6-28-2017 12:00 AM

GmMYB176 Interactome and Regulation of Isoflavonoid Biosynthesis in Soybean

Arun Kumaran Anguraj Vadivel
The University of Western Ontario

Supervisor
Dr. Sangeeta Dhaubhadel
The University of Western Ontario Joint Supervisor
Dr. Mark Bernards
The University of Western Ontario

Graduate Program in Biology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Abstract

MYB transcription factors are one of the largest transcription factor families characterized in plants. They are classified into four types: R1 MYB, R2R3 MYB, R3 MYB and R4 MYB. GmMYB176 is an R1MYB transcription factor that regulates *Chalcone synthase (CHS8)* gene expression and isoflavonoid biosynthesis in soybean. Silencing of *GmMYB176* suppressed the expression of the *GmCHS8* gene and reduced the accumulation of isoflavonoids in soybean hairy roots. However, overexpression of *GmMYB176* does not alter either *GmCHS8* gene expression or isoflavonoid levels suggesting that GmMYB176 alone is not sufficient for *GmCHS8* gene regulation. I hypothesized that GmMYB176 acts cooperatively with another factor(s) for the regulation of *GmCHS8* gene expression and it may also regulate other isoflavonoid biosynthetic genes in soybean. The objective of this research was to identify the GmMYB176 interactome for *GmCHS8* gene regulation and elucidate the role of GmMYB176 in isoflavonoid biosynthesis in soybean. GmMYB176 interacting proteins were identified using two translational fusion baits (GmMYB176-YFP and YFP-GmMYB176) by co-immunoprecipitation, followed by liquid chromatography-tandem mass spectrometry. The interaction of selected candidates with GmMYB176 was validated *in planta* and their DNA binding activities determined. GmMYB176 may form a transcriptional complex with Gm04bZIP and/or Gm05bZIP for the regulation of *GmCHS8* gene expression. RNA-seq and metabolomics analyses of soybean hairy roots in which *GmMYB176* was either silenced or over-expressed revealed that GmMYB176 regulates multiple genes in the isoflavonoid biosynthesis pathway, affecting the production of metabolites in phenylpropanoid pathway such as phenylalanine, liquiritigenin, daidzin, genistin, glycitein, and glyceollin. The knowledge and information generated on the role of GmMYB176 in the regulation of isoflavonoid biosynthesis will allow genetic manipulation of isoflavonoid level in soybean and/or introduce isoflavonoid pathway in non-legumes.

Keywords

Co-immunoprecipitation, LC-MS/MS, MYB, protein-protein interaction, transcriptomics, metabolomics, isoflavonoid, phenylpropanoid, secondary metabolism, transcriptional complex, gene regulation, gene families, soybean

Acknowledgements

I dedicate this thesis to my non-biological father, Mr. N. Thangavel. Thanks as well to my parents Mr. Anguraj Vadivel and Mrs. Thilagavathy, my sister, Mrs. Priyadharshini for their support in me pursue a research in Canada, far away from home. I also thank my cousins Ms. Kamalapriya and Mr. Suryakumar for their support.

I would like to thank my adopted family at Agriculture and Agri-Food Canada (AAFC), London. Everyone in the Dhaubhadel lab (past and present), Mehran, Hemanta, Kishor, Pravesh, Arjun, Caroline, Xuyan, Shaomin, Serina, Amelia, Jaeju, Kevin, Kelsey, Smarth, and Nishat – thank you for your support and encouragement. I would also like to thank Ling, for her lab technical advice and life advice. Moreover, I would like to thank all my friends in other labs (in AAFC) FM lab, CD lab, IS lab, AH lab, RM lab, RA lab, YC lab (and so many) for lighting up my days and for all those memories in last five years. I also thank my friends, Aagam, Milan, Preetam and Rohan from dept of chemistry, biochemistry, and geology.

I must thank the individuals who have been essential to the present research: my advisory committee, Dr. Jim Karagiannis and Dr. Frederic Marsolais for their input and guidance; I also thank my co-supervisor, Dr. Mark Bernards for his insights and his invaluable advice for my research and in general. I would like to thank Dr. Justin Renaud for his help in analyzing metabolite samples and teaching me a lot about metabolomics. I also thank London Regional Proteomics Center, UWO for their help in analyzing my protein samples. Most importantly, I must thank my supervisor, Dr. Sangeeta Dhaubhadel for giving me the opportunity to join her lab and for her guidance, support, and encouragement in science.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

2-HID	2-HydroxyIsoflavanone Dehydratase
3RT	Anthocyanidin 3-glucoside rhamnosyltransferase
5GT	Anthocyanin 5-*O*-glucosyltransferase
ABC	ATP Binding Cassette
ADT	Arogenate Dehydratase
At	*Arabidopsis thaliana*
bHLH	Basic Helix-loop-helix
BiFC	Bimolecular Fluorescence Complementation
BLAST	Basic Local Alignment Search Tool
BLASTN	Nucleotide BLAST
bZIP	Basic leucine zipper
C4H	Cinnamate-4-Hydroxylase
CHI	Chalcone Isomerase
CHR	Chalcone Reductase
CHS	Chalcone Synthase
Co-IP	Co-Immunoprecipitation
DAP	Days After Pollination
DEG	Differentially Expressed Gene
DESeq	Differential Expression Sequencing
DFR	Dihydroflavonol 4-Reductase
DNA	Deoxyribonucleic Acid
DoF	DNA binding with One Finger
ER	Endoplasmic Reticulum
F3'5'H	Flavonoid 3',5'-Hydroxylase
F3'H	Flavonoid 3'-Hydroxylase
F3H	Flavanone 3-Hydroxylase
F6H	Flavonoid 6-Hydroxylase
FAP	Fatty Acid-binding Protein
FLS	Flavonol Synthase

FRET	Förster Resonance Energy Transfer
GFP	Green Fluorescent Protein
Gm	*Glycine max*
GO	Gene Ontology Database
GST	Glutathione S-Transferase
GUS	β-glucuronidase
HIV	Human Immunodeficiency Virus
HPLC	High-Performance Liquid Chromatography
HRP	Horseradish Peroxidase
I2'H	Isoflavone 2-Hydroxylase
IFR	Isoflavone Reductase
IFS	Isoflavone Synthase
IOMT	Isoflavone O-methyltransferase
LB	Luria-Bertani Medium
LC-MS/MS	Liquid Chromatography-tandem Mass Spectrometry
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization Time-of-Flight
MS	Mass Spectrometry
Ms	*Medicago sativa*
MYB	Myeloblastosis
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NB	Nucleotide-binding
NCBI	National Centre for Biotechnology Information
Nod	Nodulation Factor
P450	Cytochrome P450
PAL	Phenylalanine Ammonia-lyase
PVDF	Polyvinylidene Difluoride
qPCR	Quantitative Polymerase Chain Reaction
RNA	Ribonucleic Acid
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
TAIR	The Arabidopsis Information Resource
TT8	Transparent Testa 8

UGT	UDP glucuronosyltransferases
UPLC	Ultra-Performance Liquid Chromatography
UV	Ultra-violet
YC	C-Terminal Fragment of YFP
YFP	Yellow Fluorescent Protein
YN	N-Terminal Fragment of YFP
Zm	*Zea mays*

1 GENERAL INTRODUCTION

1.1 Soybean

Soybean (*Glycine max* L. Merr) is a legume native to East Asia. It was domesticated between 7000 and 6600 before Common Era (BCE) in China, between 5000 and 3000 BCE in Japan and 1000 BCE in Korea (Lee et al., 2011). Soybean is one of the most important legumes in the world. Soybean seeds are comprised of ~40% protein and 19% oil content and contributes as a cheap source of proteins to 2/3 of global livestock feed and 25% of the global edible oil (Agarwal et al., 2013). Soybeans produce significantly more protein per acre than most other land plants and are major crops in the United States, Brazil, Argentina, China, and India. Global production of soybeans is estimated to be 338 million metric tons in the year 2016/2017 by the United States Department of Agriculture (USDA) and Canada is the seventh largest producer of soybean in the world (GlobalSoybeanProduction.com, 2016).

1.1.1 Soybean genome

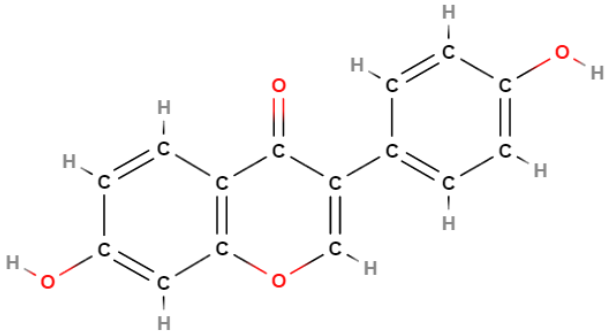
The soybean genome is 1.1 Gb in size and is predicted to contain 56,044 protein-coding loci and 88,647 transcripts (Schmutz et al., 2010; Goodstein et al., 2012). A well-annotated reference genome can boost the functional genomic analyses such as epigenomic, transcriptomic, proteomic, and non-coding RNA analyses. The soybean genome has undergone two whole genome duplications (WGD) at approximately 59 and 13 million years ago, resulting in a highly duplicated genome with nearly 75% of the genes present in multiple copies. Studies have indicated that WGD has occurred multiple times over the past 200 million years of angiosperm evolution resulting into plant genomes with an abundance of duplicate genes (Soltis et al., 2009; Soltis et al., 2014). Genome, tandem and segmental duplication events could be resulted in multiple gene families in the genome of soybean.

1.2 **Phenylpropanoid pathway**

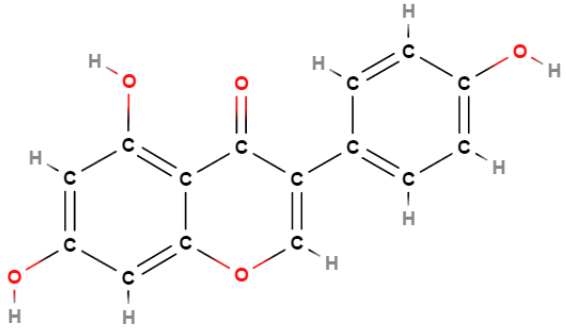
Phenylpropanoids are an important class of plant specialized metabolites such as lignins, flavonoids, and stilbenes, that are widely distributed in the plant kingdom (Jeong et al., 2016). The diversity of phenylpropanoids is the result of efficient modifications of core structures derived from the shikimate pathway (Herrmann, 1995). Phenylalanine, which is one of the end products of the shikimate pathway, is the main precursor for all compounds in the phenylpropanoid pathway, leading to a plethora of different metabolites that are involved in plant structural support, pigmentation, signaling, and attractants for pollinators. They also contribute to all aspects of plant responses towards abiotic stimuli like drought and a variation in light intensity and plant resistance towards biotic stresses (La Camera et al., 2004; Buer et al., 2010; Vogt, 2010).

1.3 **Isoflavonoids**

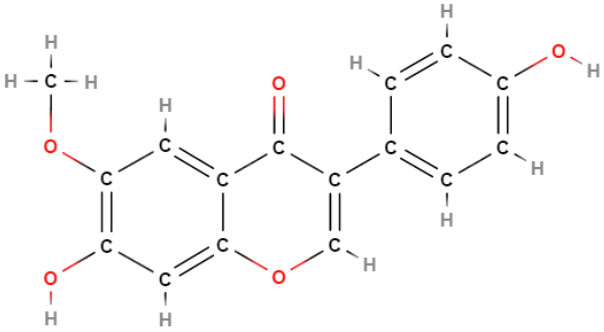
Isoflavonoids are plant natural compounds which constitute a subclass of flavonoids. The core chemical structure of isoflavonoids consists of 2 phenyl rings (A and B) and a heterologous C ring (where the B ring is attached to the C-2 carbon of the C-ring) (Fig 1.1). Isoflavonoids are biologically active metabolites that are known for their pharmacological properties and are produced almost exclusively by leguminous plants (Lapcik, 2007). Isoflavonoids also function as phytoalexins as part of the innate immune system (Dixon and Paiva, 1995). Soybean seeds contain three isoflavone aglycones: genistein, daidzein, and glycitein (Fig 1.1). These aglycones are found in 7-*O*-β-glucosides, 6''-*O*-acetyl-β-glucosides, and 6''-*O*-malonyl-β-glucosides (Kudou et al., 1991). Glycosylation and malonylation of soybean isoflavones enhance their stability and solubility and also promotes their compartmentalization into vacuoles for storage. Mostly, isoflavones are found in the ratio of 4:5:1 (daidzein: genistein: glycitein) in soybean seeds (Wang and Murphy, 1994) and their malonyl derivatives are the most abundant form in soybean. The isoflavonoid pathway is in direct competition with the concurrent flavonoid pathway for flavanone substrates. Manipulation of isoflavonoid biosynthesis or to engineer the pathway in non-legumes has underlined this effective bottleneck to metabolite accumulation (Liu et al., 2002; Yu et al., 2003).



Daidzein



Genistein



Glycitein

Figure 1.1 Chemical structure of three major isoflavones aglycones in soybean. The structures were generated using MolView version 2.4 (http://molview.org/).

1.3.1 **Role of isoflavonoids in human health and nutrition**

In humans, isoflavonoids are noted for several health benefits (Messina, 2010). As part of the human diet, isoflavonoids are associated with a reduction in the risk of cardiovascular disease and hormone-dependent cancers (Dixon and Ferreria, 2002; Chen and Rogan, 2004; Dixon, 2004). Genistein has inhibitory activity against prostate cancer cell metastasis by blocking the mitogen-activated protein kinase p38 (Xu et al., 2009).Genistein is also known to interfere with the activity of a tyrosine kinase that is involved in the actin dynamics of HIV infection (Guo et al., 2013).

Several clinical studies demonstrate the intake of a soy-rich diet is correlated with a reduction in the risk of cardiovascular disease (Hooper et al., 2008), obesity (Orgaard and Jensen, 2008), hormone-dependent cancers (Pollard and Suckow, 2006; Wu et al., 2008; Yan and Spitznagel, 2009), menopausal symptoms (Howes et al., 2006), osteoporosis (Potter et al., 1998), and also reduces total serum cholesterol (Anderson et al., 1995; Zhan and Ho, 2005). Long-term intake of soy-based food is correlated with reduced breast cancer risk in humans (Lamartiniere et al., 1995; Shu et al., 2001; Pisani et al., 2002; Duffy et al., 2007; Korde et al., 2009). Isoflavonoids are also called phytoestrogens since many of these compounds are structurally similar to estradiol and exert biological effects through their ability to bind to estrogen receptors. For example, genistein, which structurally resembles human hormone 17-β-estradiol, has a weak affinity for human estrogen receptors α, β and sex hormone-binding globulins (Barnes et al., 2000; Dixon and Ferreria, 2002). Isoflavones can act as modulators of estrogen receptors to have agonist or antagonist effects (Simons et al., 2012).

1.3.2 **Role of isoflavonoids in plants**

Legumes develop symbioses with nitrogen-fixing rhizobia to utilize atmospheric nitrogen for their growth and development (Mulligan and Long, 1985). In the rhizosphere, isoflavonoids act as signaling molecules that attract symbiotic bacteria to legume roots and induce their nodulation genes, leading to root nodule formation. They are also involved in pathogen inhibition (Phillip, 1992; Phillips and Kapulnik, 1995; Lozovaya et

al., 2004; Subramanian et al., 2006). Rhizobia help plants to consume nitrogen from the atmosphere and reduce the economic burden caused by the use of nitrogenous fertilizers.

Isoflavonoids function as phytoalexins during plant defense mechanisms against pathogens. Phytoalexins are low molecular weight compounds that are induced in response to stress such as pathogen infection. An increased level of isoflavonoids occurs during pathogen attack in soybean (Abbasi and Graham, 2001; Lozovaya et al., 2004; Kubeš et al., 2014). The isoflavone conjugates are hydrolyzed and converted to more complex derivatives such as phytoalexins in response to infection (Graham et al., 1990). Synthesis and release of phytoalexins can limit pathogen colonization, induce toxicity, and increase plant resistance (Lygin et al., 2010). Phytoalexins such as glyceollins are produced in soybean as part of the host response to plant-pathogen infection. Glyceollins can increase plant resistance and reduce pathogen colonization. Glyceollins inhibit the growth of a wide range of soybean pathogens such as *Phytophthora sojae*, *Sclerotinia sclerotiorium*, and *Macrophomina phaseolina* (Lygin et al., 2010). A reduction in isoflavonoid levels by silencing of key biosynthetic genes *IFS* or *CHR* leads to an increase in plant susceptibility to pathogens (Graham et al., 2007) demonstrating its direct role in plant defense. It is postulated that isoflavonoids play a key role in the protection of plants against stress by inhibiting the formation of reactive oxygen species through a range of different mechanisms (Mierziak et al., 2014).

1.4 Isoflavonoid biosynthesis in soybean

1.4.1 Overview of the pathway

Isoflavonoids are synthesized via the phenylpropanoid pathway (Fig 1.2). It begins with the deamination of phenylalanine to cinnamate by phenylalanine ammonialyase (PAL) (Hahlbrock and Scheel, 1989). Cinnamate is then converted into *p*-coumaric acid by cinnamate 4-hydroxylaye (C4H). *p*-Coumaric acid is converted to*p*-coumaroyl-CoA by 4-coumarate-CoA-ligase (4CL). Chalcone is produced in the first step in the branched pathway for the synthesis of isoflavonoids and flavonoids (Hahlbrock and Scheel, 1989)

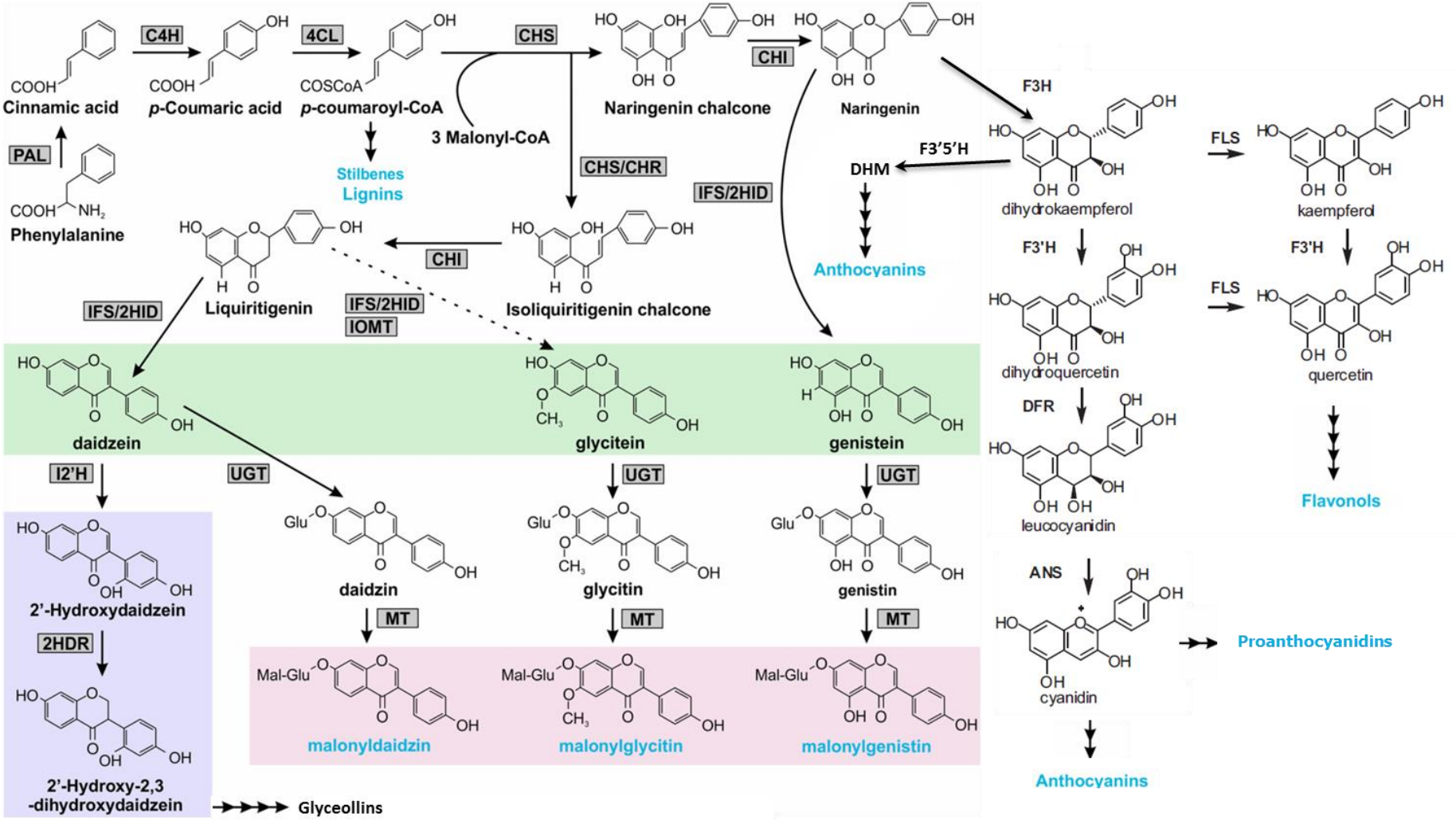

Cinnamic acid
C4H
p-Coumaric acid
4CL
p-coumaroyl-CoA
CHS
3 Malonyl-CoA
Naringenin chalcone
CHI
Naringenin
PAL
Phenylalanine
Stilbenes
Lignins
CHS/CHR
IFS/2HID
F3H
F3'5'H
DHM
Anthocyanins
dihydrokaempferol
FLS
kaempferol
F3'H
dihydroquercetin
quercetin
DFR
leucocyanidin
Flavonols
ANS
cyanidin
Proanthocyanidins
Anthocyanins
CHI
Liquiritigenin
IFS/2HID
IOMT
Isoliquiritigenin chalcone
IFS/2HID
daidzein
glycitein
genistein
I2'H
UGT
UGT
UGT
2'-Hydroxydaidzein
daidzin
glycitin
genistin
2HDR
MT
MT
MT
malonyldaidzin
malonylglycitin
malonylgenistin
2'-Hydroxy-2,3
-dihydroxydaidzein
Glyceollins

Figure 1.2 Isoflavonoid biosynthetic pathway in soybean. The multiple *arrows* indicate multiple steps in the pathway and the *dotted arrow* indicates speculated steps. The major metabolites synthesized from the phenylpropanoid pathway are shown in *blue* text. The three isoflavoneaglycones, daidzein, glycitein, and genistein are highlighted in *green* and malonylisoflavones are highlighted in *pink*. The induced isoflavonoid phytoalexin synthesis branch is highlighted in purple (light). PAL, phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylate; 4CL, 4-coumarate-CoA-ligase; CHS, chalconesynthase; CHR, chalcone reductase; CHI, chalcone isomerase; IFS, 2-hydroxyisoflavanone synthase; 2HID, 2-hydroxyisoflavanone dehydratase; IOMT, isoflavone *O*-methyltransferase; UGT- uridine diphosphate glycosyltransferase; MT, malonyltransferase; I2'H, Isoflavone 2'-hydroxylase; 2HDR, 2'-hydroxydaidzein reductase; F3H, flavanone-3-hydroxylase; F3′H, flavonoid 3'hydroxylase; FLS, flavonol synthase; F3'5'H, flavonoid 3'5'-hydroxylase; DHM, dihydromyricetin; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase. [Modified from Saito et al. (2013); Albert et al. (2014); Anguraj Vadivel et al. (2015)**]**

and the reaction is catalyzed by chalcone synthase (CHS). It is considered to be the crucial step of the phenylpropanoid pathway where the branching of flavonoids and/or isoflavonoids pathway occurs. CHS is encoded by a single gene in some plants such as parsley (*Petroselinum crispum* (Mill.) Fuss) (Hermann et al., 1988) and by multiple genes in others, for instance, soybean (Tuteja and Vodkin, 2008), barley (*Hordeum vulgare* L.) (Roth et al., 1991) and petunia (*Petunia hybrida*) (Koes et al., 1989).

