displayed on selection plates when baits and negative control prey pPR3-N were co-transformed. Contradictorily, a few yeast colonies were observed on selection plates 3 days later after baits were co-transformed with negative control prey pPR3-N (Table 5). These protein pairs are GT, PAL, CHI and F3'H, which probably indicated a weak self-interaction or protein instability.

In order to eliminate the self-activation of baits, an extra experiment 3-AT titration could have been conducted. 3-AT acts as a competitive inhibitor of *HIS3* reporter gene (Durfee et al. 1993). In this step, the bait is transformed without any preys into NMY51 reporter strain and selected by using SD -leu -his -ade + 3-AT. On this selection medium, the yeast growth that comes from self-activation is removed. The lowest 3-AT concentration without yeast growth could be applied for the next Y2H assay.

6.3 The enzymatic interactions based on Y2H assay

The knowledge of enzymatic interaction gives clues to elucidate how these enzymes function and where do they locate in cells. In my experiments, PAL showed strong self-interaction, which suggested that in arabidopsis PAL is also possible to form dimers between protein macromolecules. In parsley (*Petroselinum crispum*), the crystal structure of PAL was reported and PAL was shown to exist as homotetramers (Ritter and Schulz 2004). Similarly, CHS, a plant-specific polyketide synthase in this pathway, is homodimeric (Ferrer et al. 1999).

In my experiments, CHS in bait had no interactions with other proteins whereas CHS in prey presented interactions with CHI and DFR. It is similar to Burbulis's

study, which demonstrated the interaction between CHS and DFR in a specific orientation (Burbulis and Shirley 1999). In the Y2H systems used in my study, baits have Cub half and transcription factor linked to C-terminus and preys have NuG half linked to N-terminus. These extra parts may block the interaction sites of proteins. Therefore, the change of the orientation of NubG half in prey would affect the protein interactions. Moreover, pAX20 (GT) had no interactions with other proteins and pAX19 (GT) presented interactions with some others although it also interacted with negative control. It still indicated a possibility that enzymes involving in flavonoid biosynthesis pathway interact with others in a specific orientation. This is similar to Burbulis and Shirley's results, which showed interactions between CHS and DFR, CHI and CHS, and DFR and CHI only for specific fusion pairs in the GAL4 Y2H (Burbulis and Shirley 1999). Similarly, FLS1 only interacted with CHS while FLS was fused to activation domain in yeast two-hybrid system (Owens *et al.* 2008a). It is interesting to mention that FLS5 that is not involved in this pathway interacts with DFR and CHS, it presents the possibility that FLS5 is involved in metabolons (enzymatic complexes) formation only on structural levels rather than functional levels (Owens *et al.* 2008a). These factors on structural levels may play important roles in investigating protein-protein interactions.

The protein interactions in my experiments indicate the existence of enzymatic complexes in the phenylpropanoid biosynthetic pathway and flavonoid biosynthetic pathway. Many published data support this as well. For example, an affinity chromatography experiment showed protein-protein interactions in plant cells amongst CHS, CHI and F3H (Burbulis and Shirley 1999). But it is still unknown that if these enzymes interact transiently or stably. Besides, due to the false positives, further experiments are required. In order to further confirm the interactions between proteins, other approaches, such as co-immunoprecipitation, pull-down assay and affinity electrophoresis, could be adopted (Reviewed by Berggård *et al.* 2007).

6.4 Improvements of yeast two-hybrid assay

6.4.1 Sources of false positives in yeast two-hybrid

In my study, false positives (yeast was growing in the presence of negative control (pPR3-N)) were observed (Figure 18, Table 6 and Table 7). There are plausible reasons for this. In the bait testing, bait construct and control prey construct were co-transformed into only one yeast strain NMY51. By contrast, in the test of protein-protein interactions, two different strains NMY51 and NMY61 were used. The mating of NMY51 and NMY61 formed a diploid, which created a new physiological environment for the baits. In the future, it would be more convincing and reasonable to use the same methods in both bait testing and the assay of protein-protein interactions. Secondly, in the bait testing, weak self-activation of baits was observed. The yeast strain NMY51 without optimal conditions was still used for yeast two-hybrid testing because the yeast growth on selection plates was very low. The colony number was from 1 to 20, compared to hundreds of yeast colonies with positive prey control pOst1-Nubl on selection plates (Table 5). Still, the weak self-activation of baits is likely to cause the false positive in the yeast two-hybrid testing.

Yeast two-hybrid system is generally claimed to have false-positive and false-negative results, which has been shown on other research (Ito *et al.* 2001). For example, Matthews failed to recapitulate two-hybrid interactions reported by other studies (Matthews *et al.* 2001; Ito *et al.* 2002). False positive results are caused by variable reasons. For example, a high expression level of bait and prey and the possibilities of wrongly folded proteins cause unspecific interactions. Based on sources for false positive results, approaches to detect and eliminate false positive results in yeast two-hybrid system are developed.