The soybean genome contains a nine-member *CHS* gene family (*GmCHS1–GmCHS9*) (Tuteja and Vodkin, 2008). The enzyme CHS alone or in conjunction with a legume-specific chalcone reductase (CHR) condenses *p*-coumaroyl-CoA with three molecules of malonyl-CoA to form chalcone. There are two types of chalcones: tetrahydroxy chalcones e.g., naringenin chalcone and trihydroxy chalcone e.g., isoliquiritigenin chalcone. Naringenin chalcone is present in all plant species, whereas isoliquiritigenin chalcone is legume-specific (Ralston et al., 2005). Naringenin chalcone and isoliquiritigenin chalcone are the substrates for chalcone isomerase (CHI)-mediated production of tricyclic flavanones naringenin and liquiritigenin, respectively (Dastmalchi and Dhaubhadel, 2015). The soybean CHI family contains twelve members that are grouped into four subfamilies: GmCHI1 (GmCHI1A, GmCHI1B1, GmCHI1B2), GmCHI2, GmCHI3 (GmCHI3A1, GmCHI3A2, GmCHI3B1, GmCHI3B2, GmCHI3C1, GmCHI3C2) and GmCHI4 (GmCHI4A, GmCHI4B). Among these, the catalytically active ones are GmCHI1A, GmCHI1B1, GmCHI1B2, and GmCHI2 (Ralston et al., 2005). The expression of *GmCHI1A* and *GmCHI1B2* transcripts correlates with isoflavonoid accumulation in soybean seeds (Dastmalchi and Dhaubhadel, 2015).

The first committed step in the isoflavonoid biosynthesis pathway is catalyzed by the legume-specific membrane-bound cytochrome P450 monooxygenase, isoflavone synthase (IFS), which facilitates aryl B-ring migration to the C-3 position of C ring to generate the core isoflavonoid structure (Steele et al., 1999). The flavanone substrate, naringenin is shared by the flavonoid and isoflavonoid branches of the phenylpropanoid pathway and hence competition between the branches occurs over the same substrate. The level of competition depends on the spatiotemporal availability of the enzymes such as GmCHS,

GmCHI, GmIFS, and GmF3H in a tissue (Dastmalchi and Dhaubhadel, 2014) (Fig 1.2). The combined activity of IFS and 2-hydroxyisoflavone dehydrogenase (2HID) produces the three isoflavone aglycones daidzein, genistein, and glycitein. There are two *IFS* genes, *GmIFS1* and *GmIFS2* (Jung et al., 2000) that differ by 14 amino acids in their protein sequences and display tissue-specific gene expression (Dhaubhadel et al., 2003). Reduced levels of isoflavone accumulation in *IFS*-silenced plants increased the susceptibility of soybean hairy roots to *P. sojae* infection, suggesting that isoflavonoids are involved in defense mechanisms during plant-pathogen interactions (Subramanian et al., 2005). Typically the isoflavones aglycones, daidzein, genistein, and glycitein are conjugated to glucose by uridine diphosphate glycosyltransferases (UGTs) to produce the 7-*O*-glycosides daidzin, genistin, and glycitin, respectively. Furthermore, malonylation by malonyltransferases (MTs) produce 6"-*O*-malonyldaidzin, 6"-*O*-malonylgenistin and 6"-*O*-malonylglycitin (Kudou et al., 1991; Dhaubhadel et al., 2008). Conjugation increases the water solubility and chemical stability of isoflavones promoting their storage in cell vacuoles (Jones and Vogt, 2001). Daidzein serves as the precursor for the production of the soybean phytoalexin glyceollin. It involves the hydroxylation of daidzein that is catalyzed by isoflavone 2'-hydroxylase (Kochs and Grisebach, 1986; Akashi et al., 1998). After reduction and cyclization by 2'-hydroxydaidzein oxidoreductase (2HDR) (Fischer et al., 1990a) and pterocarpan synthase (PTS) (Fischer et al., 1990b) respectively, the compound is then converted to glycinol; subsequent prenylation of glycinol followed by hydroxylation leads to the productions of three forms of glyceollin (Anguraj Vadivel et al., 2015).

1.4.2 **Regulation and accumulation of isoflavonoids in soybean tissues**

Isoflavonoid production and accumulation in soybean is a quantitative trait that is controlled by multiple factors such as plant genotype and the environment (Hoeck et al., 2000; Lee et al., 2003; Berger et al., 2008). Soil moisture and variation in temperature during seed development affect isoflavonoid content in soybean (Bennett et al., 2004; Lozovaya et al., 2005).

Isoflavonoid and flavonoid accumulation in plant tissues is influenced by stress, although the role of individual molecules is unclear. It is difficult to determine the role of each molecule because plants accumulate a diverse spectrum of these compounds. In the phenylpropanoid pathway, the basic structures of phenylpropanoids are often converted into glycosylated or malonylated derivatives (Dixon and Paiva, 1995). Significant amounts of isoflavonoids accumulate in the soybean seed during development and the seed isoflavonoid content is believed to be the result of synthesis within the seed and also from the transport of compounds from maternal tissues (Dhaubhadel et al., 2003).

A rate-limiting enzyme for isoflavonoid synthesis is IFS (Steele et al., 1999; Jung et al., 2000). The expression level of *GmIFS2* increases during embryo development and peaks at 70 days after pollination (DAP). The gene expression patterns of three genes, *GmIFS*, *GmCHS7*, and *GmCHS8* correlate well with seed isoflavonoid accumulation in soybean (Dhaubhadel et al., 2007). This suggests the expression of these genes regulates metabolite accumulation in the soybean seed. A tissue-specific expression study of *GmCHS7* and *GmCHS8* transcript levels are correlated with seed coat colour pigments (Tuteja et al., 2004). *GmCHS7* and *GmCHS8* also showed diversified patterns in their tissue-specific expression (Yi et al., 2010a). The differences in gene expression between the two soybean cultivars RCAT Angora and Harovinton could play a role in the differential accumulation of isoflavonoids in their seeds. Comparison of global gene expressions in these two cultivars has indicated the involvement of *GmCHS7* and *GmCHS8* in isoflavonoid biosynthesis (Dhaubhadel et al., 2007). All *CHS* genes share a high degree of sequence similarity within and across species, but they may have divergent promoter sequences, including different *cis*-acting regions that could contribute to their differential expression patterns. CHS play various roles during plant development or in response to environmental stimuli. Among the family members, *GmCHS7* and *GmCHS8* genes form a distinct phylogenetic clade whereas, *GmCHS1–GmCHS6* cluster closer together (Matsumura et al., 2005). Both *GmCHS7* and *GmCHS8* genes are expressed most abundantly in roots (Tuteja et al., 2004). RNAi silencing of *GmCHS8* reduced the isoflavonoid level in soybean hairy roots, implying the involvement of GmCHS8 in isoflavonoid biosynthesis in soybean. Isoflavonoid biosynthesis pathway is

complex with multidimensional regulation at genomic, transcriptomic, proteomic and metabolite level. Each step in the pathway is catalyzed by an enzyme that contains multiple isoforms whose function is determined by multiple factors that are not fully understood. Coordinated expression of phenylpropanoid pathway genes for their co-regulation can be tuned by a transcription factor or multiple transcription factors in a complex (Albert et al., 2014).

1.5 MYB transcription factors and transcriptional complexes

Transcriptional regulation of gene expression plays an important role in plant development and secondary metabolism. Transcription factors are proteins that bind to a specific DNA sequence in a gene's promoter region, thereby regulating its transcription. This function can be performed by a single transcription factor or by a complex containing multiple proteins including transcription factors. Transcription factors are classified into various classes based on sequence homology and structural similarity (Stegmaier et al., 2004). The transcriptional regulation of phenylpropanoid biosynthetic genes occurs as a result of coordinated regulation ofstructural genes by several DNA-binding factors, such as MYB, bHLH, bZIP, WRKY, and MADS-box transcription factors (Nesi et al., 2001; Ramsay and Glover, 2005). A comprehensive analysis of *CHS* expression in *Arabidopsis* revealed the involvement of an R2R3 MYB transcription factor in *CHS* gene regulation and its effects on flavonoid composition (Mehrtens et al., 2005; Stracke et al., 2007).

MYB transcription factors are classified within the helix-turn-helix (HTH) super class that represents a large transcription factor family. All MYBs share a conserved DNA-binding domain. The first *MYB* gene was identified as an oncogene in avian myeloblastosis virus. *C-MYB* is a proto-oncogene that plays a vital role in cell growth and differentiation in human and has three repeats, which are named R1, R2, and R3. The repeats of all other MYB proteins are named based on their similarity to the R1, R2 and R3 repeats of C-MYB (reviewed in Dubos et al., 2010). Generally, MYB proteins contain an HTH domain with one to four imperfect repeats. MYB proteins are classified into four types on the basis on the number of HTH domains: R1 MYB, R2R3 MYB, R3 MYB, and

R4 MYB. R1 MYB is a single-repeat MYB transcription factor, which has either an R1 or an R3 repeat (Fig 1.3). R2R3 MYB has R2 and R3, whereas R3 MYB has R1, R2, and R3 repeats. The smallest group contains the R4 MYB transcription factors, which have two R1-like and two R2-like repeats. The HTH structure is formed by the second and third helices of each repeat (Fig 1.3) (reviewed in Dubos et al., 2010). Regularly spaced tryptophan residues form a cluster in each repeat and stabilize the DNA-binding domain structure of the protein(reviewed in Du et al., 2009). Some R1 MYB transcription factors act as competitors to R2R3 MYB for the same DNA targets (reviewed in Prouse and Campbell, 2012).

MYB transcription factors belong to the largest transcription factor superfamily in higher plants and have been studied extensively in Arabidopsis, grapevine, apple, petunia, maize, and rice (reviewed in Liu et al., 2015). They regulate many plant cellular processes including metabolism. The first plant *MYB* gene, *C1* was isolated from maize (*Zea mays*) and is involved in anthocyanin biosynthesis (Paz-Ares et al., 1987). Several MYB proteins have been reported to play a vital role in the regulation of the phenylpropanoid pathway (Jin and Martin, 1999; Boddu et al., 2006). Recently, it has been shown that overexpression or silencing of an R2R3 MYB, *MYB10* directly correlates with increased or reduced levels of anthocyanin in strawberry (Lin-Wang et al., 2014). In eukaryotes, transcription is regulated by multiple protein complexes. Studies of MYB transcription factors in *Arabidopsis* revealed that MYB proteins in complex with other proteins, through their interactive C-terminal region, to regulate transcription of genes in the phenylpropanoid pathway (Stracke et al., 2001). In maize, an MYB transcription factor C1 interacts with a basic helix-loop-helix (bHLH) transcription factor Lc and activates the expression of anthocyanin biosynthetic genes (Goff et al., 1992). Many plant genes are regulated by the BHLH-MYB-WD40 transcription complex (Payne et al., 2000; Wang and Chen, 2008; Ben-Simhon et al., 2011).

A metabolic engineering approach has been used to increase isoflavonoid content in soybean seeds by introducing maize C1 and R transcription factors (Yu et al., 2003). C1 is an R2R3 MYB transcription factor that is involved in anthocyanin accumulation in

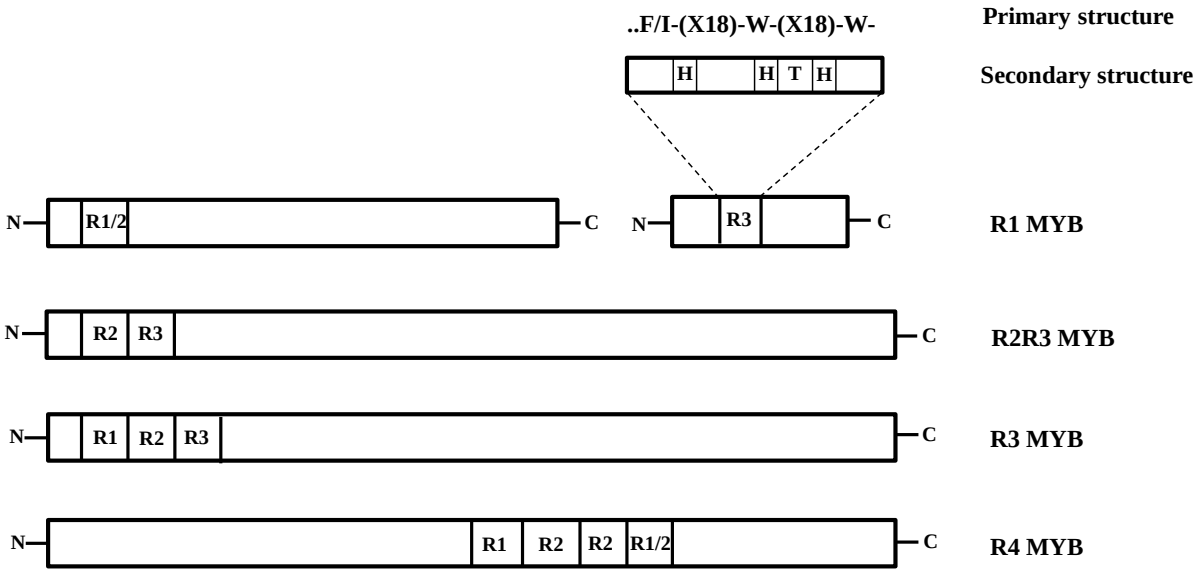


Figure 1.3 Classes of MYB transcription factors. Illustration showing the four classes of MYB proteins depending on the number of MYB repeats (R). Each repeat has three α helices. The second and third α helices in each repeat form helix-turn-helix structure in MYB proteins.[Modified from Dubos et al. (2010)] H-helix; T- turn; W- tryptophan; X- amino acid (X).

maize and requires a bHLH containing R MYC co-factor for its function (Grotewold et al., 1998). A chimeric protein, CRC, in which R was inserted between the DNA-binding and activation domains of C1, increased isoflavonoid production up to four times that of wild-type soybean seeds. This effect was observed when transcription of flavonoid biosynthetic genes was suppressed, and the complete flow of substrates was diverted to the isoflavonoid pathway (Yu et al., 2003), thereby suggesting a competition for the precursors by flavonoid and isoflavonoid branches of the phenylpropanoid pathway. The transcriptomics and metabolomics of the isoflavonoid pathway have been studied (Dhaubhadel et al., 2007; Farag et al., 2008; Jones and Vodkin, 2013; Lin et al., 2014) but, the role of regulatory elements in the pathway requires further investigation.

Co-ordinated expression of structural genes in the phenylpropanoid pathway has been studied in plants. In petunia, the expression of the flavonoid biosynthesis genes *F3H, F3'5'H, DFR, ANS, 3RT, 5GT, GST* are co-regulated by MYB27 (Albert et al., 2014). But the co-ordinated expression of isoflavonoid biosynthetic pathway genes and their spatio-temporal regulation is not known in soybean. Identification of a transcription factor that could co-regulate multiple genes in the isoflavonoid biosynthesis pathway is important and useful for metabolic engineering of isoflavonoid production in plants.

1.6 GmMYB176: the regulator of isoflavonoid biosynthesis in soybean

GmMYB176 encodes for an R1 MYB protein in soybean. It contains a single repeat HTH structure between the amino acids positions 73 and 129 and includes an SHAQKYF motif in the third predicted alpha helix that presumably functions in DNA binding (Yi et al., 2010b). GmMYB176 also forms a homodimer which possibly allows it to recognize target DNA with high affinity and specificity (Dhaubhadel and Li, 2010). Single repeat MYB transcription factors CCA1 and LHY that regulate the circadian clock in *Arabidopsis* also form homodimers and heterodimers with the other protein (Lu et al., 2009).

Yi et al. (2010b) conducted a promoter deletion study to identify the region(s) that is(are) critical for *GmCHS8* gene expression in soybean. Eight fragments (*CHS8*Δ*P1–CHS8*Δ*P8*) with 5' deletions of the *GmCHS8* promoter were generated and used to drive GUS expression as a reporter in *Arabidopsis* protoplasts. The reporters *CHS8*Δ*P1–CHS8*Δ*P8pro:GUS* and *CHS8fullpro:GUS* were co-transfected with *35Spro:* GmMYB*176* into *Arabidopsis* protoplasts (Fig 1.4). The GUS activity was similar from *CHS8full* to *CHS8*Δ*P5*, but it was drastically reduced from *CHS8*Δ*P4* to *CHS8*Δ*P1* suggesting the sequence region in the *CHS8*Δ*P5* fragment that is missing in *CHS8*Δ*P4* is critical for *GmCHS8* gene expression. The region that separates *CHS8*Δ*P5* from *CHS8*Δ*P4* is a 23 bp sequence between -795 and -818 (GCGTGAAAATATAGTTAGTATAT) in the *GmCHS8* gene promoter that contains the TAGT(T/A)(A/T) motif thought to be the recognition sequence for GmMYB176. Binding of GmMYB176 at this site in *GmCHS8* is critical for its expression (Fig 1.4) (Yi et al., 2010b).

It is known that the interaction of GmMYB176 with soybean 14-3-3 proteins affects its subcellular localization. Soybean has 16 active *SGF14* (14-3-3) genes (Li and Dhaubhadel, 2011). All 16 SGF14s interacts with GmMYB176 *in planta* (Li et al., 2012). Phosphorylated GmMYB176 binds with soybean 14-3-3 proteins (SGF14s) and localized to the nucleus and the cytoplasm whereas, unphosphorylated GmMYB176 (GmMYB176D2/GmMYB176S29A) gets retained in the nucleus only (Li et al., 2012). RNAi silencing of *GmMYB176* in soybean hairy roots culminates in reduced levels of isoflavonoids, suggesting that GmMYB176 is necessary for isoflavonoid biosynthesis. However, overexpression of GmMYB176 did not increase *GmCHS8* transcript level and isoflavonoid levels in hairy roots (Yi et al., 2010b). Hence, it is hypothesized that GmMYB176 may work in combination with other factor(s) for the regulation of *GmCHS8* gene expression and it may also regulate other isoflavonoid biosynthetic genes in soybean. Recently, a CHS protein was identified as one of the GmMYB176 interactors (Anguraj Vadivel and Dhaubhadel, unpublished), and its amino acid sequence did not match with any of the nine known CHSs in soybean. It is not known if GmMYB176 regulates the expression of other members of the *CHS* gene family.

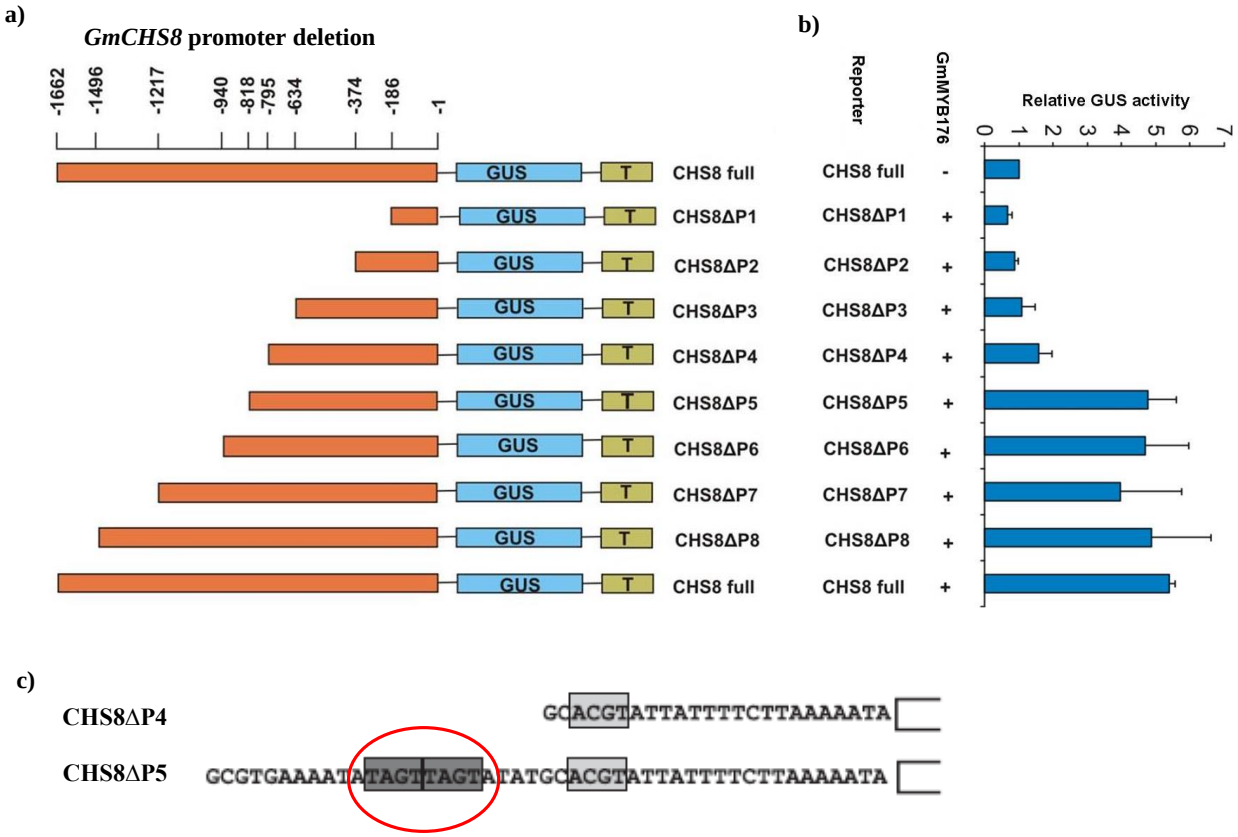


Figure 1.4 Promoter deletion study of *GmCHS8* in soybean. a) Eight promoter fragments (*CHS8ΔP1-CHS8ΔP5*) (orange boxes) were generated and used to drive GUS (blue boxes) expression in *Arabidopsis* protoplast. T (beige boxes) - Terminator. b) Relative GUS activity showed that the region in the *CHS8ΔP5* fragment is critical for *GmCHS8* gene expression. c) A 23 bp motif in the *CHS8ΔP5* fragment that is critical for *GmCHS8* gene expression. The red circle shows the GmMYB176 binding motif. [Modified from (Yi et al., 2010b)**]**.