6.4.2 Approaches to improve the reliability of yeast two-hybrid system

False positives are commonly displayed in yeast two-hybrid system. Therefore, eliminating false positives is the main way to improve the reliability of this system. In the membrane Y2H system, low-copy plasmids were applied in case

of false-positives deriving from the over-expression of baits and preys. Furthermore, a stringent reporter system was used, relying on three independent markers (two auxotrophic reporter genes $LYS2::(\text{lexAop})_4\text{-HIS3}$, $ade2::(\text{lexAop})_8\text{-ADE2}$ and a $ura3::(\text{lexAop})_8\text{-lacZ}$ marker).

Considering false-positives from bait auto-activation, the test for the expression of reporter genes can be conducted as a control in yeast cells that only containing the baits. This approach was used as well in Walhout's research to eliminate false-positives before a screen for large-scale protein interaction mapping (Walhout and Vidal 2001). For the auto-activation happening during the testing procedure, a negative selection was developed relied on negative selectable marker *CYH2*. In this approach, the prey plasmids from positive yeast cells were removed and then the activities of reporter genes were tested for those yeast cells harboring only bait plasmids (Vidalain *et al.* 2004).

Furthermore, the observation of the result of yeast two-hybrid is quite personal dependent, especially when they are compared to positive and negative controls. The yeast growth is difficult to quantify on agar medium. Therefore, incubating the transformed yeast cells in liquid media could be a supplement. McCusker and Haber (1990) found that the yeast metabolism and mutant phenotype varied on agar medium compared to liquid medium. In this way, after the transformed yeast cells were incubated in liquid selection medium for a fixed time, the OD600 could be measured and analyzed. This way, the yeast growth under selection could be quantified.

6.5 β -galactosidase assay confirms yeast two-hybrid results

In the X-gal assay, the strong blue on baits with positive prey control pOst1-Nubl demonstrated the functionality of baits. However, blue color was still found on places where baits were streaked with pPR3-N, indicating the existence of the self-activation of baits, consistent with the results of yeast two-hybrid assay. Moreover, some interactions between baits and negative control prey pPR3-N occurred on SD -trp -leu -his -ade selection plates whereas the interactions were not observed based on the expression of the reporter gene *lacZ*, such as

F3H and FLS1. This probably comes from the variance of different marker genes and how the marker genes are assayed. In Serebriiskii's research, three approaches including X-gal overlay, X-gal plate and β -galactosidase were applied to test the expression of *LacZ* reporter. The results of X-gal plate and β -galactosidase assay showed similarities. However, X-gal overlay assay demonstrated big differences compared to other two approaches. It was concluded that the observation of *lacZ* expression was not strictly related with the transcriptional activation of *lacZ* reporter (Serebriiskii *et al.* 2000).

7. CONCLUSIONS

The yeast two-hybrid system cannot completely reflect the real interaction between two proteins. In reality, apart the affinity, other factors including expression, stability and three-dimensional structure of a fused protein could influence the protein interactions (Vidal *et al.* 1999). It is more convincing to use one of other approaches, like pull down assay, BiFC, immunoprecipitation and mass spectrometry, etc., to validate identified protein-protein interactions by Y2H system. In my study, the possibilities of enzymatic interactions were tested using the DUAL hunter yeast two-hybrid system and the yeast membrane two-hybrid system. Due to the high false positives, the interaction results were not highly conclusive. However, based on the bait and prey interactions with the control preys and baits, the results still showed that PAL interacted with itself and ANS interacted with CHI. It is possible that ANS also interacts with F3H and itself depending on the yeast growth compared a very faint yeast growth in the presence of negative control pPR3-N. (Figure 18, Table 6 and Table 7). Besides, GT and CHS show no interactions with any other proteins. Meanwhile, some protein interactions happened only in a certain orientation, which is similar to Burbulis and Shirley's research (Burbulis and Shirley 1999).

8. ACKNOWLEDGMENTS

This master thesis was carried out at the Gerbera Laboratory, Department of Agricultural Sciences, and University of Helsinki.

Firstly, I express my thanks to Prof. Teemu Teeri, offering me a position in his lab for practical training and advanced training, which is the beginning of my laboratory experiences in Helsinki. After that, I continued the work of my Master's thesis in this outstanding group. I truly appreciate his critical comments and guidance on my thesis. All the experiences in Prof. Teeri's group are absolutely valuable for my study and life in the future.

Secondly, I deeply appreciate many thanks to Dr. Milla Pietiäinen. Her supports and knowledge on the yeast work are truly helpful for me. In the thesis writing period, her critical comments and patient guidance led me to finish my thesis. I express my thanks to Hany El Sayed, the supervisor of my practical training and advanced training, who gave me a lot of suggestions on my study and life. Thanks also Eija Takala for the help of the yeast two-hybrid work.

Third, I am grateful to all my friends in Helsinki: Kai Cheong, Dina, Jing, Tanja etc. Without all your companions in Helsinki, I would never have so that many happy memories.

Last but not the least, I am thankful to my parents Juqiu Xiang and Xiuli Wen. They support and love me always and forever.