1.7 Objectives of the study

The objective of this study was to identify the GmMYB176 interactome to elucidate the transcriptional complex that regulates *GmCHS8* gene expression and to identify isoflavonoid biosynthetic genes regulated by GmMYB176 in soybean. The following goals were set.

A. Identifying members of the *GmCHS* gene family

Conduct an *in silico* analysis to identify all members of the *GmCHS* gene family in soybean. Phylogenetic analysis of the gene family and their differential expression in soybean tissues will comply.

B. Identifying the GmMYB176 interactome

Evaluate the GmMYB176 interactome by co-immunoprecipitation (Co-IP) followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Transgenic soybean hairy roots carrying GmMYB176 and YFP fusion constructs were planned to be used for this study. Selected interactors will be verified for their protein-protein interactions *in planta*. GmMYB176 interactors that bind to the *GmCHS8* gene promoter were to be identified by yeast one-hybrid (Y1H) assay.

C. Identifying the regulon and metabolites regulated by *GmMYB176* in soybean

Identify isoflavonoid biosynthetic genes and metabolites controlled by GmMYB176. Silencing and overexpressing *GmMYB176* to study effects on gene expression and metabolite levels in soybean hairy roots by RNAseq and metabolomics approaches, respectively. The differentially expressed isoflavonoid biosynthetic genes in silenced or overexpressed tissues to be verified by qPCR.

1.8 **Literature cited**

Abbasi, P.A., and Graham, T.L. (2001). Age-related regulation of induced isoflavonoid responses in soybean lines differing in inherent elicitation competency. Physiological and Molecular Plant Pathology **59,** 143-152.

Agarwal, D.K., Billore, S.D., Sharma, A.N., Dupare, B.U., and Srivastava, S.K. (2013). Soybean: Introduction, Improvement, and Utilization in India—Problems and Prospects. Agricultural Research **2,** 293-300.

Akashi, T., Aoki, T., and Ayabe, S.I. (1998). Identification of a cytochrome P450 cDNA encoding (2S)-flavanone 2-hydroxylase of licorice (*Glycyrrhiza echinata*L.; Fabaceae) which represents licodione synthase and flavone synthase II. FEBS Letters **431,** 287-290.

Albert, N.W., Davies, K.M., Lewis, D.H., Zhang, H., Montefiori, M., Brendolise, C., Boase, M.R., Ngo, H., Jameson, P.E., and Schwinn, K.E. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. Plant Cell **26,** 962-980.

Anderson, J.W., Johnstone, B.M., and Cook-Newell, M.E. (1995). Meta-analysis of the effects of soy protein intake on serum lipids. New England Journal of Medicine **333,** 276-282.

Anguraj Vadivel, A.K., Sukumaran, A., Li, X., and Dhaubhadel, S. (2015). Soybean isoflavonoids: role of GmMYB176 interactome and 14-3-3 proteins. Phytochemistry Reviews **15,** 391-403.

Barnes, S., Kim, H., Darley-Usmar, V., Patel, R., Xu, J., Boersma, B., and Luo, M. (2000). Beyond ERalpha and ERbeta: estrogen receptor binding is only part of the isoflavone story. Journal of Nutrition **130,** 656S-657S.

Ben-Simhon, Z., Judeinstein, S., Nadler-Hassar, T., Trainin, T., Bar-Ya'akov, I., Borochov-Neori, H., and Holland, D. (2011). A pomegranate (*Punica granatum* L.) WD40-repeat gene is a functional homologue of *Arabidopsis* TTG1 and is involved in the regulation of anthocyanin biosynthesis during pomegranate fruit development. Planta **234,** 865-881.

Bennett, J.O., Yu, O., Heatherly, L.G., and Krishnan, H.B. (2004). Accumulation of genistein and daidzein, soybean isoflavones implicated in promoting human health, is significantly elevated by irrigation. Journal of Agricultural and Food Chemistry **52,** 7574 - 7579.

Berger, M., Rasolohery, C.A., Cazalis, R., and Daydé, J. (2008). Isoflavone accumulation kinetics in soybean seed cotyledons and hypocotyls: distinct pathways and genetic controls. Crop Science **48,** 700-708.

Boddu, J., Jiang, C., Sangar, V., Olson, T., Peterson, T., and Chopra, S. (2006). Comparative structural and functional characterization of sorghum and maize duplications containing orthologous MYB transcription regulators of 3-deoxyflavonoid biosynthesis. Plant Molecular Biology **60,** 185-199.

Buer, C.S., Imin, N., and Djordjevic, M.A. (2010). Flavonoids: new roles for old molecules. Journal of Integrative Plant Biology**52,** 98-111.

Chen, A.M., and Rogan, W.J. (2004). Isoflavones in soy infant formula: a review of evidence for endocrine and other activity infants. Annual Review of Nutrition **24,** 33-54.

Dastmalchi, M., and Dhaubhadel, S. (2014). Soybean seed isoflavonoids: biosynthesis and regulation. In Phytochemicals – Biosynthesis, Function and Application, R. Jetter, ed (Cham: Springer International Publishing), pp. 1-21.

Dastmalchi, M., and Dhaubhadel, S. (2015). Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. Planta **241,** 507-523.

Dhaubhadel, S., and Li, X. (2010). A new client for 14-3-3 proteins: GmMYB176, an R1 MYB transcription factor. Plant Signaling and Behavior **5,** 921-923.

Dhaubhadel, S., Farhangkhoee, M., and Chapman, R. (2008). Identification and characterization of isoflavonoid specific glycosyltransferase and malonyltransferase from soybean seeds. Journal of Experimental Botany **59,** 981-994.

Dhaubhadel, S., McGarvey, B.D., Williams, R., and Gijzen, M. (2003). Isoflavonoid biosynthesis and accumulation in developing soybean seeds. Plant Molecular Biology **53,** 733-743.

Dhaubhadel, S., Gijzen, M., Moy, P., and Farhangkhoee, M. (2007). Transcriptome analysis reveals a critical role of *CHS7* and *CHS8* genes for isoflavonoid synthesis in soybean seeds. Plant Physiology **143,** 326-338.

Dixon, R.A. (2004). Phytoestrogens. Annual Review of Plant Biology**55,** 225-261.

Dixon, R.A., and Paiva, N.-L. (1995). Stress-induced phenylpropanoid metabolism. Plant Cell **7,** 1085-1097.

Dixon, R.A., and Ferreria, D. (2002). Genistein. Phytochemistry **60,** 205-211.

Du, H., Zhang, L., Liu, L., Tang, X.F., Yang, W.J., Wu, Y.M., Huang, Y.B., and Tang, Y.X. (2009). Biochemical and molecular characterization of plant MYB transcription factor family. Biochemistry (Mosc) **74,** 1-11.

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

Duffy, C., Perez, K., and Partridge, A. (2007). Implications of phytoestrogen intake for breast cancer. CA Cancer Journal for Clinicians **57,** 260-277.

Farag, M.A., Huhman, D.V., Dixon, R.A., and Sumner, L.W. (2008). Metabolomics reveals novel pathways and differential mechanistic and elicitor-specific responses in phenylpropanoid and isoflavonoid biosynthesis in *Medicago truncatula* cell cultures. Plant Physiology **146,** 387-402.

Fischer, D., Ebenau-Jehle, C., and Grisebach, H. (1990a). Phytoalexin synthesis in soybean: Purification and characterization of NADPH:2′-hydroxydaidzein oxidoreductase from elicitor-challenged soybean cell cultures. Archives of Biochemistry and Biophysics **276,** 390-395.

Fischer, D., Ebenau-Jehle, C., and Grisebach, H. (1990b). Purification and characterization of pterocarpan synthase from elicitor-challenged soybean cell cultures. Phytochemistry **29,** 2879-2882.

Goff, S.A., Cone, K.C., and Chandler, V.L. (1992). Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. Genes & development **6,** 864-875.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research **40,** D1178-1186.

Graham, T.L., Kim, J.E., and Graham, M.Y. (1990). Role of constitutive isoflavone conjugates in the accumulation of glyceollin in soybean infected with *Phytophthora-Megasperma*. Molecular Plant-Microbe Interactions **3,** 157-166.

Graham, T.L., Graham, M.Y., Subramanian, S., and Yu, O. (2007). RNAi silencing of genes for elicitation or biosynthesis of 5-deoxyisoflavonoids suppresses race-specific resistance and hypersensitive cell death in *phytophthora sojae* infected tissues. Plant Physiology **144,** 728-740.

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

Guo, J., Xu, X., Rasheed, T.K., Yoder, A., Yu, D., Liang, H., Yi, F., Hawley, T., Jin, T., Ling, B., and Wu, Y. (2013). Genistein interferes with SDF-1- and HIV-

mediated actin dynamics and inhibits HIV infection of resting CD4 T cells. Retrovirology **10,** 62.

Hahlbrock, K., and Scheel, D. (1989). Physiology and molecular biology of phenylpropanoid metabolism. Annual Review of Plant Physiology and Plant Molecular Biology **40,** 347-369.

Hermann, A., Schulz, W., and Harlbrock, K. (1988). Two alleles of the single-copy *chalcone synthase* gene in parsley differ by a transposon-like element. Molecular and General Genetics **212,** 93-98.

Herrmann, K.M. (1995). The shikimate pathway as an entry to aromatic secondary metabolism. Plant Physiology **107,** 7-12.

Hoeck, J.A., Fehr, W.R., Murphy, P.A., and Welke, G.A. (2000). Influence of genotype and environment on isoflavone contents of soybean. Crop Science **40,** 48 - 51.

Hooper, L., Kroon, P.A., Rimm, E.B., Cohn, J.S., Harvey, I., Le Cornu, K.A., Ryder, J.J., Hall, W.L., and Cassidy, A. (2008). Flavonoids, flavonoid-rich foods, and cardiovascular risk: a meta-analysis of randomized controlled trials. American Journal of Clinical Nutrition **88,** 38-50.

Howes, L.G., Howes, J.B., and Knight, D.C. (2006). Isoflavone therapy for menopausal flushes: a systematic review and meta-analysis. Maturitas **55,** 203-211.

Jeong, Y.J., An, C.H., Woo, S.G., Park, J.H., Lee, K.W., Lee, S.H., Rim, Y., Jeong, H.J., Ryu, Y.B., and Kim, C.Y. (2016). Enhanced production of resveratrol derivatives in tobacco plants by improving the metabolic flux of intermediates in the phenylpropanoid pathway. Plant Molecular Biology **92,** 117-129.

Jin, H., and Martin, C. (1999). Multifunctionality and diversity within the plant MYB-gene family. Plant Molecular Biology **41,** 577-585.

Jones, P., and Vogt, T. (2001). Glycosyltransferases in secondary plant metabolism: transquilizers and stimulant controllers. Planta **213,** 164-174.

Jones, S.I., and Vodkin, L.O. (2013). Using RNA-Seq to profile soybean seed development from fertilization to maturity. PLoS One **8,** e59270.

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,** 559-559.

Kochs, G., and Grisebach, H. (1986). Enzymic synthesis of isoflavones. European journal of biochemistry / FEBS **155,** 311-318.

Koes, R.E., Spelt, C.E., van den Elzen, P.J.M., and Mol, J.N.M. (1989). Cloning and molecular characterization of the *chalcone synthase* multigene family of *Petunia hybrida*. Gene **81,** 245-257.

Korde, L.A., Wu, A.H., Fears, T., Nomura, A.M., West, D.W., Kolonel, L.N., Pike, M.C., Hoover, R.N., and Ziegler, R.G. (2009). Childhood soy intake and breast cancer risk in Asian American women. Cancer Epidemiology Biomarkers and Prevention **18,** 1050-1059.

Kubeš, J., Tůmová, L., Martin, J., Vildová, A., Hendrychová, H., and Sojková, K. (2014). The production of isoflavonoids in *Genista tinctoria* L. cell suspension culture after abiotic stressors treatment. Natural Product Research **28,** 2253-2263.

Kudou, S., Fleury, Y., Welt, D., Magnolato, D., Uchida, T., and Kitamura, K. (1991). Malonyl isoflavone glycosides in soybean seeds (*Glycine max* Merill). Agricultural and Biological Chemistry **55,** 2227-2233.

La Camera, S., Gouzerh, G., Dhondt, S., Hoffmann, L., Fritig, B., Legrand, M., and Heitz, T. (2004). Metabolic reprogramming in plant innate immunity: the contributions of phenylpropanoid and oxylipin pathways. Immunological Reviews**198,** 267-284.

Lamartiniere, C.A., Moore, J.B., Brown, N.M., Thompson, R., Hardin, M.J., and Barnes, S. (1995). Genistein suppresses mammary cancer in rats. Carcinogenesis **16,** 2833-2840.

Lapcik, O. (2007). Isoflavonoids in non-leguminous taxa: a rarity or a rule? Phytochemistry **68,** 2909-2916.

Lee, G.A., Crawford, G.W., Liu, L., Sasaki, Y., and Chen, X. (2011). Archaeological soybean (*Glycine max*) in East Asia: does size matter? PLoS One **6,** e26720.

Lee, S.J., Yan, W., Ahn, J.K., and Chung, I.M. (2003). Effects of year, site, genotype and their interactions on various soybean isoflavones. Field Crops Research **81,** 181-192.

Li, X., and Dhaubhadel, S. (2011). Soybean 14-3-3 gene family: identification and molecular characterization. Planta **233,** 569-582.

Li, X., Chen, L., and Dhaubhadel, S. (2012). 14-3-3 proteins regulate the intracellular localization of the transcriptional activator GmMYB176 and affect isoflavonoid synthesis in soybean. The Plant Journal **71,** 239-250.

Lin-Wang, K., McGhie, T.K., Wang, M., Liu, Y., Warren, B., Storey, R., Espley, R.V., and Allan, A.C. (2014). Engineering the anthocyanin regulatory complex of strawberry (*Fragaria vesca*). Frontiers in Plant Science**5,** 651.

Lin, F., Zhao, M., Baumann, D.D., Ping, J., Sun, L., Liu, Y., Zhang, B., Tang, Z., Hughes, E., Doerge, R.W., Hughes, T.J., and Ma, J. (2014). Molecular response to the pathogen *Phytophthora sojae* among ten soybean near isogenic lines revealed by comparative transcriptomics. BMC Genomics **15**.

Liu, C.J., Blount, J.W., Steele, C.L., and Dixon, R.A. (2002). Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America **99,** 14578-14583.

Liu, J., Osbourn, A., and Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. Molecular Plant **8,** 689-708.

Lozovaya, V.V., Lygin, A.V., Zernova, O.V., Li, S., Hartman, G.L., and Widholm, J.M. (2004). Isoflavonoid accumulation in soybean hairy roots upon treatment with *Fusarium solani*. Plant Physiology and Biochemistry **42,** 671-679.

Lozovaya, V.V., Lygin, A.V., Ulanov, A.V., Nelson, R.L., Daydé, J., and Widholm, J.M. (2005). Effect of temperature and soil moisture status during seed development on soybean seed isoflavone concentration and composition. Crop Science **45,** 1934-1940.

Lu, S.X., Knowles, S.M., Andronis, C., Ong, M.S., and Tobin, E.M. (2009). CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL function synergistically in the circadian clock of *Arabidopsis*. Plant Physiology **150,** 834-843.

Lygin, A.V., Hill, C.B., Zernova, O.V., Crull, L., Widholm, J.M., Hartman, G.L., and Lozovaya, V.V. (2010). Response of soybean pathogens to glyceollin. Phytopathology **100,** 897-903.

Matsumura, H., Watanabe, S., Harada, K., Senda, M., Akada, S., Kawasaki, S., Dubouzet, E.G., Minaka, N., and Takahashi, R. (2005). Molecular linkage mapping and phylogeny of the *chalcone synthase* multigene family in soybean. Theoretical and Applied Genetics **110,** 1203 - 1209.

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

Messina, M. (2010). Insights gained from 20 years of soy research. Journal of Nutrition **140,** 2289S-2295S.

Mierziak, J., Kostyn, K., and Kulma, A. (2014). Flavonoids as important molecules of plant interactions with the environment. Molecules **19,** 16240-16265.

Mulligan, J.T., and Long, S.R. (1985). Induction of *Rhizobium meliloti* nodC expression by plant exudate requires nodD. Proceedings of the National Academy of Sciences of the United States of America **82,** 6609-6613.

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

Orgaard, A., and Jensen, L. (2008). The effects of soy isoflavones on obesity. Exp Biol Med (Maywood) **233,** 1066-1080.

Payne, C.T., Zhang, F., and Lloyd, A.M. (2000). *GL3* encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with GL1 and TTG1. Genetics **156,** 1349-1362.

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A., and Saedler, H. (1987). The regulatory c1 locus of Zea mays encodes a protein with homology to MYB proto-oncogene products and with structural similarities to transcriptional activators. The EMBO journal**6,** 3553-3558.

Phillip, D.A. (1992). Flavonoids:plant signals to soil microbes. In Recent advances in phytochemistry:phenolic metabolism in plants, H.A. Stafford and R.K. Ibrahim, eds (New York: Plenum Press), pp. 201-223.

Phillips, D.A., and Kapulnik, Y. (1995). Plant isoflavonoids, pathogens and symbionts. Trends in Microbiology **3,** 58-64.

Pisani, P., Bray, F., and Parkin, M.D. (2002). Estimates of the world-wide prevalence of cancer for 25 sites in the adult population. International Journal of Cancer **97,** 72-81.

Pollard, M., and Suckow, M.A. (2006). Dietary prevention of hormone refractory prostate cancer in Lobund-Wistar rats: a review of studies in a relevant animal model. Comparative Medicine **56,** 461-467.

Potter, S.M., Baum, J.A., Teng, H., Stillman, R.J., Shay, N.F., and Erdman, J.W., Jr. (1998). Soy protein and isoflavones: their effects on blood lipids and bone density in postmenopausal women. The American Journal of Clinical Nutrition **68,** 1375S-1379S.

Prouse, M.B., and Campbell, M.M. (2012). The interaction between MYB proteins and their target DNA binding sites. Biochimica et Biophysica Acta **1819,** 67-77.

Ralston, L., Subramanian, S., Matsuno, M., and Yu, O. (2005). Partial reconstruction of flavonoid and isoflavonoid biosynthesis in yeast using soybean type I and type II chalcone isomerases. Plant Physiology **137,** 1375-1388.

Ramsay, N.A., and Glover, B.J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends in Plant Science **10,** 63-70.

Roth, M.B., Zahler, A., and Stolk, J.H. (1991). A conserved family of nuclear phsphoproteins localized to sites of polymerase II transcription. Journal of Cell Biology **115,** 587-596.

Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., and Fernie, A.R. (2013). The flavonoid biosynthetic pathway in *Arabidopsis*: structural and genetic diversity. Plant Physiology and Biochemistry **72,** 21-34.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., and Jackson, S.A. (2010). Genome sequence of the palaeopolyploid soybean. Nature **463,** 178-183.

Shu, X.O., Jin, F., Dai, Q., Wen, W., Potter, J.D., Kushi, L.H., Ruan, Z., Gao, Y.-T., and Zheng, W. (2001). Soyfood intake during adolescence and subsequent risk of breast cancer among Chinese women. Cancer Epidemiology Biomarkers and Prevention **10,** 483-488.

Simons, R., Gruppen, H., Bovee, T.F., Verbruggen, M.A., and Vincken, J.P. (2012). Prenylated isoflavonoids from plants as selective estrogen receptor modulators (phytoSERMs). Food and Function **3,** 810-827.

Soltis, DE., Albert, VA., Leebens-Mack, J., Bell, CD., Paterson, AH., Zheng, C.,Sankoff, D., Depamphilis, CW., Wall, PK., and Soltis, PS. (2009) Polyploidy and angiosperm diversification. American Journal of Botany **96**, 336–348

Soltis, DE., Visger, CJ., and Soltis, PS. (2014) The polyploidy revolution then…. and now: Stebbins revisited. American Journal of Botany **101**, 1057–1078

Steele, C.L., Gijzen, M., Qutob, D., and Dixon, R.A. (1999). Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. Archives of Biochemistry and Biophysics **367,** 146-150.

Stegmaier, P., Kel, A.E., and Wingender, E. (2004). Systematic DNA-binding domain classification of transcription factors. Genome Informatics. International Conference on Genome Informatics **15,** 276-286.

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. Current Opinion in Plant Biology **4,** 447-456.

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

Subramanian, S., Stacey, G., and Yu, O. (2006). Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. The Plant Journal **48,** 261-273.

Subramanian, S., Graham, M.A., Yu, O., and Graham, T.L. (2005). RNA interference of soybean isoflavone synthase genes leads to silencing in tissues distal to the transformation site and to enhanced susceptibility to *Phytophthora sojae*. Plant Physiology **137,** 1345-1353.

Tuteja, J.H., and Vodkin, L.Q. (2008). Structural features of the endogenous *CHS* silencing and target loci in the soybean genome. Crop Science **48,** S49-S68.

Tuteja, J.H., Clough, S.J., Chan, W.-C., and Vodkin, L.O. (2004). Tissue-specific gene silencing mediated by a naturally occurring *chalcone synthase* gene cluster in *Glycine max*. Plant Cell **16,** 819-835.

Vogt, T. (2010). Phenylpropanoid biosynthesis. Molecular Plant **3,** 2-20.

Wang, H., and Murphy, P.A. (1994). Isoflavone composition of american and japanese soybeans in iowa: effects of variety, crop year, and location. Journal of Agricultural and Food Chemistry **42,** 1674 - 1677.

Wang, S., and Chen, J.G. (2008). Arabidopsis transient expression analysis reveals that activation of GLABRA2 may require concurrent binding of GLABRA1 and GLABRA3 to the promoter of GLABRA2. Plant & Cell Physiology **49,** 1792-1804.