Helsinki, September 2014

Jiale Xiang

9. REFERENCES

- Aoki, T., Akashi, T. and Ayabe, S. 2000. Flavonoids of leguminous plants: structure, biological activity, and biosynthesis. *Journal of Plant Research* 113: 475–488.
- Bell-Lelong, D.A., Cusumano, J.C., Meyer, K. and Chapple, C. 1997. Cinnamate-4-hydroxylase expression in *Arabidopsis*. Regulation in response to development and the environment. *Plant Physiology* 113: 729–738.
- Berggård, T., Linse, S. and James, P. 2007. Methods for the detection and analysis of protein-protein interactions. *Proteomics* 7: 2833-2842.
- Bernstein, D.S., Buter, N., Stumpf, C. & Wickens, M. 2002. Analyzing mRNA–protein complexes using a yeast three-hybrid system. *Methods* 26: 123-141.
- Brown, D.E., Rashotte, A.M., Murphy, A.S., Normanly, J., Tague, B.W., Peer, W.A., Taiz, L. and Muday, G.K. 2001. Flavonoids act as negative regulators of auxin transport in vivo in *Arabidopsis*. *Plant physiology* 126: 524-535
- Brückner, A., Polge, C., Lentze, N., Auerbach, D. and Schlattner, U. 2009. Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Science* 10: 2763-2788.
- Brugliera, F., Barri-Rewell, G., Holton, T. and Mason, J.G. 1999. Isolation and characterization of a flavonoid 3'-hydroxylase cDNA clone corresponding to the *Ht1* locus of *Petunia hybrida*. *The Plant Journal* 19: 441-451.
- Buchweitz, D., Brauch, J., Carle, R. and Kammerer, D.R. 2013. Application of ferric anthocyanin chelates as natural blue food colorants in polysaccharide and gelatin based gels. *Food Research International* 51: 274-282.
- Burbulis, I.E. and Winkel-Shirley, B. 1999. Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. *Proceedings of the National Academy of Sciences* 96: 12929-12934.
- Butelli, E., Titta, L., Giorgio, M., Mock, H., Matros, A., Peterek, S., Schijlen, E.G.W.M., Hall, R.D., Bovy, A.G., Luo, J. and Martin, C. 2008. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nature Biotechnology* 26: 1301-1308.

- Castellarin, S., Pfeiffer, A., Sivilotti, P., Degan, M., Peterlunger, E. and Gaspero, G.D. 2007. Transcriptional regulation of anthocyanin biosynthesis in ripening fruits of grapevine under seasonal water deficit. *Plant, Cell and Environment* 30: 1381-1399.
- Christie, P.J., Alfenito, M.R. and Walbot, V. 1994. Impact of low-temperature stress on general phenylpropanoid and anthocyanin pathway: enhancement of transcript abundance and anthocyanin pigmentation in maize seedlings. *Planta* 194: 541-549.
- Cochrane, F.C., Davin, L.B., Lewis, N.G. 2004. The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Phytochemistry* 65: 1557-1564.
- Coskun, O., Kanter, M., Korkmaz, A., and Oter, S. 2005. Quercetin, a flavonoid antioxidant, prevents and protects streptozotocin-induced oxidative stress and β -cell damage in rat pancreas. *Pharmacological Research* 51: 117-123.
- Deng, X., Bashandy, H., Ainasoja, M., Kontturi, J., Pietiäinen, M., Laitinen, R.A.E., Albert, V., Valkonen, J.P.T., Elomaa, P. and Teeri, T.H. 2014. Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of *Gerbera hybrida*. *New Phytologist* 201: 1469-1483.
- Deplancke, B., Dupuy, D., Vidal, M. and Walhout, A.J.M. 2004. A gateway-compatible yeast one-hybrid system. *Genome Research* 14: 2093-2101.
- Du, W., Vidal, M., Xie, J.E. and Dyson, N. 1996. RBF, a novel RB-related gene that regulates E2F activity and interacts with cyclin E in Drosophila. *Genes & Development* 10: 1206-1218.
- Durfee, T., Becherer, K., Chen, P.L., Yeh, S.H., Yang, Y., Kilburn, A.E., Lee, W.H. and Elledge, S. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes & Development* 7: 555-569.
- Dykes, L., Seitz, L.M., Rooney, W.L. and Rooney, L.W. 2009. Flavonoid composition of red sorghum genotype. *Food Chemistry* 116: 313-317.
- Elomaa, P., Honkanen, J., Puska, R., Seppänen, P., Helariutta, Y., Mehto, M., Kotilainen, M., Nevalainen, L. and Teeri, T.H. 1993. *Agrobacterium-*