Wu, A.H., Yu, M.C., Tseng, C.C., and Pike, M.C. (2008). Epidemiology of soy exposures and breast cancer risk. British Journal of Cancer **98,** 9-14.

Xu, L., Ding, Y., Catalona, W.J., Yang, X.J., Anderson, W.F., Jovanovic, B., Wellman, K., Killmer, J., Huang, X., Scheidt, K.A., Montgomery, R.B., and Bergan, R.C. (2009). MEK4 function, genistein treatment, and invasion of human prostate cancer cells. Journal of the National Cancer Institute **101,** 1141-1155.

Yan, L., and Spitznagel, E.L. (2009). Soy consumption and prostate cancer risk in men: a revisit of a meta-analysis. American Journal of Clinical Nutrition **89,** 1155-1163.

Yi, J., Derynck, M.R., Chen, L., and Dhaubhadel, S. (2010a). Differential expression of *CHS7* and *CHS8* genes in soybean. Planta **231,** 741-753.

Yi, J., Derynck, M.R., Li, X., Telmer, P., Marsolais, F., and Dhaubhadel, S. (2010b). A single-repeat MYB transcription factor, GmMYB176, regulates *CHS8* gene expression and affects isoflavonoid biosynthesis in soybean. The Plant Journal **62,** 1019-1034.

Yu, O., Shi, J., Hession, A.O., Maxwell, C.A., McGonigle, B., and Odell, J.T. (2003). Metabolic engineering to increase isoflavone biosynthesis in soybean seed. Phytochemistry **63,** 753-763.

Zhan, S., and Ho, S.C. (2005). Meta-analysis of the effects of soy protein containing isoflavones on the lipid profile. The American Journal of Clinical Nutrition **81,** 397-408.

2 IDENTIFICATION OF SOYBEAN *CHS* GENE FAMILY MEMBERS

2.1 Introduction

Whole-genome duplication has occurred multiple times over the past 200 millions of years of plant evolution. Many plant species are comprised of mixed populations of polyploid and diploid individuals (reviewed in Panchy et al., 2016). Polyploidization, which is a mechanism of gene duplication, results in an increase in both genome size and the entire gene set. Gene duplication generates two copies of a gene either to acquire novel gene functions for adaptation or evolution under reduced selective constraint. Gene duplication is a type of genomic change that can lead to novel functions of pre-existing genes (True and Carroll, 2002; Magadum et al., 2013). New genes can also arise *de novo* from intergenic space (Schlotterer, 2015) or new transcriptional regulatory sites on a promoter that alters gene expression (Wray et al., 2003). Duplicate genes are theorized to have introduced novel traits in plants over the course of evolution (Van de Peer et al., 2009; Magadum et al., 2013). The availability of whole genome sequences of a large number of plant species has shown that approximately 64.5% of plant genes are duplicated (reviewed in Panchy et al., 2016).

Soybean is a paleopolyploid that has undergone two whole genome duplication events (Shoemaker et al., 1996; Blanc and Wolfe, 2004; Gill et al., 2009). The genome sequence data indicates that 75% of soybean genes are duplicated (Schmutz et al., 2010). Duplicated genes responsible for nodulation and oil production were retained even after whole genome duplication events (Schmutz et al., 2010). Genome duplication results in an increased number of gene family members in the plant genome. For example, soybean 14-3-3 protein family has 16 active members (Li and Dhaubhadel, 2011). Although members of a gene family are similar at the sequence level, their expression levels in different tissues may differ due to their various promoter regulatory regions. The potential *cis*-elements in the promoter regions can also be subject to changes in sequence and specificity in response to developmental stage and environment (Dey et al., 2015).

Polyketide synthases (PKS) play a critical role in bridging primary and secondary metabolism in plants by catalyzing the sequential condensation of two-carbon acetate units into a growing polyketide chain. PKS enzymes are classified into type I, II, and III based on their catalytic mechanism, domain structure, and subunit organization. A type I PKS is a protein complex with a large subunit containing ketosynthase (KS), acylcarrier protein (ACP), and acyltransferase (AT) domains. A type II PKS is a protein complex composed of a heterodimeric KS and an ACP, whereas type III PKS is a homodimer of subunits based on a KS domain. Single KS domains in the type III PKS perform all the functions that are played by the essential domains in type I/II PKSs (Hertweck, 2009). In the absence of ACP, the type III PKSs are essential condensing enzymes that can act directly on acyl- CoA substrates (Hopwood and Sherman, 1990). Type III PKS enzymes are also known as the chalcone synthase (CHS)-like in plants. The CHS-like family of genes includes enzymes such as CHS, stilbene synthase (STS), 2-pyronesynthase, acridone synthase, benzophenone synthase, bibenzyle synthase, phlorisovalerophenone synthase, benzalacetone synthase, *C*-methylchalcone synthase, homoeriodictyol/eriodictyol synthase, aloesone synthase, coumaroyltriacetic acid synthase, hexaketide synthase, biphenyl synthase, stilbene carboxylate synthase, octaketide synthase, penta ketide chromone synthase, and anther-specific CHS-like (Austin and Noel, 2003).

CHSs are key enzymes of (iso)flavonoid biosynthesis. It was first crystallized from alfalfa (Ferrer et al., 1999). There are four catalytic active sites that are conserved in CHS and CHS-like protein such as STS. CHS and STS use the same substrate and are structurally similar but differ in function. It is not clear how these two enzymes will use the same substrate to produce different products. Crystallization of these enzymes from different species will be helpful to characterize the enzymes (Shomura et al., 2005). CHS catalyze a claisen condensation reaction (Ferrer et al., 1999) whereas, STS catalyze aldol cyclization (Tropf et al., 1995). No sequence difference is observed between CHS and STS, however, mutations named “aldol switch” responsible for aldol cyclization specificity that change CHS to STS activity has been defined (Austin et al., 2004).

In soybean, *GmCHS8* is regulated by GmMYB176 and silencing of *GmCHS8* or *GmMYB176* reduces the transcript level of *GmCHS8* and affects the isoflavonoid biosynthesis in soybean (Yi et al., 2010a; Yi et al., 2010b). Previously, 9 *CHS* were reported in soybean (Tuteja and Vodkin,

2008). A CHS protein was identified as one of the GmMYB176 interactors (Anguraj Vadivel and Dhaubhadel, unpublished), though its amino acid sequence did not match with any of the nine known CHSs in soybean. The whole genome sequencing of soybean was completed in the year 2010 (Schmutz et al., 2010). With the advantage of soybean genome sequence availability, here I performed a genome-wide search of *CHS* genes in soybean and identified 5 additional *GmCHS* genes. The analysis of *GmCHS* genes identifies 14 *GmCHS* (*GmCHS1*-*GmCHS14*), with 3 additional copies of *GmCHS3* and 2 additional copies of *GmCHS4* in soybean.

2.2 Materials and methods

2.2.1 *In silico* and phylogenetic analysis

To identify putative *GmCHS* genes in soybean, the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012) was used for a keyword search using 'chalcone synthase' in the annotated *G. max* Wm82.a2.v1 genome. To ensure no *GmCHSs* were missed in the keyword search, each *CHS* identified in the soybean genome was used as a query for a nucleotide BLAST (BLASTn) search. Protein sequences were retrieved for all GmCHSs from Phytozome and their molecular weight, PI were determined using the web-based tool ExPASy (http://web.expasy.org/compute_pi/) and their subcellular localizations were predicted in TargetP (http://www.cbs.dtu.dk/services/TargetP/) with default parameters.

A phylogenetic tree was constructed by using MEGA7 (Kumar et al., 2016). The amino acid sequences were aligned in ClustalOmega (ClustalO), prior to constructing a neighbour-joining tree with 1000 bootstrap replications. Pairwise nucleotide and amino acid comparison were performed using the sequence identity matrix function in BioEdit Sequence Alignment Editor Version 7.2.5. The amino acid alignment from ClustalO was used in boxshade server (http://www.ch.embnet.org/software/BOX_form.html) for shading of multiple-alignment. Active sites, malonyl-CoA binding sites and product binding sites on sequences of GmCHSs were identified using NCBI conserved domain search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

2.2.2 Generation of a heat map

RNAseq data from different soybean tissues are publically available and the expression values are presented in FPKM. FPKM values of all *GmCHS*s in soybean tissues were retrieved from

Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012). A heatmap for expression levels of *GmCHS*s in soybean tissues was constructed based on the average linkage hierarchical clustering with Pearson correlation distance using the software Multi Experimental Viewer (MeV) (http://mev.tm4.org/).

2.3 Results

2.3.1 The GmCHS family contains 14 putative members

Identification of *GmCHS*s was performed through keyword searches within the annotated *G. max* Wm82.a2.v1 genome on the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012). Each identified *GmCHS* was then used as a query for a BLAST search to ensure all *GmCHSs* were identified. A benzophenone synthase and three hydroxymethylglutaryl-CoA synthases, which also showed up in the BLAST result, were not used for further analysis. There are 19 putative *GmCHS* genes identified in the genome, out of which, 2 are not expressed according to RNAseq data in the Phytozome database and are therefore considered as pseudogenes (Glyma.05G153100 and Glyma.13G034300). There are three copies of *GmCHS3* and two copies of *GmCHS4* to make a total of 14 *GmCHS* candidates distributed on 8 different chromosomes. Six of the *GmCHS* genes are located on chromosome 8 (Table 2.1; Fig 2.1). Two of the three copies of *GmCHS3* are in the anti-sense strand and their promoter sequences are diverse (Table 2.2; Fig 2.1). Similarly, two copies of *GmCHS4* are in close proximity to each other on chromosome 8, but in opposite orientation (one in sense and the other one in the anti-sense strand). The promoter regions (1000 bp upstream of ATG) of all *GmCHS*s were analyzed and a pairwise sequence comparison between all candidate genes' promoters showed a range from 0.4% to 100% identical (Table 2.2). *GmCHS5* and *GmCHS12* share 100% identical promoter sequences. The promoter sequences of two copies of *GmCHS4* are 81.6% identical and the distance between the 5'UTR of *GmCHS4a* and 3'UTR of *GmCHS4b* is 9.4 Kb (Table 2.2; Fig 2.1).

CHSs have conserved residues of Cys 164, Phe 215, His 303 and Asn 336 as active sites (Ferrer et al., 1999). To evaluate if all candidates have these conserved active sites, amino acid sequences of all putative GmCHSs were retrieved from Phytozome database and aligned with a well characterized and crystallized CHS from alfalfa, MsCHS2. The active sites of the enzymes

are conserved among all 14 CHS from soybean (Fig 2.2). GmCHSs have ~389 amino acids and have a molecular weight of ~41.5 to 43 kDa and a PI of ~5.6 to 6.2 except for GmCHS12 which is 340 aa long and has a molecular weight of 37 kDa and a PI of 5.3.

A neighbour joining tree was produced from the amino acid sequences of 14 putative GmCHSs along with previously characterized CHS and STS from other plants (Fig 2.2). GmCHS7 and GmCHS8 together formed a separate clade as previously reported (Yi et al., 2010a). Both GmCHS7 and GmCHS8 could be active GmCHSs and formed a close clade with previously characterized PvCHS17 and MsCHS2. GmCHS14 was part of a distinct clade with MtCHS in the phylogenetic tree (Fig 2.3). STS from *Vitis riparia*, *Vitis vinifera*, and *Arachis hypogaea* formed a close clade separate from CHS from soybean and other species.

The putative *GmCHS*s identified in the soybean genome are listed in Table 2.1. Pairwise coding sequence comparison between all candidate genes showed a range from 47.5% to 100% identity whereas, pairwise amino acid sequence comparison yielded a range from 50.8% to 100% (Table 2.3). Even though amino acid sequences were 100% identical for three copies of *GmCHS3* genes, the three copies are referred as *GmCHS3a*, *GmCHS3b*, and *GmCHS3c*. Similarly, two copies of *GmCHS4* are referred as *GmCHS4a* and *GmCHS4b*.

2.3.2 *GmCHS* transcript accumulation vary in soybean tissues

The transcript abundance of *GmCHS* genes in soybean tissues (flower, shoot apical meristem, seed, pod, stem, root, nodules, leaves, and root hairs) were retrieved from publically available RNAseq data in Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012). Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values of *GmCHS*s in their highly expressed tissues varied from 9.17 to 599 (Table 2.1). Fig 2.4 shows that most of the *GmCHS*s were highly expressed in leaves except for *GmCHS13* (in flower), *GmCHS8* and *GmCHS11* (in root). *GmCHS7* was expressed highly in roots and root hairs, whereas *GmCHS10* was shown to be highly expressed in roots and leaves. The expression patterns of *GmCHS7* and *GmCHS8* were consistent with the previous studies (Yi et al., 2010). The expression patterns of three copies of *GmCHS3* were different in soybean tissues (Fig 2.4). The two copies of *GmCHS4* showed almost similar expression patterns in all soybean tissues. Most of the *GmCHS*s expression was confined to the leaves and roots.

Table 2.1 List of *GmCHS* genes identified in soybean genome

Gene name	Locus name	Gene location	CDS	Splice variants	Highly expressed in (FPKM)	Protein molecular weight (kDa)	Predicted subcellular localization
Gm*CHS1*	Glyma.08G109400	Chr08:8391364 - 8394840	1167	1	leaves (23.6)	42.5	cytoplasm
Gm*CHS2*	Glyma.05G153200	Chr05: 34687009 - 34693243	1167	1	leaves (30.29)	42.4	cytoplasm
Gm*CHS3a*	Glyma.08G109300	Chr08: 8387509 – 8391327	1167	1	leaves (12.9)	42.4	cytoplasm
Gm*CHS3b*	Glyma.08G110900	Chr08: 8517799 – 8519303	1167	2	leaves (20.9)	42.4	cytoplasm
Gm*CHS3c*	Glyma.08G110300	Chr08:8475793 - 8477410	1167	1	leaves (22.4)	42.4	cytoplasm
Gm*CHS4a*	Glyma.08G110700	Chr08: 8513952 - 8515719	1167	1	leaves (22.5)	42.5	cytoplasm
Gm*CHS4b*	Glyma.08G110500	Chr08: 8504479 – 8506020	1167	1	leaves (18.6)	42.5	cytoplasm
Gm*CHS5*	Glyma.08G109200	Chr08: 8384742 – 8386542	1167	1	leaves (23.4)	42.5	cytoplasm
Gm*CHS6*	Glyma.09G075200	Chr09: 8145494 – 8147595	1167	1	leaves (7.8)	42.5	cytoplasm
Gm*CHS7*	Glyma.01G228700	Chr01: 55659010 – 55660950	1170	1	Root (447)	42.8	cytoplasm
Gm*CHS8*	Glyma.11G011500	Chr11: 802453 – 804663	1170	2	Root (599)	42.8	cytoplasm
Gm*CHS9*	Glyma.08G109500	Chr08: 8397944 – 8399751	1167	1	leaves (33.9)	42.4	cytoplasm
Gm*CHS10*	Glyma.02G130400	Chr02: 13399253 – 13401493	1167	1	leaves (9.17)	42.5	cytoplasm
Gm*CHS11*	Glyma.01G091400	Chr01: 27621455 – 27623628	1167	1	Root (21.6)	42.4	cytoplasm
Gm*CHS12*	Glyma.08G110400	Chr08: 8478834 – 8480215	1023	1	leaves (12.8)	37	cytoplasm
Gm*CHS13*	Glyma.19G105100	Chr19: 35466392 - 35469297	1176	1	Flower (381)	43	cytoplasm
Gm*CHS14*	Glyma.06G118500	Chr06: 9644661 - 9650144	1170	1	leaves (60.5)	43	cytoplasm

Table 2.2 Promoter sequence (1000 bp) identity matrix of *GmCHS*s

	Promoter sequence (% identity)																
	CHS1	***CHS2***	***CHS3a***	***CHS3b***	***CHS3c***	***CHS4a***	***CHS4b***	***CHS5***	***CHS6***	***CHS7***	***CHS8***	***CHS9***	***CHS10***	***CHS11***	***CHS12***	***CHS13***	***CHS14***
CHS1		35.8	35.0	0.6	35.7	37.0	34.8	41.4	38.8	31.3	33.9	38.6	38.4	37.2	41.4	34.0	33.3
CHS2	35.8		39.3	2.4	41.2	40.6	37.6	38.9	39.3	31.0	31.8	37.1	33.9	38.4	38.9	32.6	31.9
CHS3a	35.0	39.3		2.6	48.9	38.4	35.2	35.0	37.4	29.8	30.4	35.2	33.3	35.8	35.0	29.6	30.
CHS3b	0.6	2.4	2.6		2.6	1.6	11.0	2.0	1.7	1.2	1.9	1.6	2.0	1.8	2.0	1.2	0.4
CHS3c	35.7	41.2	48.9	2.6		39.8	37.6	37.3	39.6	33.1	34.2	37.5	35.3	38.9	37.3	31.6	32.0
CHS4a	37.0	40.6	38.4	1.6	39.8		81.6	40.0	48.2	32.8	32.7	39.9	37.8	47.8	40.0	30.3	31.8
CHS4b	34.8	37.6	35.2	11.0	37.6	81.6		37.8	43.7	30.4	30.8	36.3	34.1	41.4	37.8	27.7	29.8
CHS5	41.4	38.9	35.0	2.0	37.3	40.0	37.8		40.9	34.1	36.7	45.8	40.7	40.3	100.0	33.2	30.5
CHS6	38.8	39.3	37.4	1.7	39.6	48.2	43.7	40.9		32.4	34.2	41.8	45.8	92.5	40.9	32.1	33.0
CHS7	31.3	31.0	29.8	1.2	33.1	32.8	30.4	34.1	32.4		44.0	32.8	29.3	30.9	34.1	31.9	30.8
CHS8	33.9	31.8	30.4	1.9	34.2	32.7	30.8	36.7	34.2	44.0		34.1	31.6	33.6	36.7	32.6	28.4
CHS9	38.6	37.1	35.2	1.6	37.5	39.9	36.3	45.8	41.8	32.8	34.1		51.7	40.8	45.8	32.6	32.0
CHS10	38.4	33.9	33.3	2.0	35.3	37.8	34.1	40.7	45.8	29.3	31.6	51.7		45.8	40.7	32.9	29.8
CHS11	37.2	38.4	35.8	1.8	38.9	47.8	41.4	40.3	92.5	30.9	33.6	40.8	45.8		40.3	30.6	31.7
CHS12	41.4	38.9	35.0	2.0	37.3	40.0	37.8	100.0	40.9	34.1	36.7	45.8	40.7	40.3		33.2	30.5
CHS13	34.0	32.6	29.6	1.2	31.6	30.3	27.7	33.2	32.1	31.9	32.6	32.6	32.9	30.6	33.2		31.1
CHS14	33.3	31.9	30.5	0.4	32.0	31.8	29.8	30.5	33.0	30.8	28.4	32.0	29.8	31.7	30.5	31.1	

Table 2.3 The CDS and protein sequence identity matrix of *GmCHS*s.

		Amino acid (% identity)																
		CHS1	CHS2	CHS3a	CHS4a	CHS5	CHS6	CHS7	CHS8	CHS9	CHS10	CHS11	CHS12	CHS13	CHS3b	CHS14	CHS3c	CHS4b
Nucleotide (% identity)	CHS1		99.2	99.4	99.4	99.2	97.1	89.9	91.0	98.9	96.6	97.4	86.0	85.1	99.4	51.3	99.4	99.4
	CHS2	94.8		99.4	99.2	99.4	96.9	90.2	91.2	98.7	96.3	97.1	86.3	84.9	99.4	51.3	99.4	99.2
	CHS3a	98.4	95.2		99.4	99.4	97.1	89.9	91.0	98.9	96.6	97.4	86.3	85.1	100.0	51.3	100.0	99.4
	CHS4a	98.2	95.1	97.7		99.7	97.1	89.7	90.7	98.9	96.6	97.4	86.5	85.1	99.4	51.3	99.4	100.0
	CHS5	98.0	95.3	97.8	99.7		96.9	89.9	91.0	98.7	96.3	97.1	86.8	85.1	99.4	51.3	99.4	99.7
	CHS6	92.8	92.9	92.4	92.6	92.5		89.7	90.2	96.6	97.4	99.2	83.7	83.8	97.1	52.4	97.1	97.1
	CHS7	81.3	81.0	81.4	80.4	80.6	80.2		98.7	89.2	89.4	89.9	77.3	82.8	89.9	51.9	89.9	89.7
	CHS8	69.0	68.4	68.8	68.2	68.2	67.7	82.0		90.2	89.9	90.4	78.4	83.1	91.0	51.9	91.0	90.7
	CHS9	97.3	95.2	97.5	97.5	97.6	92.8	81.0	68.6		96.1	96.9	85.8	84.6	98.9	50.8	98.9	98.9
	CHS10	91.4	91.8	91.2	91.2	91.1	95.2	80.3	67.9	91.5		97.6	83.5	83.6	96.6	51.9	96.6	96.6
	CHS11	93.1	93.2	92.7	92.8	92.7	99.5	80.3	67.7	93.0	95.4		84.0	84.3	97.4	52.4	97.4	97.4
	CHS12	85.5	83.2	85.6	87.1	87.4	80.5	70.2	58.0	85.0	79.4	80.7		73.4	86.3	43.0	86.3	86.5
	CHS13	73.0	73.2	72.9	72.8	72.8	72.1	70.8	59.6	73.3	70.7	72.2	63.1		85.1	52.9	85.1	85.1
	CHS3b	98.5	95.1	99.9	97.8	97.9	92.3	81.4	68.8	97.6	91.1	92.6	85.6	72.9		51.3	100.0	99.4
	CHS14	56.3	55.7	56.5	56.0	56.0	56.5	56.5	50.9	56.0	56.2	56.3	47.5	57.4	56.4		51.3	51.3
	CHS3c	98.5	95.1	99.9	97.8	97.9	92.3	81.4	68.8	97.6	91.1	92.6	85.6	72.9	100.0	56.4		99.4
	CHS4b	98.2	95.1	97.7	100.0	99.7	92.6	80.4	68.2	97.5	91.2	92.8	87.1	72.8	97.8	56.0	97.8	

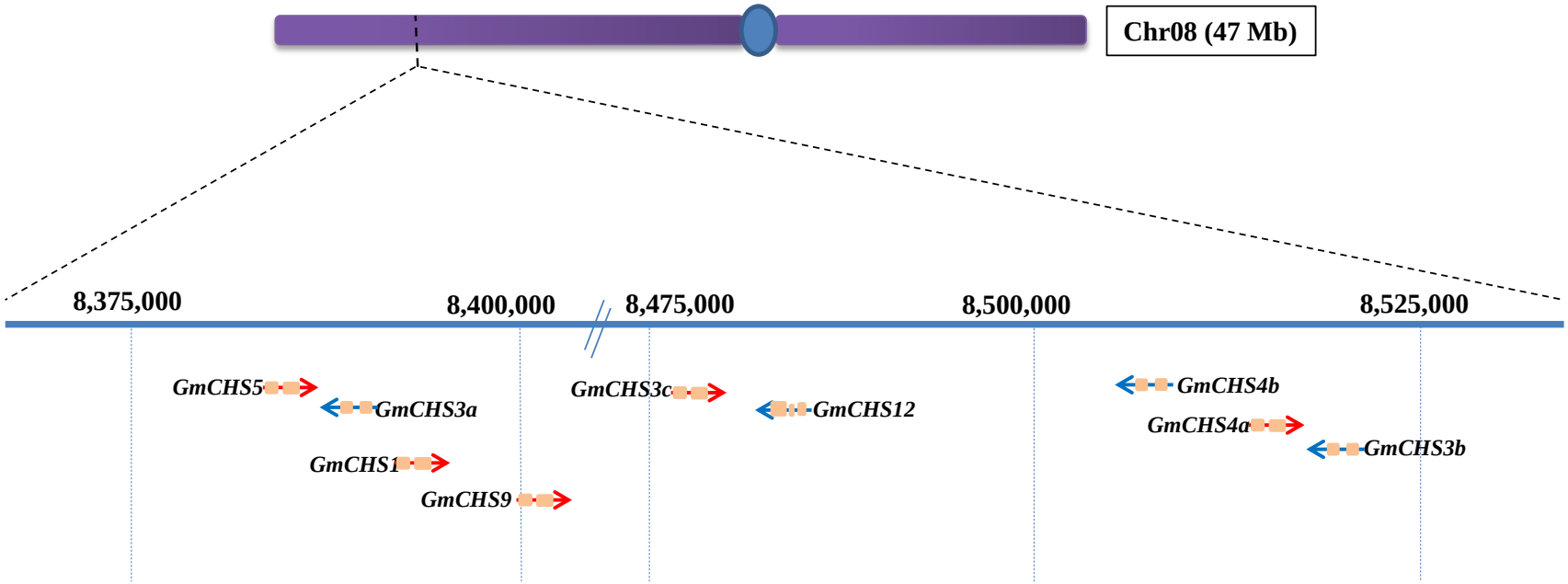


Figure 2.1 Genomic distribution of six *GmCHS* genes on chromosome 8 in soybean. Purple colored is the chromosome8; Blue oval is centromere; Red arrows show the orientation of gene in sense strand; Blue arrow shows the orientation of genes in anti-sense strand; Numbers on the chromosome are in bp units.