- Mediated Transfer of Antisense Chalcone Synthase cDNA to *Gerbera hybrida* Inhibits Flower Pigmentation. *Nature Biotechnology* 11: 508-511.
- Ehltling, J., Buttner, D., Wang, Q., Douglas, C.J., Somssich, I.E., Kombrink, E. 1999. Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *The Plant Journal* 19: 9-20.
- Espín, J.C., Soler-Rivas, C., Wichers, H.J. and García-Viguera, C. 2000. Anthocyanin-based natural colorants: A new source of anti-radical activity for foodstuff. *Journal of Agricultural and food chemistry* 48: 1588-1592.
- Feinbaum, R.L. and Ausubel, F.M. 1988. Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. *Molecular and Cellular Biology* 8: 1985-1992.
- Feinbaum, R.L., Storz, G. and Ausubel, F.M. 1991. High intensity and blue light regulated expression of chimeric chalcone synthase genes in transgenic *Arabidopsis thaliana* plants. *Molecular and General Genetics* 226: 449-456.
- Ferrer, J.L., Jez, J.M., Bowman, M.E., Dixon, R.A. and Noel, J.P. 1999. Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nature Structural Biology* 6: 775-784.
- Firestine, S.M., Salinas, F., Nixon, A.E., Baker, S.J. & Benkovic, S.J. 2000. Using an AraC-based three-hybrid system to detect biocatalysts in vivo. *Nature Biotechnology* 18: 544-547.
- Forkmann, G. and Ruhnau, B. 1987. Distinct substrate specificity of dihydroflavonol 4-reductase from flowers of *Petunia hybrida*. *Z Naturforsch* 42C: 1146-1148
- Gitz, D.C., Liu, L., McClure, J.W. 1998. Phenolic metabolism, growth and UV-B tolerance in phenylalanine ammonia lyase inhibited red cabbage seedlings. *Phytochemistry* 49: 377-386.
- Hamberger, B. and Hahlbrock, K. 2004. The *4-coumarate:CoA ligase* gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proceedings of the National Academy of Sciences* 101: 2209-2214.
- Harborne, J.B., Williams, C.A. 2000. Advanced in flavonoid research since 1992. *Phytochemistry* 55: 481-504.

- Halbwirth, H., Martens, S., Wienand, U., Forkmann, G., Stich, K. 2003. Biochemical formation of anthocyanins in silk tissue of *Zea mays*. *Plant Science* 164: 489-495.
- He, J. and Giusti, M.M. 2010. Anthocyanins: Natural colorants with health-promoting properties. *Food Science and Technology* 1: 163-187.
- Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R.J., Schröder, J. and Teeri, T.H. 1995. Chalcone synthase-like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera Hybrida* (Asteraceae). *Plant Molecular Biology* 28: 47-60.
- Holton, T.A. and Cornish, E.C. 1995. Genetics and biochemistry of anthocyanin biosynthesis. *The Plant Cell* 7: 1071-1083.
- Hrazdina, G. & Wagner, G.J. 1985a. Metabolic pathways as enzyme complexes: Evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. *Archives of Biochemistry and Biophysics* 237: 88-100.
- Hrazdina, G. and Wagner, G.J. 1985b. Compartmentation of plant phenolic compounds; sites of synthesis and accumulation. *Annual Proceedings of the Phytochemical Society of Europe* 25: 120-133.
- Hrazdina, G., Zobel, A.M. and Hoch, H.C. 1987. Biochemical, immunological, and immunocytochemical evidence for the association of chalcone synthase with endoplasmic reticulum membranes. *Proceedings of the National Academy of Sciences* 84: 8966-8970.
- Hu, G., Zhang, S., Vidal, M., La Baer, J., Xu, T. and Fearon, E.R. 1997. Mammalian homologs of seven in absentia regulate DCC via the ubiquitin-proteasome pathway. *Gene & Development* 11: 2701-2714.
- Huang, J., Gu, M., Lai, Z., Fan, B., Shi, K., Zhou, Y.H., Yu, J.Q. and Chen, Z. 2010. Functional analysis of the Arabidopsis *PAL* gene family in plant growth, development, and response to environmental stress. *Plant Physiology* 153: 1526-1538.
- Huguene, P., Provenzano, S., Verries, C., Ferrandino, A., Meudec, E., Batelli, G., Merdinoglu, D., Cheynier, V., Schubert, A. and Ageorges, A. 2009. A novel cation-dependent O-methyltransferase involved in anthocyanin methylation in grapevine. *Plant Physiology* 150: 2057-2070.

- Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. and Sakaki, Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences* 98: 4569-4574.
- Ito, T., Ota, K., Kubota, H., Yamaguchi, Y., Chiba, T., Sakuraba, K. and Yoshida, M. 2002. Roles for the two-hybrid system in exploration of the yeast protein interactome. *Molecular & Cellular Proteomics* 1: 561-566.
- Iwashina, T. 2000. The structure and distribution of the flavonoids in plants. *Journal of Plant Research* 113: 287-299.
- Iwashina, T. 2003. Flavonoid function and activity to plants and other organisms. *Biological Sciences in Space* 17: 24-44.
- Jayaprakasam, B., Vareed, S.K., Olson, L.K. and Nair, M.G. 2005. Insulin secretion by bioactive anthocyanins and anthocyanidins present in fruits. *Journal of Agricultural and Food Chemistry* 53: 28-31.
- Johnson, E.T., Ryu, S., Yi, H., Shin, B., Cheong, H. and Choi, G. 2001. Alternation of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. *The Plant Journal* 25: 325-333.
- Johnsson, N. & Varshavsky, A. 1994. Split ubiquitin as a sensor of protein interactions *in vivo*. *Proceedings of the National Academy of Sciences* 91: 10340-10344.
- Jørgensen, K., Rasmussen, A.V., Morant, M., Nielsen, A.H., Bjarnholt, N., Zagrobelny, M., Bak, S. and Møller, L. 2005. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology* 8: 280-291.
- Jung, W., Yu, O., Lau, S.M.C., O'Keefe, D.P., Odell, J., Fader, G. and McGonigle, B. 2000. Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nature Biotechnology* 18: 208-212.
- Katsumoto, Y., Mizutani, M.F., Fukui, Y., Brugliera, F., Holton, T.A., Karan, M., Nakamura, N., Sakakibara, K.Y., Togami, J., Pigeaire, A., Tao, G.Q., Nehra, N.S., Lu, C.Y., Dyson, B.K., Tsuda, S., Ashikari, T., Kusumi, T., Mason, J.G. and Tanaka, Y. 2007. Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating delphinidin. *Plant Cell Physiology* 48: 1589-1600.