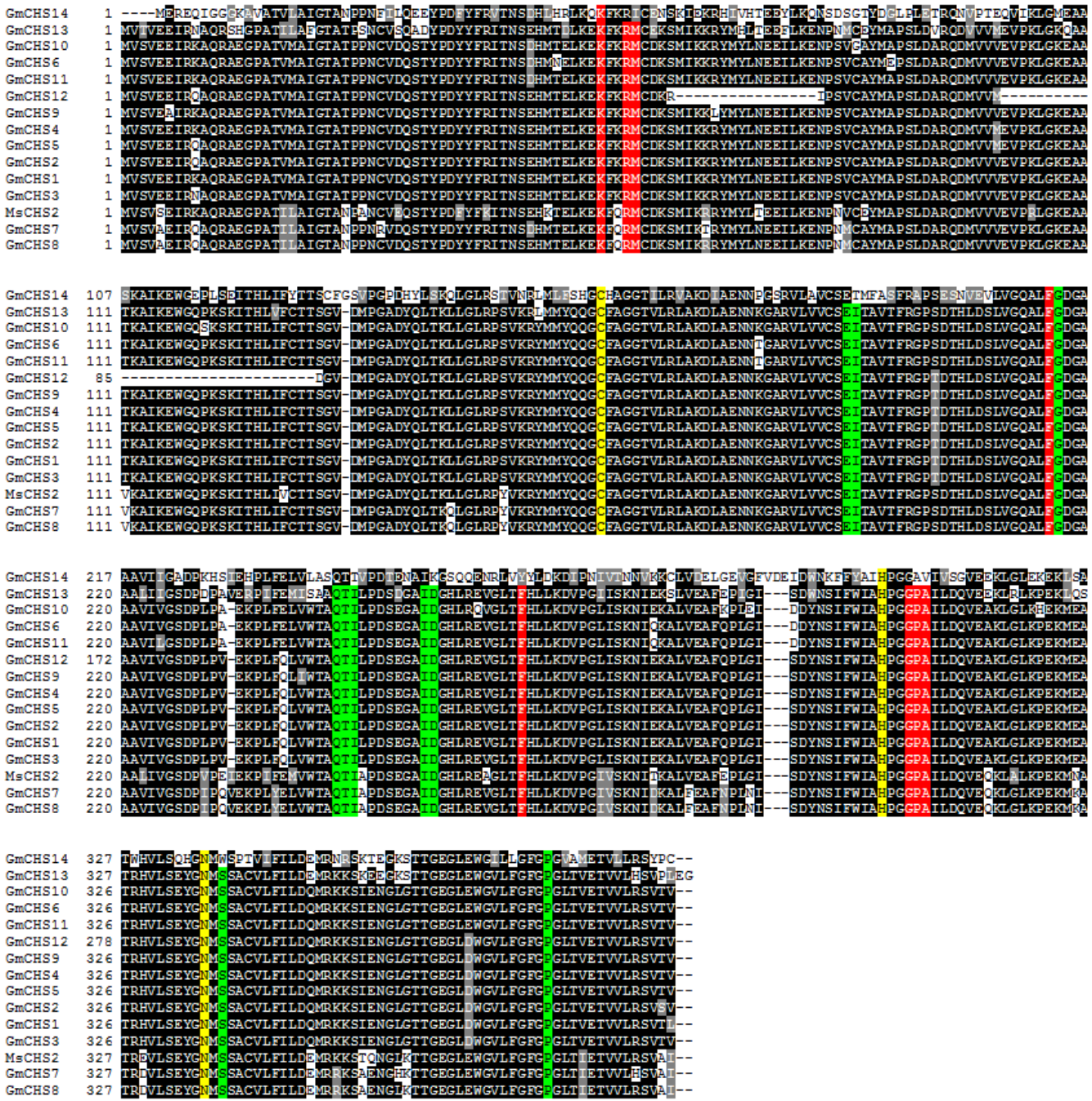

GmCHS14
GmCHS13
GmCHS10
GmCHS6
GmCHS11
GmCHS12
GmCHS9
GmCHS4
GmCHS5
GmCHS2
GmCHS1
GmCHS3
MsCHS2
GmCHS7
GmCHS8

Figure 2.2 Aminoacid alignments of GmCHSs. The alignment is shown in the figure. CHS2 from alfalfa (MsCHS2) was used as a reference CHS. Active sites, which are conserved in CHS and CHS-like proteins, are highlighted in yellow colour. Active sites: Cys 164, Phe 215, His 303 and Asn 336 (number represents its position in MsCHS2). MalonyCoA binding sites are highlighted in red; Product binding sites are highlighted in green. Identical residues are shown in black and similar residues are in grey. A hyphen indicates a gap.

Figure 2.3 Phylogenetic analyses of soybean chalcone synthases. The amino acid sequences of the CHSs from soybean were aligned with those from other plant species and the evolutionary tree was generated using the Neighbor-Joining method(Saitou and Nei, 1987). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test is shown next to the branches. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016). AtCHS (AT4G00040); OsCHS1 (A2ZEX7); MsCHS2 (P30074); MtCHS (KEH35360); MdCHS2 (AFX71920); PvCHS17 (P49440); VrSTS (AAF00586); VvSTS (AEB00549); AhSTS (1Z1E_A); XP_015644205 (OsFAS); At – *Arabidopsis thaliana*; Os - *Oryza sativa*; Ms - *Medicago sativa*; Mt – *Medicago truncatula*; Md – *Malus domestica*; Pv - *Phaseolus vulgaris*; Vr - *Vitis riparia*; Vv - *Vitis vinifera*; Ah - *Arachis hypogaea*. Scale bar indicates branch length representing residue substitution per site.

2.4 Discussion

2.4.1 Soybean genome contains 14 putative *GmCHS* genes

Multigene families in a plant genome can be a result of multiple factors such as gene duplications, whole genome duplications, and domestication. Plant genomes tend to evolve at higher rates, resulting in higher genome diversity and dynamics than in mammals (Murat et al., 2012). Recently, 14 *CHS* gene family members were identified in maize (Han et al., 2016). Similarly, 14 *GmCHS*s were identified in soybean (Table 2.1). There are three copies of *GmCHS3*, two copies of *GmCHS4* and two pseudogenes in the soybean genome that makes a total of 19 loci of *CHS* in the genome. Duplicated *CHS* genes have shown functional diversification in other species. Members of CHS gene family in other species showed functional variations and tissue-specific expression patterns, for example, among three *CHS* genes that showed different spatial and temporal regulation in *Gerbera hybrida*, only *GCHS1* contributing to flavonoid biosynthesis (Deng et al., 2014). All CHS and STS use the same substrate in the pathway and the catalytic active sites area consensus among these proteins. The activity of CHS/STS will not be known until enzyme activity assays are conducted. The folding of CHS/STS determines its catalytic activity.

CHS forms a homodimer for enzymatic activity. The CHS homodimer contains two functionally independent active sites. CoA-thioesters and product analogs occupy both active sites of the homodimer in the CHS complex structures. These structures identify the location of the active site at the cleft between the lower and upper domains of each monomer, where few chemically reactive residues are present in the active site (Ferrer et al., 1999; Austin et al., 2004).

There are four conserved amino acid residues, specifically Cys 164, Phe 215, His 303 and Asn 336, which form active sites in all CHS-related enzymes(Ferrer et al., 1999). Cys 164 serves as the nucleophile and as the attachment site for polyketide intermediates in both CHS and STS. The nitrogen electron of His 303 is within hydrogen-bonding distance of the sulfur atom of Cys 164 and so His 303 most likely acts as a general base during the generation of a nucleophilic thiolate anion from Cys 164. The active site architecture of CHS consists of three interconnected cavities intersect with these four residues and these cavities include a coumaroyl-binding pocket, CoA-binding tunnel, and a cyclization pocket (Ferrer et al., 1999).

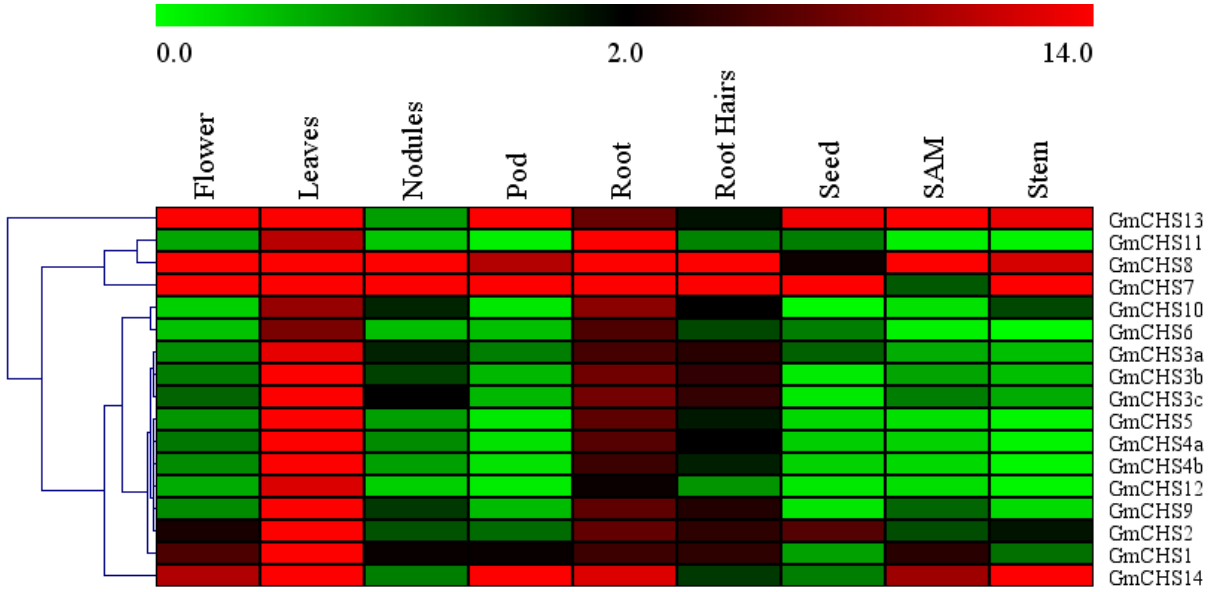


Figure 2.4 Expression levels of *chalcone synthases* in different tissues of soybean. The average-linkage hierarchical clustering with Pearson correlation distance was applied. The FPKM values of each gene were retrieved from Phytozome database. The color scale above represents expression values with green indicating low levels and red indicating high levels of transcript abundance. SAM- Shoot apical meristem.

GmCHS7 and GmCHS8 could be the active CHSs in soybean and shared the same clade with PvCHS17 and MsCHS2 in the phylogenetic tree (Fig 2.3). As CHS and STS are postulated to have evolved from FAS, OsFAS was used as an outgroup. GmCHS12 has fewer numbers of amino acids (340) and a lower molecular weight (37 kDa) than the other GmCHSs (~389aa; ~42 kDa) (Table 2.1). However, GmCHS12 contains conserved residues which should be tested for its activity *in vivo* to determine whether GmCHS12 has the CHS activity or CHS-like protein such as STS activity. Though STS and CHS use the same substrate, their unique tertiary structure results in different end products. The aldol condensation and claisen condensation reactions differentiate these two enzymes (Austin et al., 2004).

2.4.2 *GmCHS*s display tissue-specific expression in soybean

Gene family members containing different promoter regions result in differential expression patterns within a species. The potential *cis*-elements inthe promoter regions can differ among promoters and undergo changes in response to developmental or environmental regulation (Dey et al., 2015). Members of *CHS* gene family in other plant species showed functional variations and tissue-specific expression patterns (Deng et al., 2014). *GCHS3* and *GCHS4* showed different spatial and temporal expression in *Gerbera hybrida*. *GCHS4* was found to be regulated by an MYB transcription factor, GMYB10. The expression of *GCHS4* was upregulated in the tissue overexpressing *GMYB10* (Laitinen et al., 2008). The possibility remains that differential regulation of different putative *GmCHS* genes is due to diverse *cis*-acting elements in their promoters and *trans*-acting factors that activate the promoters. The differential expression of gene family members shows the dynamic of plant pathway regulation.

Most of the *GmCHS*s are highly expressed in soybean leaves and roots suggestive of their importance in these tissues. The expression of these genes in soybean roots is highly important since downstream of the CHS-catalyzed step is the production of isoflavonoids that participate in plant defense mechanisms and also in the symbiotic relationship between soybean and bacteria for nitrogen fixation. The expression of *GmCHS7* and *GmCHS8* in soybean tissues have already been studied (Yi et al., 2010). The high expression levels of *GmCHS7* and *GmCHS8* in roots (Fig 2.4) is consistent with Yi et al. (2010). Most GmCHSs were expressed in soybean leaves and roots which could explain the requirement of these genes in the respective tissues for (iso)flavonoid biosynthesis (Fig 2.4). Diverse expression of *GmCHS* genes in soybean tissues

may be due to their diverse promoters regions except for *GmCHS5* and *GmCHS12* as their promoters are 100% identical (Table 2.2). Identical promoter regions with conserved *cis*-regulatory elements could be a result of segmental duplication and it has been observed previously among certain duplicated genes (Haberer et al., 2004). Gene family members showing diverse gene expression in soybean have been documented. For example, soybean 14-3-3 protein (*SGF14*s) (Li and Dhaubhadel, 2011) and chalcone isomerase (*GmCHIs*) (Dastmalchi and Dhaubhadel, 2015) family members also display differential expression patterns in soybean tissues. The spatio-temporal expression of *CHS* genes may important for the production of flavonoids, isoflavonoids, and their derived phytoalexins.

2.5 **Conclusion**

In this study, with the advantage of soybean genome sequence availability, I identified 14 *GmCHS* genes (5 new *GmCHS* genes) including three additional copies of *GmCHS3* and two additional copies of *GmCHS4* in soybean. Their CDS and protein sequences are highly conserved but, diverse in their promoter sequences except for *GmCHS5* and *GmCHS12*. *GmCHS* gene family members show diverse tissue-specific expression patterns in soybean.

2.6 Literature cited

Austin, M.B., and Noel, J.P. (2003). The chalcone synthase superfamily of type III polyketide synthases. Natural Product Reports**20,** 79-110.

Austin, M.B., Bowman, M.E., Ferrer, J.L., Schroder, J., and Noel, J.P. (2004). An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. Chemistry and Biology**11,** 1179-1194.

Blanc, G., and Wolfe, K.H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell **16,** 1667-1678.

Dastmalchi, M., and Dhaubhadel, S. (2015). Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. Planta **241,** 507-523.

Deng, X., Bashandy, H., Ainasoja, M., Kontturi, J., Pietiainen, M., Laitinen, R.A., Albert, V.A., Valkonen, J.P., 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.

Dey, N., Sarkar, S., Acharya, S., and Maiti, I.B. (2015). Synthetic promoters in planta. Planta **242,** 1077-1094.

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 and Molecular Biology **6,** 775-784.

Gill, N., Findley, S., Walling, J.G., Hans, C., Ma, J., Doyle, J., Stacey, G., and Jackson, S.A. (2009). Molecular and chromosomal evidence for allopolyploidy in soybean. Plant Physiology **151,** 1167-1174.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research **40,** D1178-1186.

Haberer, G., Hindemitt, T., Meyers, B., and Mayer, K. (2004). Transcriptional similarities, dissimilarities, and conservation of cis-elements in duplicated genes of *Arabidopsis*. Plant Physiology **136**, 3009-3022.

Han, Y., Ding, T., Su, B., and Jiang, H. (2016). Genome-wide identification, characterization and expression analysis of the *Chalcone synthase* family in maize. International Journal of Molecular Science **17,** 161.

Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. Angewandte Chemie International Edition in English **48,** 4688-4716.

Hopwood, D.A., and Sherman, D.H. (1990). Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annual Review of Genetics**24,** 37-66.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution**33,** 1870-1874.

Laitinen, R.A., Ainasoja, M., Broholm, S.K., Teeri, T.H., and Elomaa, P. (2008). Identification of target genes for a MYB-type anthocyanin regulator in *Gerbera hybrida*. Journal of Experimental Botany **59,** 3691-3703.

Li, X., and Dhaubhadel, S. (2011). Soybean 14-3-3 gene family: identification and molecular characterization. Planta **233,** 569-582.

Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., and Ravikesavan, R. (2013). Gene duplication as a major force in evolution. Journal of Genetics**92,** 155-161.

Murat, F., Van de Peer, Y., and Salse, J. (2012). Decoding plant and animal genome plasticity from differential paleo-evolutionary patterns and processes. Genome Biology and Evolution**4,** 917-928.

Panchy, N., Lehti-Shiu, M.D., and Shiu, S.-H. (2016). Evolution of gene duplication in plants. Plant Physiology **171,** 2294-2316.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution**4,** 406-425.

Schlotterer, C. (2015). Genes from scratch--the evolutionary fate of de novo genes. Trends in Genetics **31,** 215-219.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., and Jackson, S.A. (2010). Genome sequence of the palaeopolyploid soybean. Nature **463,** 178-183.

Shoemaker, R.C., Polzin, K., Labate, J., Specht, J., Brummer, E.C., Olson, T., Young, N., Concibido, V., Wilcox, J., Tamulonis, J.P., Kochert, G., and Boerma, H.R. (1996). Genome duplication in soybean (*Glycine* subgenus *soja*). Genetics **144,** 329-338.

Shomura, Y., Torayama, I., Suh, D.Y., Xiang, T., Kita, A., Sankawa, U., and Miki, K. (2005). Crystal structure of stilbene synthase from *Arachis hypogaea*. Proteins **60,** 803-806.

Tropf, S., Karcher, B., Schroder, G., and Schroder, J. (1995). Reaction mechanisms of homodimeric plant polyketide synthase (stilbenes and chalcone synthase). A single active

site for the condensing reaction is sufficient for synthesis of stilbenes, chalcones, and 6'-deoxychalcones. Journal of Biological Chemistry **270,** 7922-7928.

True, J.R., and Carroll, S.B. (2002). Gene co-option in physiological and morphological evolution. Annual Review of Cell and Developmental Biology**18,** 53-80.

Tuteja, J.H., and Vodkin, L.Q. (2008). Structural features of the endogenous *CHS* silencing and target loci in the soybean genome. Crop Science **48,** S49-S68.

Van de Peer, Y., Maere, S., and Meyer, A. (2009). The evolutionary significance of ancient genome duplications. Nature Reviews Genetics **10,** 725-732.

Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V., and Romano, L.A. (2003). The evolution of transcriptional regulation in eukaryotes. Molecular Biology and Evolution**20,** 1377-1419.

Yi, J., Derynck, M.R., Chen, L., and Dhaubhadel, S. (2010a). Differential expression of *CHS7* and *CHS8* genes in soybean. Planta **231,** 741-753.

Yi, J., Derynck, M.R., Li, X., Telmer, P., Marsolais, F., and Dhaubhadel, S. (2010b). A single-repeat MYB transcription factor, GmMYB176, regulates *CHS8* gene expression and affects isoflavonoid biosynthesis in soybean. The Plant Journal **62,** 1019-1034.

3 GmMYB176 INTERACTOME: KEY FACTORS IN ISOFLAVONOID BIOSYNTHESIS IN SOYBEAN

3.1 Introduction

MYB TFs are classified within the HTH TF family that shares a conserved DNA-binding domain. The first *MYB* gene, *C-MYB*, was identified as an oncogene in avian myeloblastosis virus (Craig and Bloch, 1984). C-MYB has three repeats R1, R2, and R3. Based on their similarity to the repeats of C-MYB, MYB TFs are classified into four groups: R1 MYB, R2R3 MYB, R3 MYB, and R4 MYB. R1 MYB is a single-repeat MYB TF that has either R1 or R3 repeat (reviewed in Dubos et al., 2010). Regularly spaced tryptophan residues in each of the repeat forms a hydrophobic core that stabilizes the DNA-binding activity (Ogata et al., 1996). MYBs, belonging to the largest TF family in plants, and regulate many aspects of plant metabolism during the development in Arabidopsis, grapevine, apple, petunia, maize, and rice (reviewed in Du et al., 2009; Fernández-Marcos et al., 2016).