- Keegan, L., Gill, G., Ptashne, M. 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* 231: 699-704.
- Knekt, P., Kumpulainen, J., Järvinen, R., Rissanen, H., Heliövaara, M., Reunanen, A., Hakulinen, T. and Aromaa, A. 2002. Flavonoid intake and risk of chronic diseases. *The American Journal of Clinical Nutrition* 76: 560-568.
- Koes, R., Verweij, W., Quattrocchio, F. 2005. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science* 10: 236-242.
- Koostra, A. 1994. Protection from UV-B-induced DNA damage by flavonoids. *Plant Molecular Biology* 26: 771-774.
- Lee, D., Ellard, M., Wanner, L.A., Davis, K.R. and Douglas, C.J. 1995. The *Arabidopsis thaliana* 4-coumarate: CoA ligase (*4CL*) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA. *Plant Molecular Biology* 28: 871–884.
- Li, H., Deng, Z., Zhu, H., Hu, C., Liu, R., Young, J.C. and Tsao, R. 2012. Highly pigmented vegetables: Anthocyanin compositions and their role in antioxidant activities 46: 250-259.
- Li, J.J. & Herskowitz, I. 1993. Isolation of ORC6, a component of the yeast origin recognition complex by a One-Hybrid System. *Science* 262: 1870-1874.
- Li, J., Ou-Lee, T.M., Raba, R., Amundson, R.G. & Last, R.L. 1993. *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. *The Plant Cell* 5: 171-179.
- Licitra, E.J. and Liu, J.O. 1996. A three-hybrid system for detecting small ligand–protein receptor interactions. *Proceedings of the National Academy of Sciences* 93: 12817-12821.
- Lin, Y.M., Anderson, H., Flavin, M.T. and Pai, Y.H.S. 1997. In vitro anti-HIV activity of biflavonoids isolated from *Rhus succedanea* and *Garcinia multiflora*. *Journal of Natural Products* 60: 884-888.
- Lorenc-Kukula, K., Jafra, S., Oszmianski, J. and Szopa, J. 2005. Ectopic expression of anthocyanin 5-o-glucosyltransferase in potato tuber causes

increased resistance to bacteria. *Journal of Agricultural Food Chemistry* 53: 272-281.

- Ma, J. and Ptashne, M. 1987. A new class of yeast transcriptional activator. *Cell* 51: 113-119.
- Matthews, L. R., Vaglio, P., Reboul, J., Ge, H., Davis, B. P., Garrels, J., Vincent, S., and Vidal, M. 2001. Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or "interologs." *Genome Research* 11: 1771-1775
- Meng, X., Brodsky, M.H. and Wolfe, S.A. 2005. A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nature Biotechnology* 23: 988-994.
- McCusker, J.H. and Haber, J.E. 1990. Mutations in *Saccharomyces cerevisiae* which confer resistance to several amino acid analogs. *Molecular and Cellular Biology* 10: 2941-2949.
- Mol, J., Cornish, E., Mason, J. and Koes, R. 1999. Novel coloured flowers. *Current Opinion in Biotechnology* 10: 198-201.
- Ohl, S., Hedrick, S.A., Chory, J., Lamb, C.J. 1990. Functional properties of a phenylalanine ammonia-lyase promoter from *Arabidopsis*. *The Plant Cell*: 2: 837-847.
- Olsen, K.M., Lea, U.S., Sliemstad, R., Verheul, M. and Lillo, C. 2008. Differential expression of four *Arabidopsis* PAL genes; *PAL1* and *PAL2* have functional specialization in abiotic environmental-triggered flavonoid synthesis. *Journal of Plant Physiology* 165: 1491-1499.
- Olsson, L.C., Veit, M., Weissenböck, G., Bornman, J.F. 1998. Flavonoid response to UV-B radiation in *Brassica napus*. *Phytochemistry* 49: 1021-1028.
- Ovadi, J. and Srere, P.A. 2000. Macromolecular compartmentation and channeling. *International Review of Cytology* 192: 255-280.
- Owens, D.K., Alerding, A.B., Crosby, K.C., Bandara, A.B., Westwood, J.H. and Winkel, B.S.J. 2008a. Functional analysis of a predicted flavonoid synthase gene family in *Arabidopsis*. *Plant Physiology* 147: 1046-1061.

- Owens, D.K., Crosby, K.C., Runac, J., Howard, B.A. and Winkel, B.S.J. 2008b. Biochemical and genetic characterization of *Arabidopsis* flavanone 3 β -hydroxylase. *Plant Physiology and Biochemistry* 46: 833-843.
- Peer, W.A., Brown, D.E., Tague, B.W., Muday, G.K., Taiz, L. and Murphy, A.S. 2001. Flavonoid accumulation patterns of transparent testa mutants of *Arabidopsis thaliana*, *Plant Physiology* 126: 536-548.
- Pelletier, M.L. and Shirley, B.W. 1996. Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings. Coordinate regulation with chalcone synthase and chalcone isomerase. *Plant Physiology* 111: 339-345.
- Pelletier, M.K., Murrell, J.R. and Shirley, B.W. 1997. Characterization of flavonol synthase and leucoanthocyanidin dioxygenase genes in *Arabidopsis*. Further evidence for differential regulation of "early" and "late" genes. *Plant Physiology* 113: 1437-1445.
- Preuß, A., Stracke, R., Weisshaar, B., Hillebrecht, A., Matern, U. and Martens, S. 2009. *Arabidopsis thaliana* expresses a second functional flavonol synthase. *FEBS letters* 583: 1981-1986.
- Provenzano, S., Spelt, C., Hosokawa, S., Nkamura, N., Brugliera, F., Demelis, L., Geerke, D.P., Schubert, A., Tanaka, Y., Quattrocchio, F. and Koes, R. 2014. Genetics and evolution of anthocyanin methylation. *Plant Physiology* 113.
- Raes, J., Rohde, A., Christensen, J.H., De Peer, Y.V. and Boerjan, W. 2003. Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiology* 133: 1051-1071.
- Reuber, S., Bornman, J.F., Weissenböck, G. 1996. A flavonoid mutant of barley (*Hordeum Vulgare* L.) exhibits increased sensitivity to UV-B radiation in primary leaves. *Plant Cell Environment* 19: 593-599.
- Rohde, A., Morreel, K., Ralph, J., Goeminne, G., Hostyn, V., De Rycke, R., Kushnir, S., Doorselaere, J.V., Joseleau, J.P., Vuylsteke, M., Driessche, G.V., Beeumen, J.V., Messens, E. and Boerjan, W. 2004. Molecular phenotyping of the *pal1* and *pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. *The Plant Cell* 16: 2749-2771.

- Ritter, H. and Schulz, G.E. 2004. Structural Basis for the Entrance into the Phenylpropanoid Metabolism Catalyzed by Phenylalanine Ammonia-Lyase. *The Plant Cell* 16: 3426-3436.
- Saito, K., Kobayashi, M., Gong, Z.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*. *The Plant Journal* 17: 181-189.
- Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tokayuki, T. and Fernie, A.R. 2013. The flavonoid biosynthetic pathway in *Arabidopsis*: Structural and genetic diversity. *Plant Physiology and Biochemistry* 72: 21-34.
- Saslowsky, D. and Winkel-Shirley, B. 2001. Localization of flavonoid enzymes in *Arabidopsis* root. *The Plant Journal* 27: 37-48.
- Saslowsky, D.E., Warek, U. and Winkel, B.S.J. 2005. Nuclear localization of flavonoid enzymes in *Arabidopsis*. *The Journal of Biological Chemistry* 280: 23735-23740.
- Schillmiller, A.L., Stout, J., Weng, J.K., Humphreys, J., Ruegger, M.O. and Chapple, C. 2009. Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in *Arabidopsis*. *The Plant Journal* 60: 771-782.
- Schoenbohm, C., Martens, S., Eder, C., Forkmann, G. and Weisshaar, B. 2000. Identification of the *Arabidopsis thaliana* flavonoid 3'-hydroxylase gene and functional expression of the encoded P450 enzyme. *Biological Chemistry* 381: 749-753.
- SenGupta, D.J., Zhang, B., Kraemer, B., Pochart, P., Fields, S. and Wichens, M. 1996. A three-hybrid system to detect RNA-Protein interactions in vivo. *Proceedings of the national Academy of Sciences* 96: 8496-8501.
- Serebriiskii, I., Estojak, J., Berman, M. and Golemis, E.A. 2000. Approaches to detecting false positives in yeast two-hybrid systems. *BioTechniques* 28: 328-336.
- Shirley, B.W., Hanley, S. and Goodman, H.M. 1992. Effects of ionizing radiation on a plant genome: Analysis of two *Arabidopsis transparent testa* mutations. *The Plant Cell* 4: 333-347.