The first plant *MYB* gene, *C1*, was identified in *Zea mays* and was shown to be involved in anthocyanin biosynthesis (Paz-Ares et al., 1987). MYB proteins have been reported to play a vital role in the regulation of the phenylpropanoid pathway in plants by either activating or suppressing the target gene expression. For example, Y1 an R2R3 MYB protein, regulates the biosynthesis of 3-deoxyflavonoids by regulating the expressions of *CHS*, *CHI*, and *DFR* genes in sorghum (Boddu et al., 2006). In maize, ZmMYBC1 regulates anthocyanin biosynthesis by activating upstream genes in the flavonoid biosynthetic pathway (Jin and Martin, 1999). Similarly, TaMYB14 has been confirmed to affect proanthocyanin biosynthesis in *Trifolium arvense* (Hancock et al., 2012). Some MYB proteins can negatively regulate genes involved in the phenylpropanoid pathway. For example, overexpression of VvMYBC2-L1 in grapevine hairy roots showed a reduction in accumulation of proanthocyanidin (Huang et al., 2014). Similarly, overexpression of GbMYBF2 from *Ginkgo biloba* in transgenic *Arabidopsis* down regulated the expression of the flavonoid biosynthetic pathway genes and thus affecting flavonoids and anthocyanins accumulation in transgenic *Arabidopsis* (Xu et al., 2014). Overexpression of an

R2R3 MYB TF, GmMYB100 negatively regulate the expressions of *GmCHS7* and *GmCHS8* genes in soybean and thus downregulates the flavonoid biosynthesis (Yan et al., 2015).

GmMYB176 (Glyma.05G032200) is a member of the MYB TF superfamily in soybean. GmMYB176 belongs to the R1 MYB group and regulates the isoflavonoid biosynthesis by activating *GmCHS8* gene expression (Yi et al., 2010). Isoflavonoids are a group of legume-specific plant natural compounds that are derived from central flavanone intermediates naringenin and liquiritigenin, which in turn are derived from tetrahydroxychalcone (naringenin chalcone) and trihydroxychalcone (isoliquiritigenin chalcone), respectively. The enzyme CHS is involved in the condensation of *p*-coumaroyl-CoA with three acetate moieties, derived from malonyl-CoA, to form naringenin chalcone, and is the first step in the branched pathway for the synthesis of flavonoids and isoflavonoids (Hahlbrock and Scheel, 1989). The soybean genome contains nine *GmCHS* genes (*GmCHS1–GmCHS9*) (Tuteja and Vodkin, 2008). All *CHS* genes share a high degree of sequence similarity within and across species. However, they play various roles during plant development or in response to environmental stimuli. Comparison of global gene expression in two soybean cultivars that differ in seed isoflavonoid level revealed that *GmCHS7* and *GmCHS8* transcript abundance are directly associated with isoflavonoid level in soybean seeds (Dhaubhadel et al., 2007).

RNAi silencing of *GmMYB176* leads to a decrease in *GmCHS8* gene expression and isoflavonoid levels, which demonstrates the importance of GmMYB176 in isoflavonoid biosynthesis (Yi et al., 2010). Interestingly, the overexpression of GmMYB176 does not affect either the expression level of *GmCHS8* or isoflavonoid content in soybean hairy roots. In eukaryotes, transcriptional regulation is often mediated by multi-protein complexes or the concerted action of several proteins. Such proteins are part of an interactome where members of the complex may bind DNA directly or may facilitate the interaction of other proteins in the complex. MYB TFs have been found to form protein complexes with basic helix loop helix (bHLH) TF and WD40 for phenylpropanoid gene regulation. For examples, AtMYB0 forms a protein complex with AtbHLH1 in *Arabidopsis* (Goff et al., 1992), whereas, in maize, ZmMYBC1 forms a protein complex with ZmR for the regulation of anthocyanin biosynthesis (Stracke et al., 2001). Since the overexpression of GmMYB176 does not alter *GmCHS8* expression and isoflavonoid level in

soybean hairy roots (Yi et al., 2010), I have speculated that GmMYB176 alone is insufficient for *GmCHS8* gene activation and that it requires other factor(s) for its function.

Here I performed co-immunoprecipitation (Co-IP) using GmMYB176 as the bait followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) to identify a number of putative GmMYB176 interactors. Thereafter, the candidate proteins whose orthologs or paralogs have been shown to function in transcriptional regulation or phenylpropanoid metabolism were validated for their protein-protein interaction (PPI) with GmMYB176 *in planta*. My results demonstrate that GmMYB176 works in combination with bZIP TFs (Gm04bZIP or Gm05bZIP) for the regulation of *GmCHS8* gene expression. The characterization of GmMYB176 interactome will provide a broader knowledge of the regulation of isoflavonoid biosynthesis in soybean.

3.2 Materials and methods

3.2.1 Plant materials and growth conditions

Nicotiana benthamiana Domin seeds were sprinkled onto wet Pro-Mix BX Mycorrhizae^{TM} soil (Riviere-du-Loup, Canada). The seeds were germinated in a growth chamber with a 16 h light cycle at 25ºC, and an 8 h dark cycle at 20ºC, with 60-70% relative humidity, with light intensity 100-150 mol $\mu m^{-2}s^{-1}$. After two weeks, seedlings were transferred to new pots and watered regularly. A nutrient mixture of nitrogen, phosphorous, and potassium (20-8-20) was applied to the plants twice a week. To obtain soybean cotyledons for hairy root transformation, soybean (*Glycine max* L. Merr.) cv. Harosoy63 (H63) seeds were surface sterilized with 70% ethanol (v/v) containing 3% H_2O_2 (v/v) for 2 minutes and then rinsed with sterile water prior to planting. Seeds were planted in soaked autoclaved vermiculite and grown for 6 days in a growth chamber with a 16 h light cycle at 25ºC, and an 8 h dark cycle at 20ºC, with 60-70% relative humidity.

3.2.2 Plasmid construction

For the PPI study, genes of interest (GOI) were PCR amplified, from cDNAs that were prepared from RNA of different soybean tissues, using attB1 and attB2 site-containing Gateway primers (Table 3.1). The PCR products were separated on 1% (w/v) agarose gel in 0.5X TBE buffer. A single amplicon corresponding to the gene of interest was excised from the gel and DNA was

extracted from the gel using EZ-10 Spin Column DNA Gel Extraction Kit (BioBasicInc, Canada), quantified using a NanoDrop spectrophotometer (ThermoScientific, USA), and recombined with an entry vector pDONR-Zeo (Invitrogen, USA) using BP clonase reaction mix (Invitrogen, USA) according to Gateway cloning technology. The BP reaction product was transformed into *E. coli* DH5α cells by electroporation, and cells were grown on Luria-Bertani (LB) media supplemented with zeocin (50 µg/mL). Five bacterial colonies were screened by PCR, using gene-specific primers, to select colonies carrying the recombinant plasmids. Plasmid DNA was prepared from positive colonies using EZ-10 Spin Column Plasmid DNA Kit (BioBasicInc., Canada) and sequence verified. The pDONR/Zeo-GOI plasmids were used in an LR recombination reaction with the destination vectors, pEarleyGate201-YN and pEarleyGate202-YC (Lu et al., 2010) (Invitrogen, USA), transformed as above into *E. coli* DH5α, plasmid extracted, and then transformed into *Agrobacterium tumefaciens* strain GV3101 by electroporation.

For Co-IP, the recombinant pDONR/Zeo-GmMYB176 plasmid was used in an LR recombination reaction with the destination vectors, pEarleyGate101 and pEarleyGate104 vectors (Earley et al., 2006), using Gateway technology (Invitrogen, USA) to obtain pEG101-GmMYB176-YFP and pEG104-YFP-GmMYB176, respectively. The plasmids were transformed as above into *E. coli* DH5α cells, plasmid extracted, and then transformed into *Agrobacterium rhizogenes* K599. For yeast one-hybrid (Y1H) assays, a 30 bp region (containing a 23 bp critical sequence) and *CHS8*Δ*P5* fragments of the *GmCHS8* promoter were chosen as baits (Yi et al., 2010). Both the promoter fragments were purified using the EZ-10 spin column PCR purification kit (BioBasicInc, Canada) for cloning into a pABAi vector. The 30bp tandem repeat (30bpTR) fragments, *CHS8*Δ*P5*, and pAbAi were digested with HindIII and XhoI sequentially and purified using the EZ-10 spin column PCR purification kit (BioBasicInc, Canada). The digested promoter fragments and vector were ligated using DNA ligase (Invitrogen,USA) and transformed into *E. coli* DH5α cells by electroporation to obtain p30bpTR-AbAi and pCHS8ΔP5-AbAi. Five positive colonies were screened by PCR amplification, plasmid extracted, and sequence verified. Promoter fragments, *proCHI1B1* (Dastmalchi and Dhaubhadel, 2015b), *proGmCYP1* (Mainali, 2014), and *CHS8*Δ*P1* (Yi et al., 2010) were amplified using the gene-specific primers (Table 3.1) and cloned into a pAbAi vector as described above and transformed into yeast Y1HGold strain to

integrate the promoter bait fragments into yeast genome. The pDONR/Zeo-GOI plasmids 'entry clone' was used in an LR recombination reaction with the destination vector, pGADT7 using Gateway technology (Invitrogen, USA). The recombinant plasmid, pGADT7-GOI was used as preys in Y1H assays.

3.2.3 Generation of soybean hairy roots

Six-day-old soybean cotyledons were harvested for hairy root generation following Subramanian et al. (2005). The *A. rhizogenes* K599 containing pEG101-GmMYB176-YFP or pEG104-YFP-GmMYB176 was transformed into soybean cotyledons. Transgenic hairy roots were selected 20-30 days post-inoculation, using a Leica MZ FL III fluorescence stereo microscope with a YFP filter (excitation 510/20 nm; barrier filter 560/40 nm). The selected transgenic hairy roots were flash frozen and stored at -80ºC until use.

3.2.4 Subcellular localization and Bimolecular Fluorescence Complementation (BiFC)

To study the PPI of fusion proteins, the constructs in *A. tumefaciens* strain GV3101 were transformed into *N. benthamiana* leaves by infiltration (Sparkes et al., 2006). Briefly, a single colony of *A. tumefaciens* GV3101 was used to inoculate 'infiltration culture' media (LB broth containing 10mM 2-*N*-morpholinoethanesulfonic acid [MES] pH 5.6, and 100 µM acetosyringone) supplemented with kanamycin (50 µg/mL), rifampicin (10 µg/mL), and gentamycin (50 µg/mL) and grown at 28ºC with shaking at 225rpm until OD_{600} reached 0.5-0.8. The culture was centrifuged in a microfuge tube at 775x *g* for 30 minutes at room temperature and the pellet was washed with infiltration culture media to remove the residual antibiotics. The pellet was re-suspended in Gamborg's solution (3.2 g/L Gamborg's B5 and vitamins, 20 g/L sucrose, 10 mM MES pH5.6, and 200 µM acetosyringone) to a final OD_{600} of 1.0 and incubated at room temperature for at least one hour with gentle agitation to activate the virulence gene required for transformation. To verify PPI, pEarleyGate201-geneA constructs were co-transformed, in a 1:1 mixture, with pEarleyGate202-geneB. To determine subcellular localization, *A. tumefaciens* strain GV3101 containing pEG101-GmMYB176-YFP and pEG104-YFP-GmMYB176 were used.

The leaves of *N. benthamiana* plants (4-6 week old) were infiltrated with the bacterial culture using a 1 mL syringe. Bacteria were slowly injected into the abaxial side of the leaf. The infiltrated leaves were labeled and plants were returned to the growth room at normal growth condition as described in section 3.2.1. Confocal microscopy was carried out after 48 hours using a Leica TCS SP2 inverted confocal microscope. The excitation and emission of YFP were 514 nm and 530-560 nm, respectively. Leaves infiltrated with empty vectors were used as negative controls in the experiment.

3.2.5 **Co-IP**

Proteins were extracted from 1 g of soybean hairy roots in a lysis buffer containing 50 mM HEPES (pH7.5), 150 mM NaCl, 0.5% Triton X-100 (v/v), 0.1% Tween-20 (v/v) and protease inhibitor cocktail (Sigma, USA). Co-IP was conducted using µMACS epitope tag protein isolation kit (Miltenyi Biotech Inc., USA). Briefly, 2 mL of crude protein extract was incubated on ice with 100 µL of anti-GFP micro beads for 30 min. The samples were then passed through a µcolumn provided in the kit, washed with lysis buffer ten times and eluted with 50 µL of 0.1 M Na_2CO_3 (pH11.3) in a tube containing 5 µL of 1 M MES (pH3.0). The eluate was quantified using the Bradford assay (Bradford, 1976).

3.2.6 **SDS-PAGE, *in gel* digestion and LC-MS/MS**

The protein eluates (1 µg) were separated by SDS-PAGE and silver stained with ProteoSilver kit (Sigma, USA). Protein bands were excised from gels and destained by using 30 mM $K_3[Fe(CN)_6$ and 100 mM $Na_2S_2O_3$ solution. Destained protein bands were subjected to *ingel* digestion with trypsin, quick frozen in liquid nitrogen and then ground to extract peptides using a buffer containing 20 mM Tris-HCl (pH7.5) and 150 mM NaCl. *In gel* digestion was performed using a MassPREP automated digester station (PerkinElmer, USA). The *in gel* digested samples were centrifuged at 10,000x *g* for 15 minutes and the supernatant was analyzed by LC-MS/MS, Waters nanoacquity UPLC coupled with ThermoOrbitrap Elite ETD at the Biological Mass Spectrometry Laboratory (London Regional Proteomics Centre, Western University). MS/MS data were analyzed by MassLynx 4.1 software with Mascot (http://www.matrixscience.com)and compared against the soybean protein database extracted from Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) (Schmutz et al., 2010; Goodstein et al., 2012),

with parameters of monoisotopic peptide masses adopted, mass window ranged from 1000 to100 000 Da, mass tolerance set as 50 ppm and allowance of one missed cleavage.

3.2.7 **Western blot**

For Western blot analysis, crude extracts (5 µg), flow through (5 µg), wash samples (20 µL), and eluate proteins (1 µg) were separated on 10% SDS-PAGE (w/v) gel and electro-blotted onto a PVDF membrane by using a semi-dry electrophoretic transfer apparatus (Bio-Rad, Canada) following manufacturer's instructions. Translational fusions of GmMYB176 with YFP were identified by sequential incubation of the blot with the anti-GFP monoclonal antibody (Clontech, USA) at a dilution of 1:3000 followed by the HRP-conjugated goat anti-mouse secondary antibody (Pierce, USA) at a dilution of 1:5000. HRP detection was performed by using Super Signal West Femto (Thermo Scientific, Canada) following the manufacturer's instructions.

3.2.8 **Y1H assay**

The recombinant plasmids p30bpTR-AbAi, pCHS8ΔP5-AbAi, pProGmCYP1-AbAi, pProCHI1B1-AbAi, and pCHS8ΔP1-AbAi were transformed into yeast (Y1HGold strain) using the Yeastmaker™ Yeast Transformation System 2 (Clontech,USA) to integrate promoter fragments into the yeast genome. Thereafter, the transformants were cultivated in culture plates on SD/-Ura media at 30°C for 3 days. The transformed colonies were PCR screened after 3 days of incubation at 30°C using Matchmaker Insert check PCR mix1 (Clontech, USA). The positive colonies were used for competent cell preparation and Y1H assays with prey candidates (pGADT7-GOI). Y1H assays were performed by following the Matchmaker Y1H user manual (Clontech, USA).The minimum inhibitory concentration of aeurobasidin A for yeasts carrying *CHS8*Δ*P5*, *ProGmCYP1*, or *proCHI1B1* promoter bait was 200 ng/mL; an inhibitory concentration of 150 ng/mL was implemented for yeasts carrying *30bpTR* or *CHS8*Δ*P1*. The prey constructs (pGADT7-GOI) were transformed into yeast carrying a promoter bait fragment, plated on SD/-Leu/AbA250, and incubated at 30ºC for 3-5 days.

Table 3.1 Sequence of oligonucleotides used for gene/promoter amplification

Gene/ promoter	Primer name	Sequence (5'-3')	Amplicon size (bp)
GmMYB176	GmMYB176F	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGTCTCGCGCCTCTTCCGCC	855
	GmMYB176R	GGGGACCACTTTGTACAAGAAAGCTGGGTCGGGTCAAGCAACACTAATGATGCTATC	
Gm04bZIP	Gm04bZIP-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcCCCCCAAAACCCAAAACGAT	906
	Gm04bZIP-R	GGGGACCACTTTGTACAAGAAAGCTGGGTcAAAGGCATTGTTAATATTATTAACAGAAATG	
Gm05bZIP	Gm05bZIP-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcCTCCTGCGAGTGTTGTTTGG	732
	Gm05bZIP-R	GGGGACCACTTTGTACAAGAAAGCTGGGTcCAATGATTGTGCTGAATTTACTCTCCG	
Gm07bHLH	Gm07bHLHb-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcTCGCAACTCGAACTTGTCCACCG	1038
	Gm07bHLHb-R	GGGGACCACTTTGTACAAGAAAGCTGGGTcGCACAAACAAGATCCGTGGTTGCATGAGG	
Gm05bHLH	Gm05bHLH-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcAGTACAGGAGGAGACAGGCCA	1548
	Gm05bHLH-R	GGGGACCACTTTGTACAAGAAAGCTGGGTcATAGCTGTTTGTATGTGGAACTGCTGC	
Gm13bHLH/ Gm15bHLH	Gm13bHLH-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcCATATAGCTAGGTCCCCCACCCC	924
	Gm13bHLH-R	GGGGACCACTTTGTACAAGAAAGCTGGGTcTTGTTGCACATCACATGACTTCATC	
Gm19DoF	Gm19Dof-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcGTGCAACACAACACCAGCAAC	1470
	Gm19Dof-R	GGGGACCACTTTGTACAAGAAAGCTGGGTcTAAGCTCTCCTGAAATGCATGAGCACG	
Gm02ZF	Gm02ZFF	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcTTCCCCACAATTAGAGAGATCATCA	762
	Gm02ZFR	GGGGACCACTTTGTACAAGAAAGCTGGGTcCATCCTTCTTAGAGAACTTCTTG	
Gm17WD40	WD40Gm17F	GGGGACAAGTTTGTACAAAAAAGCAGGCTtcGGCCGTAGCTCATAAAGCCC	2037

	WD40Gm17R	GGGGACCACTTTGTACAAGAAAGCTGGGTcGCTTTCATCATCAGTACTTATTTCAGA	
CHS8ΔP5	CHS8P5-Y1HF	CGGAAGCTTTGATTGAATAGCGTGA	828
	CHS8P5-Y1HR	GCCTCGAGCTTTCCTTCAAATTAAG	
proCHI1B1	1B1pro-Y1HF	CGGAAGCTTACGGCACGAAGCGTTGATCT	481
	1B1pro-Y1HR	CACTCGAGATGGCACGCACGAACACCAA	
CHS8ΔP1	CHS8P1-Y1HF	CGGAAGCTTTTGGTGCATGCACGTGATGATAC	186
	CHS8pro-	CACTCGAGCGGCTTTCCTTCAAATTAAGTGAT	
30bpTR	Sense strand	CGGAAGCTTGCGTGAAAATATAGTTAGTATATGCACGTAGCGTGAAAATATAGTTAGTATATG CACGTAGCGTGAAAATATAGTTAGTATATGCACGTACTCGAGTG	Annealed size -107
	Antisense strand	CACTCGAGTACGTGCATATACTAACTATATTTTCACGCTACGTGCATATACTAACTATATTTTC ACGCTACGTGCATATACTAACTATATTTTCACGCAAGCTTCCG	

3.3 Results

3.3.1 Identification of GmMYB176 interacting proteins

To identify the GmMYB176 interactome in soybean, translational fusions of GmMYB176 with YFP were created where YFP was used as a tag in Co-IP experiments to obtain GmMYB176 interactors. Two separate fusions of GmMYB176 with YFP at both C-terminus (GmMYB176-YFP) and N-terminus (YFP-GmMYB176) were overexpressed in soybean hairy roots and subcellular localization of the fusion protein was determined. Despite the YFP tag position, both GmMYB176-YFP and YFP-GmMYB176 were localized to the nucleus and the cytoplasm (Fig. 3.1).

GmMYB176 interacting proteins from soybean hairy roots overexpressing either GmMYB176-YFP or YFP-GmMYB176 were precipitated in two separate Co-IP experiments. As predicted, many other proteins co-eluted with the fusion protein (Figure 3.2a). The presence of bait in the crude protein sample and in the eluate was confirmed by Western blot analysis (Fig 3.2b). Although the expected size of the GmMYB176-YFP fusion protein was seen in the Western blot (Fig 3.2b), some of the fusion proteins appeared to be cleaved. Similarly, the YFP-GmMYB176 fusion protein was also verified by SDS-PAGE and Western blot. A total of 802 proteins were identified in the eluate of all three replicates for both GmMYB176-YFP and YFP-GmMYB176 fusion protein baits. Soybean hairy root proteins were shown to interact with YFP, and co-eluate with it in Co-IP experiment (Dastmalchi et al., 2016). Therefore, to obtain GmMYB176-specific interacting candidates and to remove non-specific YFP interactors, I obtained the list of YFP interacting proteins from Dastmalchi et al. (2016) were subtracted from the list containing GmMYB176-YFP and YFP-GmMYB176 interacting proteins (Fig 3.3). This process yielded a total of 716 candidates from both GmMYB176-YFP and YFP-GmMYB176 fusion protein baits, which were then compared against the soybean database (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012) to get annotations (Appendix 1). Among them, 105 proteins were found in both GmMYB176-YFP and YFP-GmMYB176 fusion baits (Fig 3.3 and Appendix 1). GmMYB176 interacting candidates identified solely with GmMYB176-YFP or YFP-GmMYB176 were included in the study as they could have been missed due to the position of protein fusion. Hence, all 716 GmMYB176

interactors were chosen for further analysis. The biological activity and domain enrichment of 716 candidates were retrieved from GO annotation provided for the soybean genome (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012) and grouped into the categories based on their biological process, cellular component and molecular function (Fig 3.4). Four proteins were annotated as belonging to the flavonoid biosynthetic pathway category under biological process (Fig 3.4a). The pathway enrichment analysis using soybean identifiers in PhytoMine database revealed that phenylpropanoid biosynthesis is enriched but with a low p-value (Appendix 2). The enzymes that were enriched in the Co-IP experiment and are involved in the synthesis of flavonoids include three GmCHIs (GmCHI1A, GmCHI1B1, and GmCHI4A) and GmCHS14.