- Shirley, B.W., Kubasek, W.L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F.M. and Goodman, H.M. 1995. Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. *The Plant Journal* 8: 659-671.
- Stagljar, I., Korostensky, C., Johnsson, N. and Te Heesen, S. 1998. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins *in vivo*. *Proceedings of the National Academy of Sciences* 95: 5187-5192.
- Stafford, H.A. 1974. Possible multi-enzyme complexes regulating the formation of C6-C3 phenolic compounds and lignins in higher plants. *Recent Advances in Phytochemistry* 8: 53–79.
- Stapleton, A.E. & Walbot, V. 1994. Flavonoids can protect maize DNA from the induction of ultraviolet radiation damage. *Plant Physiology* 105: 881-889.
- Stevenson, P.C., Haware, M.P. 1999. Maackiain in *Cicer bijungum* associated with resistance to Botrytis. *Biochemical Systematics and Ecology* 27: 761-767.
- Stracke, R., De Vos, R.C., Bartelniewoehner, L., Ishihara, H., Sagasser, M., Martens, S., Weisshaar, B. 2009. Metabolomic and genetic analyses of flavonol synthesis in Arabidopsis thaliana support the *in vivo* involvement of leucoanthocyanidin dioxygenase. *Planta* 229: 427-445.
- Taylor, L.P., Grotewold, E. 2005. Flavonoids as developmental regulators. *Current Opinion in Plant Biology* 8: 317–323.
- Tohge, T., Nishiyama, Y., Hirai, M.Y., Yano, M., Nakajima, J., Awazuhara, M., Inoue, E., Takahashi, H., Goodenowe, D.B., 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 Journal* 42: 218-235.
- Treutter, D. 2006. Significance of flavonoids in plant resistance: a review. *Environmental Chemistry Letter* 4: 147-157.
- Vidal, M., Brachmann, R., Fattaey, A., Harlow, E. and Boeke, J.D. 1996. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proceedings of the National Academy of Sciences* 93: 10315-10320.
- Vidal, M., Legrain, P. 1999. Yeast forward and reverse 'n'-hybrid systems. *Nucleic Acids Research* 27: 919-929.

- Vidalain, P.O., Boxem, M., Ge, H., Li, S. and Vidal, M. 2004. Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods* 32: 363-370.
- Wagner, G. and Hrazdina, G. 1984. Endoplasmic reticulum as a site of phenylpropanoid and flavonoid metabolism in *Hippeastrum*. *Plant Physiology* 74: 901-906.
- Walhout, A.J.M. and Vidal, M. 1999. A Genetic strategy to eliminate self-activator baits prior to high-throughput yeast two-hybrid screens. *Genome Research* 9: 1128-1134.
- Walhout, A.J.M. and Vidal, M. 2001. High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. *Methods* 24: 297-306.
- Wanner, L.A., Li, G., Ware, D., Somssich, I.E., Davis, K.R. 1995. The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. *Plant Molecular Biology* 27: 327-338.
- Wilmouth, R.C., Turnbull, J.J., Welford, R.W.D., Clifton, I.J., Prescott, A.G. and Schofield, C.J. 2002. Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. *Structure* 10: 93-103.
- Winefield, C.S., Lewis, D.H., Swinny, E.E., Zhang, H., Arathoon, H.S., Fischer, T.C., Halbwirth, H., Stich, K., Gosch, C., Forkmann, G. and Davies, K.M. 2005. Investigation of the biosynthesis of 3-deoxyanthocyanins in *Sinningia cardinalis*. *Physiologia Plantarum* 124: 419-430.
- Winkel-Shirley, B. 1999. Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. *Physiologia Plantarum* 107: 142-149.
- Winkel-Shirley, B. 2001. Flavonoid Biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* 126: 485-493.
- Yu, O., Jung, W., Shi, J., Croes, R.A., Fader, G., McGonigle, B. and Odell, J.T. 2000. Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant Physiology* 124: 781-794
- Zhu, G., Li, C. and Cao, Z. 2007. Inhibitory effect of flavonoid baicalin on degranulation of human polymorphonuclear leukocytes induced by interleukin-8: Potential role in periodontal diseases. *Journal of Ethnopharmacology* 109: 325-330.

10. APPENDICES

Appendix 1: Medium used

YPAD plates and medium:

10 g Yeast extract
20 g bacto peptone
20 g agar (no agar for medium)
20 g glucose
0.1 g adenine hemi sulfate
MQ water to 1000 ml
autoclave

SD-trp-leu:

20 g bacto agar
6.7 g yeast nitrogen base without amino acids
800 ml MQ
100 ml appropriate 10x amino acid stock solution (contains all other amino acids except trp and leu)
adjust pH to 5.8 with NaOH
autoclave
Add 100 ml 20 % glucose

SD-leu-trp-his-ade:

20 g bacto agar
6.7 g yeast nitrogen base without amino acids
800 ml MQ
100 ml appropriate 10x amino acid stock solution (contains all other amino acids except leu, trp, his and ade)
adjust pH to 5.8 with NaOH
autoclave
Add 100 ml 20 % glucose