As the research objective was to identify the GmMYB176 interacting transcriptional complex for the regulation of *GmCHS8* gene expression, I focused on the TFs that have putatively bound to the *GmCHS8* promoter. The *GmCHS8* promoter analysis for regulatory elements binding sites was performed with the *CHS8*Δ*P5* promoter fragment in NSITE-PL (http://www.softberry.com/) (Solovyev et al., 2010; Shahmuradov and Solovyev, 2015) (Appendix 3). The TFs from Co-IP (based on GO annotation-Biological process- "Transcription, DNA-dependent") were retrieved. GmMYB176 and bHLHs were missing in the annotated category and thus were added to the list of TFs that could bind to *GmCHS8* gene promoter (Appendix 3). TFs that were predicted to bind to *GmCHS8* promoter was chosen for further studies (Fig 3.5). Four TFs that were identified in the Co-IP experiment and that could bind to the *GmCHS8* promoter are bZIPs, bHLHs, WRKY, and GmMYB176. The WRKY is known to be a stress responsive TF that is active in plants during bioticand abiotic stress (Kalde et al., 2003; Rushton et al., 2010; Jiang et al., 2014; Huang et al., 2016; Zhang et al., 2016). Excluding the WRKY, two bZIPs (Gm04bZIP and Gm05bZIP), three bHLHs (Gm05bHLH, Gm07bHLH, and Gm15bHLH) and GmMYB176 were chosen for further studies. In addition, four of the flavonoid pathway enzymes GmCHI1A, GmCHI1B1, GmCHI4A, and GmCHS14, which were identified as GmMYB176 interactors, were also used for further PPI validation *in planta*.

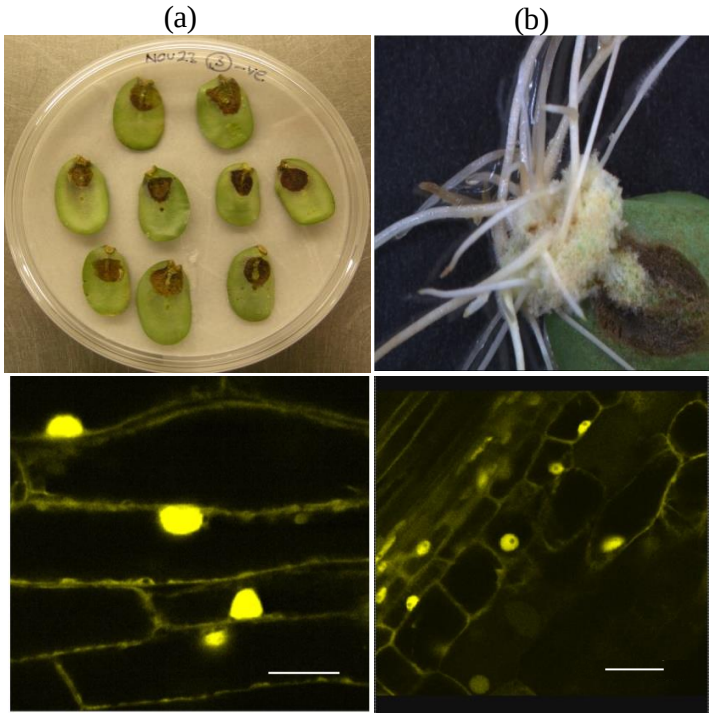


Figure 3.1 Soybean hairy roots showing localization of GmMYB176 fusion proteins. (a) Six day old soybean cotyledons were used to generate hairy roots using *A. rhizogenes* K599 transformed with GmMYB176-YFP or YFP-GmMYB176. (b) Photograph showing hairy roots three weeks after inoculation. Both (c) GmMYB176-YFP and (d) YFP-GmMYB176 fusion proteins were localized in nucleus and cytoplasm of the hairy root cells as observed with confocal microscopy. Scale bar = 50 µm.

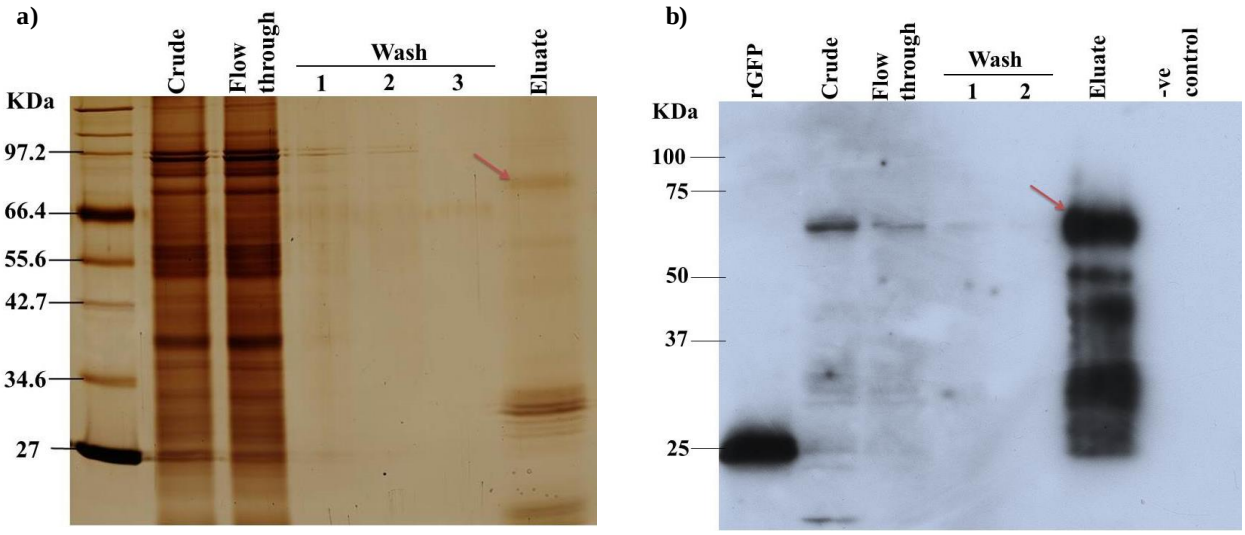


Figure 3.2 Co-immunoprecipitation of GmMYB176 interacting proteins from soybean hairy roots. (a) Total proteins were extracted from soybean hairy roots overexpressing GmMYB176-YFP, co-immunoprecipitated, separated on SDS-PAGE and stained with silver stain. (b) Western blot of the replicate gel as in (a) along with a positive control, rGFP and a negative control, crude extract from non-transgenic hairy roots. The proteins in the gel were transferred to a PVDF membrane and the blot was incubated sequentially with anti-GFP monoclonal antibody followed by HRP-conjugated goat anti-mouse secondary antibody, and then the HRP detection was performed with Super Signal West Femto. The arrow indicates the theoretical size of GmMYB176-YFP protein in the eluate. Crude: crude protein extract from soybean hairy roots; Flow through: crude extract incubated with anti-GFP microbeads and applied to µcolumn, with the flowthrough collected; Wash: sequential wash steps with lysis buffer; Eluate: elution of bound proteins from the column; -ve control: crude extract from non-transgenic hairy roots.

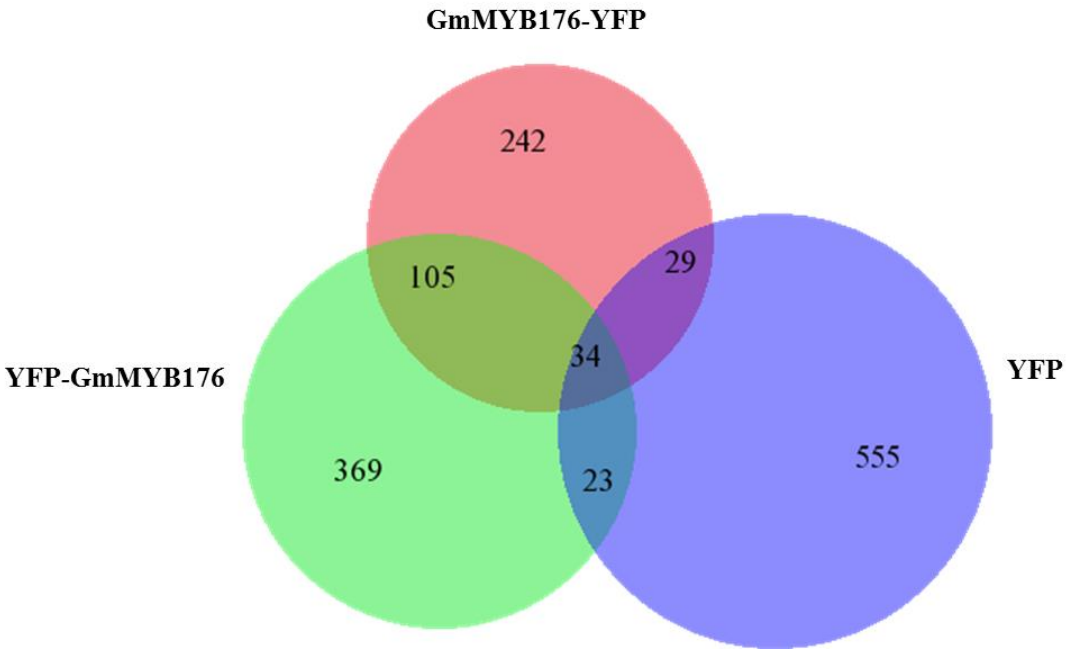
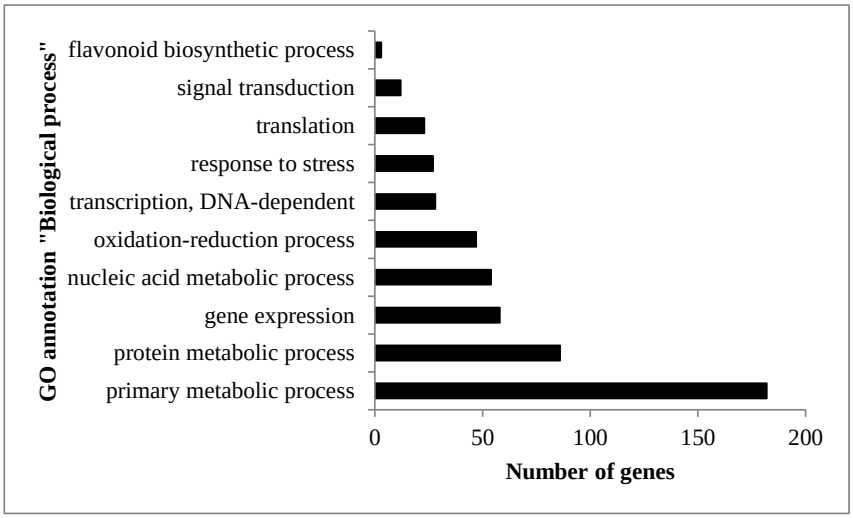
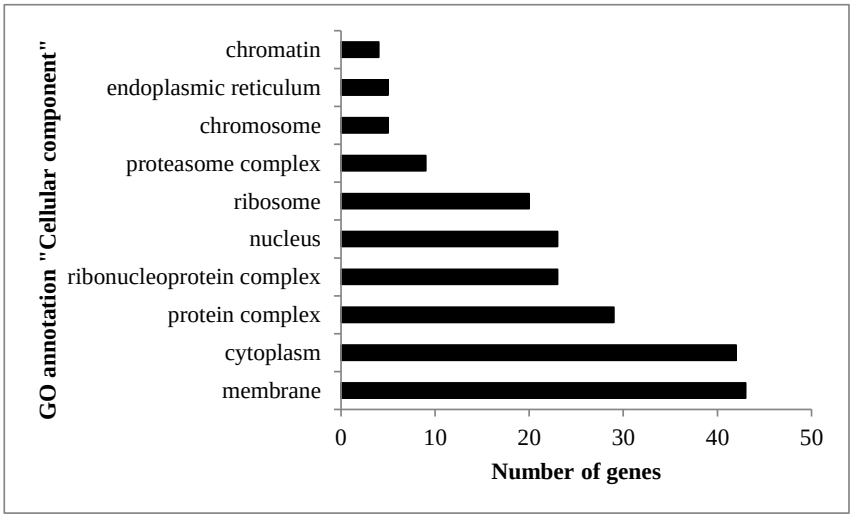


Figure 3.3 Overlap of YFP, GmMYB176-YFP, and YFP-GmMYB176 interacting candidate proteins identified by LC-MS/MS analysis. Venn diagram was generated using BioVenn (http://www.cmbi.ru.nl/cdd/biovenn/index.php). The YFP interacting protein candidates were obtained from Dastmalchi et al. (2016). A total of 716 GmMYB176 specific interactors were identified.


GO annotation "Biological process"
flavonoid biosynthetic process
signal transduction
translation
response to stress
transcription, DNA-dependent
oxidation-reduction process
nucleic acid metabolic process
gene expression
protein metabolic process
primary metabolic process
0
50
100
150
200
Number of genes


GO annotation "Cellular component"
chromatin
endoplasmic reticulum
chromosome
proteasome complex
ribosome
nucleus
ribonucleoprotein complex
protein complex
cytoplasm
membrane
0
10
20
30
40
50
Number of genes

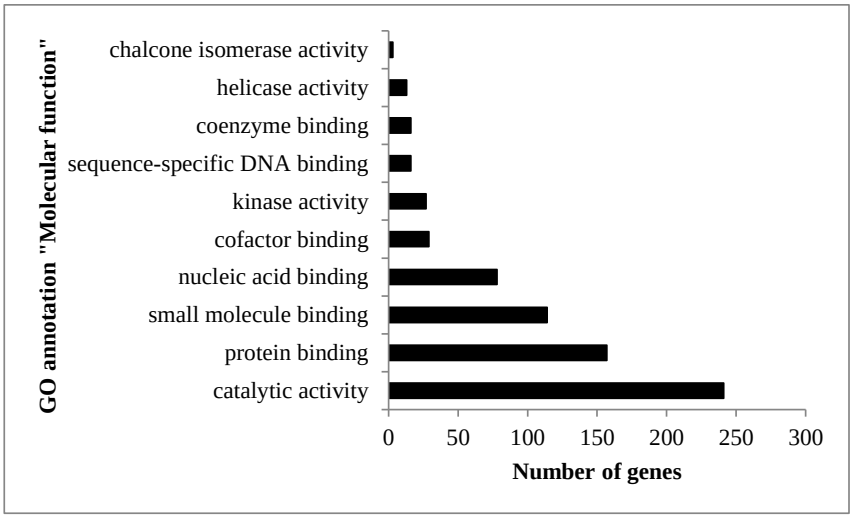

GO annotation "Molecular function"
chalcone isomerase activity
helicase activity
coenzyme binding
sequence-specific DNA binding
kinase activity
cofactor binding
nucleic acid binding
small molecule binding
protein binding
catalytic activity
0
50
100
150
200
250
300
Number of genes

Figure 3.4 ‘GO’ annotations of the 716 candidate GmMYB176-interacting proteins. The list of soybean genes encoding the candidate proteins was used in PhytoMine (http://www.phytozome.net/)(Goodstein et al., 2012) to generate annotations regarding (a) the biological process, (b) cellular component, and (c) the molecular function of the candidates.

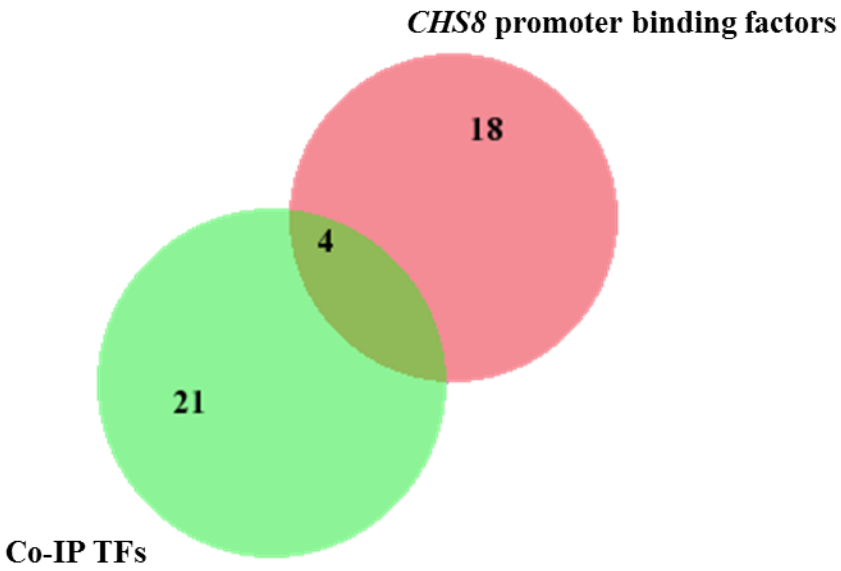


Figure 3.5 Overlap of TFs from Co-IP with that of TFs obtained from *in silico* analysis of *GmCHS8* promoter. The Co-IP list of TFs was retrieved from GO annotation "Biological process". *GmCHS8* promoter binding factors were identified by analysis of *CHS8ΔP5* promoter fragment in NSITE-PL (http://www.softberry.com/) (Solovyev et al., 2010; Shahmuradov and Solovyev, 2015) (Appendix 3).

Table 3.2 List of GmMYB176 interacting candidates used for *in planta* validation.

Name	Glyma Id	Family	*In planta* interaction with GmMYB176	Comments
GmMYB176	Glyma.05G032200.1.p	R1MYB	Yes (homo-dimer)	Binding site on *GmCHS8* promoter
GmMYB176S29A		R1MYB	Yes (homo-dimer)	Binding site on *GmCHS8* promoter
Gm04bZIP	Glyma.04G222200.1.p	bZIP	Yes	Binding site on *GmCHS8* promoter
Gm05bZIP	Glyma.05G122400.1.p	bZIP	Yes	Binding site on *GmCHS8* promoter
Gm05bHLH	Glyma.05G134400.1.p	bHLH	Yes	Binding site on *GmCHS8* promoter
Gm07bHLH	Glyma.07G205800.1.p	bHLH	No	Binding site on *GmCHS8* promoter
Gm15bHLH	Glyma.15G005100.1.p	bHLH	Yes	Binding site on *GmCHS8* promoter
Gm19DoF	Glyma.19G118000.1.p	DoF	Yes	Known to work with R1MYB
SGF14l	Glyma.08G115800.1.p	14-3-3 protein	Yes	Regulates intracellular localization of GmMYB176
GmCHI1A	Glyma.20G241500.1.p	Chalcone isomerase	Yes	Isoflavonoid biosynthesis pathway enzyme
GmCHI1B1	Glyma.20G241600.1.p	Chalconeisomerase	Yes	Isoflavonoid biosynthesis pathway enzyme
GmCHI4A	Glyma.06G143000.1.p	Chalconeisomerase	Yes	Isoflavonoid biosynthesis pathway enzyme
GmCHS14	Glyma.06G118500.1.p	Chalcone synthase	Yes	Isoflavonoid biosynthesis pathway enzyme
Gm17WD40	Glyma.17G029800.1.p	WD40	Yes	Transcriptional regulator
Gm02ZF	Glyma.02G029400.1.p	Zinc finger	Yes	-

3.3.2 Validation of PPIs between GmMYB176 and its interacting proteins obtained by Co-IP

To validate the GmMYB176 protein interaction candidates identified by Co-IP, BiFC assays were used. This method is widely used to study the PPIs *in planta* (Bracha-Drori et al., 2004). BiFC assay was conducted by using split YFP, where N- and Y-terminal halves of YFP were attached to two candidate proteins. In the event of a PPI, candidate proteins bring together the split YFP halves and form a functional fluorophore that can be detected by confocal microscopy. Thirteen candidate proteins including TFs and flavonoid biosynthetic pathway enzymes were used in the BiFC assay to confirm their interactions with GmMYB176 (Table 3.2). The results confirmed that GmMYB176 interacts with Gm04bZIP (Glyma.04G222200), Gm05bZIP (Glyma.05G122400), Gm05bHLH (Glyma.05G134400), Gm15bHLH (Glyma.15G005100), Gm19DoF (Glyma.19G118000) and Gm02ZF (Glyma.02G029400), and the interaction was observed in the nucleus (Fig. 3.6a). Although the interaction of GmMYB176 was observed with Gm17WD40 (Glyma.17G029800), GmCHI1A (Glyma.20G241500), GmCHI1B1 (Glyma.20G241600), GmCHI4A (Glyma.06G143000), and GmCHS14 (Glyma.06G118500), it was not only restricted to the nucleus but also occurred in the cytoplasm (Fig 3.6a and Table 3.2). Gm13bHLH (Glyma.13G368500) was identified as one of the GmMYB176 interactors by Co-IP. However, in an attempt to clone *Gm13bHLH*, *Gm15bHLH* (Glyma.15G005100) was amplified using *Gm13bHLH*-specific primers. Hence, Gm15bHLH was used for further analysis. Both *Gm13bHLH* and *Gm15bHLH* share 88.3% sequence identity at the nucleotide level. Gm07bHLH (Glyma.07G205800) did not show any interaction with GmMYB176 *in planta* and hence it was not used for further studies. SGF14l, a 14-3-3 protein, was used as a positive control as the interaction between SGF14l and GmMYB176 is well known (Li et al 2012).

Phosphorylation of GmMYB176 plays an important role in its intracellular localization in soybean. SGF14s binds with phosphorylated GmMYB76 and regulates the shuttling of GmMYB176 from the nucleus to cytoplasm whereas unphosphorylated GmMYB176 (GmMYB176D2/GmMYB176S29A) localized to the nucleus (Li et al., 2012). Based on this observation, I speculated that GmMYB176 transcriptional complex may include a non-phosphorylated GmMYB176. To confirm this, *in planta* PPI study of the candidate interactors were also performed with GmMYB176S29A. The confocal microscopy parameters for

fluorescence detection were kept under the same. The results demonstrated that GmMYB176S29A can interact with Gm04bZIP and Gm05bZIP and that based on the intensity of fluorescence, the interaction appeared stronger in the nucleus compared to their interaction with GmMYB176 (compare Fig 3.6a and b). The interactions of GmMYB176S29A with GmCHI1B1, GmCHI4A, and GmCHS14 were observed in the nucleus only (Fig 3.6b). However, GmMYB176S29A did not show any interaction with Gm05bHLH, Gm07bHLH, Gm15bHLH, Gm19DoF, Gm02ZF, GmCHI1A and Gm17WD40 *in planta*.