10X SD amino acids 600 ml

0.12 g L-Adenine hemi sulfat salt
0.12 g L-arginine
0.12 g L-histidine
0.18 g L-isoleucin
0.6 g L-leucin
0.18 g L-lycine
0.12 g L-methionine
0.3 g L-phenylalanine
1.2 g L-threonine
0.12 g L-tryptophane
0.18 g L-tyrosine
0.12 g L-urasil
0.9 g L-valine
Add 600 MilliQ water

Appendix 2: Agarose gel for checking the quality of RNA

Agarose 1%
TBE buffer 0.5X
Loading sample: 1 μ l RNA, 5 μ l loading dye (5X).
Running: 100 Volts for 20 min

Appendix 3: Agarose gel for PCR products

Agarose 1%
TBE buffer 0.5X
Loading sample: 2 μ l PCR products, 5 μ l loading dye (5X).
Running: 150 Volts for 40 min

Appendix 4: Details and sequence of primers used for PCR

Gene name	Sequences (5-3)	Annealing T (°C)	Expected size of product (bp)
<i>PAL1</i>	AtPAL1_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GAG ATT AAC GGG GCA C	64	2178
	AtPAL1_R (Prey): AAC TGA TT GGC CGA GGC GGC C TTAACA TAT TGG AAT GGG AGC TC	63.1	
	AtPAL1_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC ACA TAT TGG AAT GGG AGC TCC	64.2	
<i>C4H</i>	AtC4H_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GAC CTC CTC TTG CTG G	64.7	1518
	AtC4H_R (Prey & Bait): AAC TGA TT GGC CGA GGC GGC C CC ACA GTT CCT TGG TTT CAT AAC GA	63.8	
<i>CHS</i>	AtCHS_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GTG ATG GCT GGT GCT Tc	67.7	1188
	AtCHS_R (Prey): AAC TGA TT GGC CGA GGC GGC C TTA GAG AGG AAC GCT GTG CAA g	65.9	
	AtCHS_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC GAG AGG AAC GCT GTG CAA ga	66.1	
<i>CHI</i>	AtCHI_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATGTCTTCATCCAACGCCTG	65.4	741
	AtCHI_R (Prey): AAC TGA TT GGC CGA GGC GGC C TCAGTTCTCTTTGGCTAGTTTTTCC	64.8	
	AtCHI_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC GTTCTCTTTGGCTAGTTTTTCCTCA	64.8	
<i>F3H</i>	AtF3H_ (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GCT CCA GGA ACT TTG ACT	64.2	1077
	AtF3H_R (Prey): AAC TGA TT GGC CGA GGC GGC C CTAAGC GAA GAT TTG GTC GAC	63	
	AtF3H_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC AGC GAA GAT TTG GTC GAC AG	64.7	

<i>DFR</i>	AtDFR_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GTT AGT CAG AAA GAG ACC GTG	64.7	1149
	AtDFR_R (Prey): AAC TGA TT GGC CGA GGC GGC C CTA GGC ACA CAT CTG TTG TGC	63.8	
	AtDFR_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC GGC ACA CAT CTG TTG TGC T	63.1	
<i>F3'H</i>	AtF3'H_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GCA ACT CTA TTT CTC ACA ATC CTC	67.3	1542
	AtF3'H_R (Prey & Bait): AAC TGA TT GGC CGA GGC GGC C CC ACC CGA CCC GAG TCC ATA AAC	68.3	
<i>FLS1</i>	AtFLS1_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GAG GTC GAA AGA GTC CAA	65.2	1011
	AtFLS1_R (Prey): AAC TGA TT GGC CGA GGC GGC C TCA ATC CAG AGG AAG TTT ATT GAG C	65.6	
	AtFLS1_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC ATC CAG AGG AAG TTT ATT GAG CTT G	64.7	
<i>ANS</i>	AtANS_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GTT GCG GTT GAA AGA GT	63.8	1071
	AtANS_R (Prey): AAC TGA TT GGC CGA GGC GGC C TTA ATC ATT TTT CTC GGA TAC CAA TTC	64.8	
	AtANS_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC ATC ATT TTT CTC GGA TAC CAA TTC	63.2	
<i>GT</i>	AtGT_F (Prey & Bait): ATT AAC AA GGC CAT TAC GGC C ATG GAG CAT AAG AGA GGA CAT GTA	63	1350
	AtGT_R (Prey): AAC TGA TT GGC CGA GGC GGC C CTA TTT GCT CTG AAC CCT TGA TA C A	64.4	
	AtGT_R (Bait): AAC TGA TT GGC CGA GGC GGC C CC TTT GCT CTG AAC CCT TGA TAC A	63.8	

Appendix 5: Primers and sequences for colony PCR and sequencing

Vectors' name	Primer Sequences
pPR3-N	Forward: GTCGAAAATTCAAGACAAGG Reverse: AAGCGTGACATAACTAATTAC
pPR3-SUC	Forward: TTTCTGCACAATATTTCAAGC Reverse: CTTGACGAAAATCTGCATGG
pDHB1	Forward: TTTCTGCACAATATTTCAAGC Reverse: GTAAGGTGGACTCCTTCT
pBT3-SUC	Forward: TGGCATGCATGTGCTCTG Reverse: GTAAGGTGGACTCCTTCT