3.3.3 Two GmMYB176 interactors, Gm04bZIP and Gm05bZIP, bind *GmCHS8* promoter

GmMYB176 interacting TFs that are involved in *GmCHS8* gene activation may also possess the ability to bind *GmCHS8* promoter. Therefore, I analyzed the DNA binding ability of all the GmMYB176 interactors that were identified by Co-IP and validated for their interactions *in planta* (Fig 3.6), using the Y1H assay. It is known that the *CHS8*Δ*P5* promoter region contains all the elements necessary for *GmCHS8* gene activation (Yi et al., 2010). By using a deletion study, Yi et al. (2010) also determined the region critical for *GmCHS8* activation to a 23 bp region within the *GmCHS8* promoter. Therefore, I chose *CHS8*Δ*P5* promoter fragment and 23 bp region within the *GmCHS8* promoter with an additional flanking sequence containing a putative bZIP binding region (30 bp in a tandem repeat, 30bpTR) for theY1H assay. Out of six TFs used, two bZIPs (Gm04bZIP and Gm05bZIP) showed DNA binding activity with *CHS8*Δ*P5* and *30bpTR* in yeast (Fig 3.7a). A targeted Y1H assay was also performed with GmMYB176S29A as a prey proteinand *GmCHS8* promoter fragments (*CHS8*Δ*P5* and *30bpTR*) as baits (Fig 3.7a).The results revealed that GmMYB176S29A binds to the *GmCHS8* promoter suggesting that the DNA binding activity of GmMYB176 does not depend on its phosphorylation state.

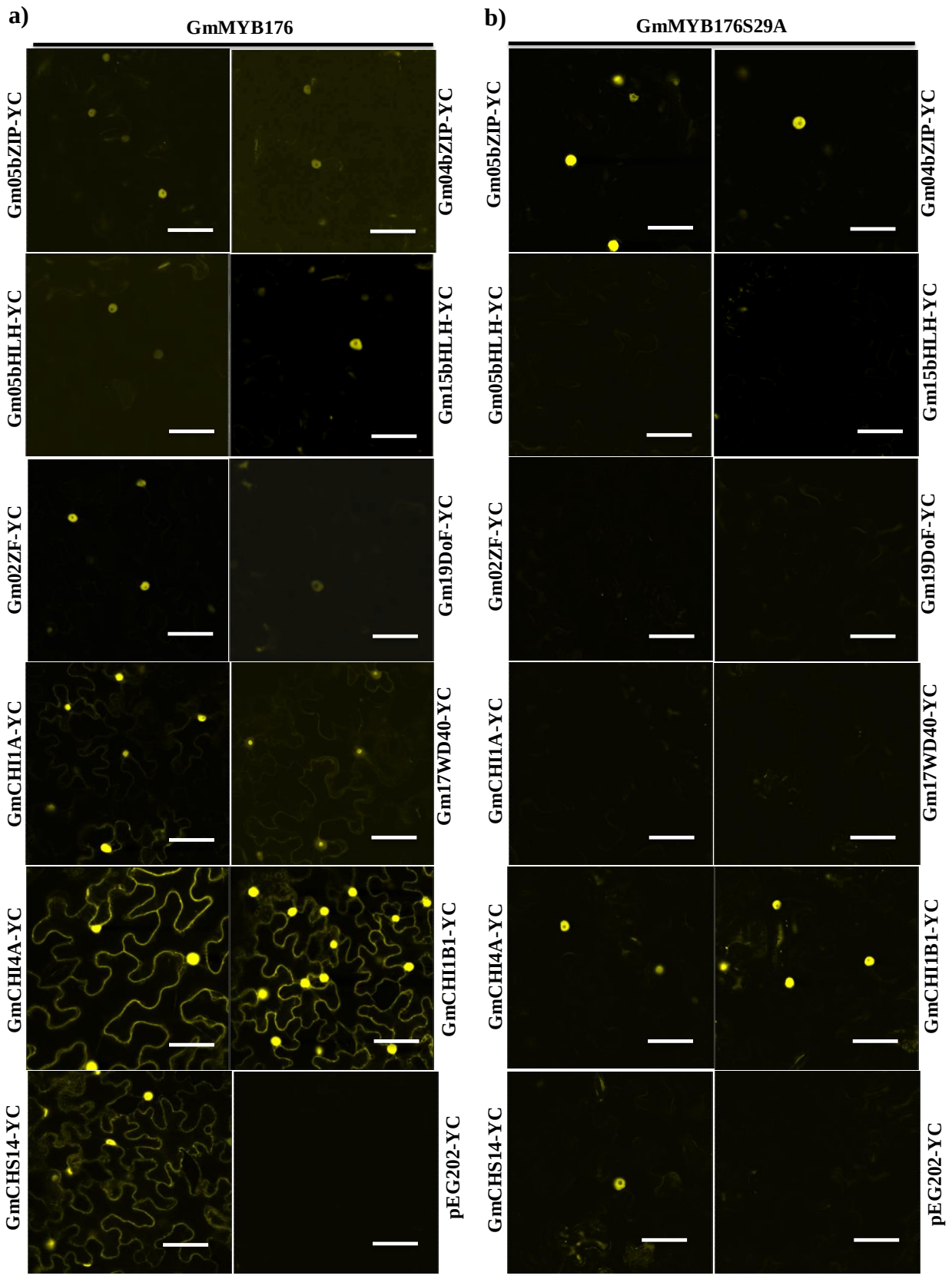

a)
GmMYB176
Gm05bZIP-YC
Gm04bZIP-YC
Gm05bHLH-YC
Gm15bHLH-YC
Gm02ZF-YC
Gm19DoF-YC
GmCHI1A-YC
Gm17WD40-YC
GmCHI4A-YC
GmCHI1B1-YC
GmCHS14-YC
pEG202-YC
b)
GmMYB176S29A
Gm05bZIP-YC
Gm04bZIP-YC
Gm05bHLH-YC
Gm15bHLH-YC
Gm02ZF-YC
Gm19DoF-YC
GmCHI1A-YC
Gm17WD40-YC
GmCHI4A-YC
GmCHI1B1-YC
GmCHS14-YC
pEG202-YC

Figure 3.6 Protein-protein interaction of GmMYB176 and GmMYB176S29A with Co-IP candidates *in planta*. The interaction between proteins was assayed by co-expression of translational fusions of candidate proteins with N–terminal (YN) and C–terminal (YC) fragments of YFP. The proximity of the two fragments results in a functional fluorophore. The fluorescence indicates the presence and location of the interaction between (a) GmMYB176 and (b) GmMYB176S29A with the candidate proteins as indicated. Fluorescent intensity parameters were kept same for all images. Scale bars = 50 µm.

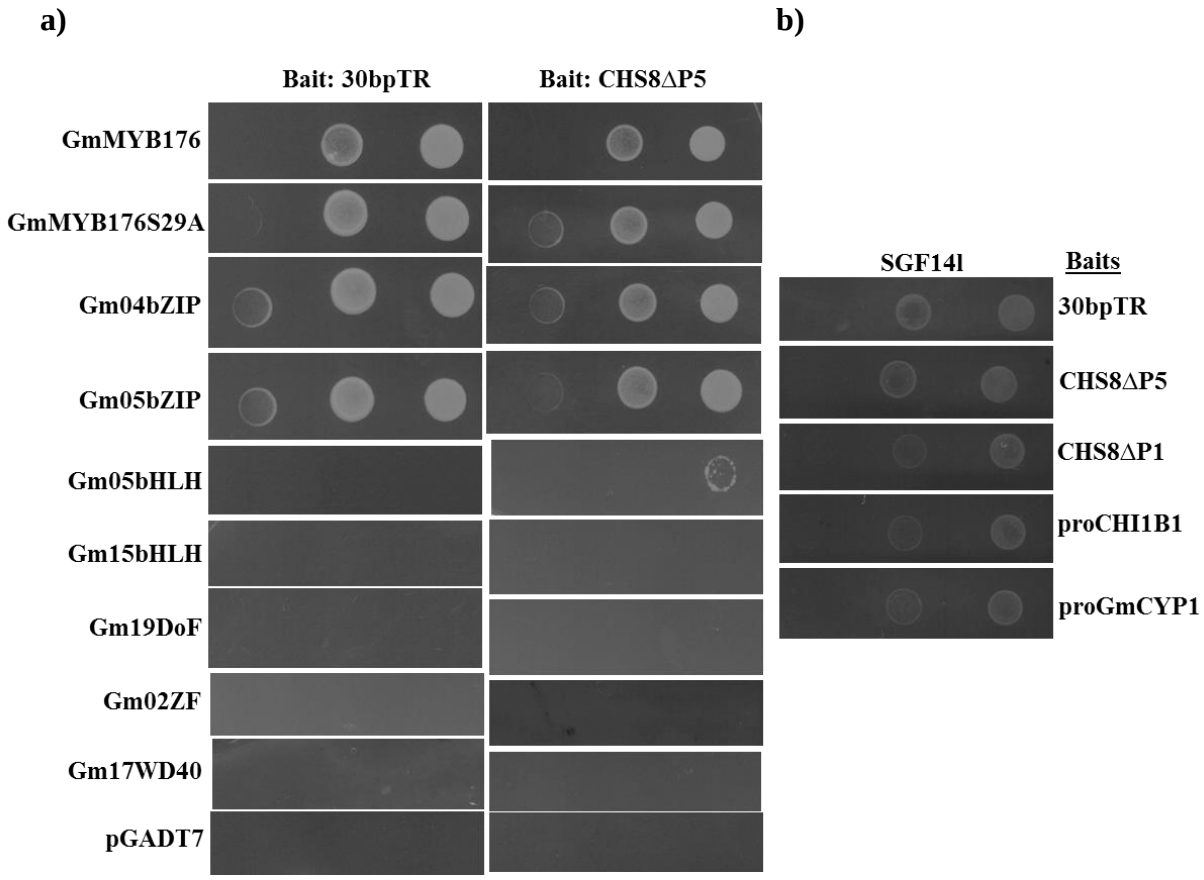


Figure 3.7 *GmCHS8* promoter binding activity of GmMYB176 interacting candidates. Yeast cells carrying *30bpTR, CHS8ΔP5, CHS8ΔP1, proCHI1B1*, or *pro*Gm*CYP1* promoter bait fragments, were transformed with prey constructs fused to a GAL4 activation domain. A series of 5 µL of 10x diluted yeast suspension culture transformed with prey constructs were spotted onto synthetic defined (SD) selection plates. Growth on SD lacking leucine and in the presence of Aureobasidin A (SD/-Leu/AbA) shows the activation of the reporter and indicates DNA binding activity. As a negative control, pGADT7 vector only was used.

Interestingly, SGF14l also showed DNA binding activity with *CHS8*Δ*P5* and *30bpTR*. However, similar results were obtained for SGF14l with other promoter fragments, such as *CHS8*Δ*P1, proCHI1B1*, and *proGmCYP1* (Fig 3.7b).

3.4 Discussion

3.4.1 YFP tag position did not alter the localization of GmMYB176

A protein-protein fusion could affect its natural three-dimensional structural integrity, which in turn could affect the interaction of some proteins with the fusion protein. Routzahn and Waugh (2002) reported that the nature of the tag and its location affects the solubility of fusion proteins when *E. coli'*s maltose binding (MBP) protein was used as a tag for purification of its fusion partners. In that study, the addition of C-terminal His-tag reduced the solubility of MBP fusion proteins, whereas an N-terminal His-tag showed less impact on the solubility. Moreover, the fusion of His tag at the C-terminus of a PhoP protein affected its biochemical properties and its interaction with target DNA in *Salmonella enterica*. This could be due to changes in conformation or oligomerization state of the protein PhoP associated with the fusion tag (Perron-Savard et al., 2005). YFP-tagging altered the biochemical property and strength of glycosylation of some transmembrane proteins like transient receptor potential vanilloid 6 channels (Park et al., 2009). The subcellular localization of GmIFS2 is altered when YFP is fused at the N-terminal end of GmIFS2 in soybean (Dastmalchi et al., 2016). So, it is possible that position of YFP fusion with GmMYB176 could affect its folding and the recognition of its interacting partner. Although the position of YFP-tag did not alter the subcellular localization of GmMYB176 (Fig 3.1), it does not exclude the possibility that the position of tag affects its interaction with other proteins. Therefore, translational fusions of GmMYB176 with YFP at both N- and C- termini were created to use as baits in Co-IP to identify all potential interactors.

3.4.2 GmMYB176 interacts with isoflavonoid pathway enzymes

All the candidates (716) obtained by Co-IP (Appendix1) were categorized based on GO annotations (Fig 3.4). In addition to GO annotations, pathway enrichment was also performed. The interactions of GmMYB176 with flavonoid biosynthesis pathway enzymes such as GmCHI1A, GmCHI1B1, GmCHI4A, and GmCHS14 were found in both nucleus and cytoplasm.

Nuclear localization of phenylpropanoid pathway enzymes has been previously reported in *Arabidopsis* (Saslowsky et al., 2005) and in soybean (Dhaubhadel et al., 2008; Dastmalchi and Dhaubhadel, 2015a). In addition to their localization in the cytoplasm, three GmCHIs (subfamilies 1 and 2), GmCHS8 (Dastmalchi et al., 2016), isoflavonoid-specific glycosyltransferase (UGT73F2) and malonyltransferase (GmMT7) occur in the nucleus possibly indicating their function in the nucleus (Dhaubhadel et al., 2008; Dastmalchi and Dhaubhadel, 2015a). Here I demonstrate that GmMYB176 interacts with three GmCHI isoforms, and also with GmCHS14. The phosphorylated GmMYB176 interacts with GmCHIs and GmCHS14 in both cytoplasm and nucleus (Fig 3.6a) whereas unphosphorylated GmMYB176 (GmMYB176S29A) interacts with the GmCHI4A, GmCHI1B1, and GmCHS14 in the nucleus only (Fig 3.6). This is the first observation of an isoflavonoid biosynthetic pathway enzyme interacting with a TF. Since the key cytochrome P450 enzymes GmC4H and GmIFS involved in isoflavonoid biosynthesis are anchored to the ER membrane (Dastmalchi et al., 2016), it is unlikely that these nuclear enzymes GmCHIs and GmCHS14 perform the same function in the nucleus. It is also possible that these enzymes have multiple functions, an occurrence commonly referred to as 'moonlighting'. Moonlighting proteins can perform multiple functions simultaneously or at different times, in response to environmental stimuli (Jeffery, 2014).

3.4.3 GmMYB176 forms a transcriptional complex with bZIPs and may regulate *GmCHS8* gene expression

It is known that phosphorylation of GmMYB176 affects its intracellular localization (Li et al., 2012). Activation of a gene transcription by the unphosphorylated state of TFs has been previously reported (Kim et al., 2009; Takahashi et al., 2013; Li et al., 2016). In *Arabidopsis,* three bHLH TFs (AKS1, AKS2, and AKS3) in their unphosphorylated state, activate the genes that are responsible for stomatal opening (Takahashi et al., 2013). In human cell lines, both phosphorylated and unphosphorylated form of the Forkhead box O3 (FOXO3) TF bind to its target promoter; however, only the unphosphorylated form of FOXO3 TF serves as a transcriptional activator (Li et al., 2016). Since both GmMYB176 and GmMYB176S29A can bind to the *GmCHS8* gene promoter (Fig. 3.7a), it is possible that the transcriptional complex for the regulation of *GmCHS8* gene contains either phosphorylated or unphosphorylated state of GmMYB176.

Among the GmMYB176 interactors obtained from the Co-IP experiment, 28 candidates were categorized as "Transcription, DNA-dependent" by GO annotation. It is not necessary that all the proteins that interact with GmMYB176 bind to the *GmCHS8* gene promoter. Therefore, to select the TFs that are involved in *GmCHS8* regulation, a list of TFs whose binding sites are predicted in the *GmCHS8* gene promoter were selected from Co-IP experiment and validated for PPI with both GmMYB176 and GmMYB176S29A for the *GmCHS8* promoter binding activity (Fig 3.6 and Fig 3.3). Not all the candidate proteins confirmed for PPI with GmMYB176 (Fig 3.6) interacted with the *GmCHS8* promoter (Fig 3.7a) suggesting that these complexes could regulate other genes in either isoflavonoid biosynthesis pathway or some other pathway yet to be identified. A previous study showed the level of *IFS* transcript was reduced in GmMYB176 silenced soybean hairy roots (Yi et al., 2010). GmMYB176 could regulate multiple genes in the pathway by forming different transcriptional complexes.

My results provide strong evidence that unphosphorylated GmMYB176 interact with Gm04bZIP and Gm05bZIP in the nucleus and also bind to the *GmCHS8* promoter at the region that is critical for the gene activation. An R1 MYB, HvMYBS3, forms a ternary complex with a bZIP TF and a DoF TF and enhances the expression of developing endosperm-specific gene *ltr1* in Barley (Rubio-Somoza et al., 2006). Though GmMYB176 showed DNA binding activity with *GmCHS8* promoter fragments in yeast, the phosphorylation state of GmMYB176 in yeast nucleus is unknown, especially when GmMYB176 binds to the promoter fragments for the reporter gene activation in yeast.

DNA binding activities of Gm04bZIP and Gm05bZIP with *GmCHS8* promoter suggests that these two bZIPs work in combination with GmMYB176 for the regulation of *GmCHS8* gene expression. The coding sequences of the Gm04bZIP (Glyma.04G222200) and Gm05bZIP (Glyma.05G122400) genes were 36.3% identical. Plant bZIP TFs bind to a specific promoter region that contains an ACGT core, such as A-box, C-box, G-box, hybrid C/G-box, or C/A-box motifs (Izawa et al., 1993; Jakoby et al., 2002; Song et al., 2008). A transcriptional complex with an R1 MYB TF and a bZIP TF for the regulation of a gene occurs in barley (Rubio-Somoza et al., 2006). The *GmCHS8* promoter contains 12 predicted GmMYB176 binding sites and 5 predicted bZIP binding sites. However, only the deletion of a 23 bp motif containing GmMYB176 binding sites affected the promoter activity. Yi et al. (2010) conducted a promoter

deletion study to identify the region that is critical for *GmCHS8* gene expression. Eight fragments with 5' deletion of the *GmCHS8* promoter were generated and used to drive GUS expression in a co-transfection assay using *Arabidopsis* protoplast. The result suggested that a region of the 23bp motif in the *CHS8*Δ*P5* fragment is critical for *GmCHS8* gene expression. Although the *GmCHS8* promoter has multiple predicted bZIP binding sites (ACGT core), only one is proximal to the GmMYB176 recognition site (23bp motif) of *GmCHS8* gene promoter (Fig 3.7). Hence, GmMYB176S29A could form complex with Gm04bZIP and/or Gm05bZIP for *GmCHS8* gene regulation in soybean (Fig 3.8).

SGF14l, a 14-3-3 protein also showed a DNA binding activity in yeast. SGF14l interacts with all the baits *CHS8*Δ*P5*, *CHS8*Δ*P1*, *30bpTR*, *proCHI1B1*, and *pro*Gm*CYP1* used in the present study (Fig 3.7). Many biological processes such as protein trafficking, primary metabolism, and cell signaling are regulated by 14-3-3 proteins in plants (Bunney et al., 2002; Chevalier et al., 2009; Denison et al., 2011; Catala et al., 2014; Zhou et al., 2015). 14-3-3 proteins are also part of a florigen activation complex (FAC) in *Arabidopsis* (Hou and Yang, 2016). In wheat, 14-3-3 protein interacts with FT1 protein and then translocate from cytoplasm to the nucleus to form a FAC with a bZIP TF. This FAC including 14-3-3 protein showed DNA binding activity with the *Vernalization 1* gene (*VRN1*) promoter *in vitro* (Li et al., 2015). The presence of 14-3-3 protein in FAC showed stronger DNA binding activity than FAC without 14-3-3 protein *in vitro*. Ta14-3-3C enhances the binding activity of FAC to the VRN1 promoter(Li et al., 2015). Similarly, in soybean, SGF14l may directly or indirectly activate the reporter gene in Y1H assays. SGF14l may activate an endogenous TF that could bind to the promoters and/or may have direct DNA-binding activity. A TF can be involved in several transcriptional complexes. TFs such as Gm05bHLH, Gm15bHLH, Gm02ZF, and Gm19DoF which were identified as GmMYB176 interactors did not show DNA binding activity with *GmCHS8* promoter fragments.

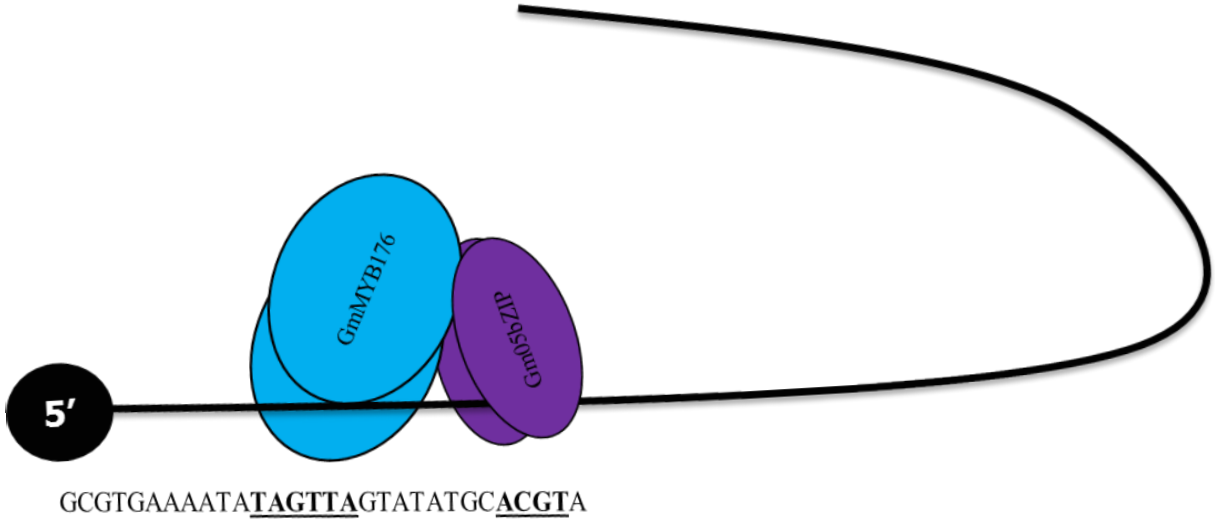


Figure 3.8 A proposed model of binding of GmMYB176 and bZIP on the *GmCHS8* gene promoter. Unphosphorylated GmMYB176 and Gm04bZIP/ Gm05bZIP bind with the *GmCHS8* promoter and also interact with each other suggesting a tight transcriptional complex for *GmCHS8* gene activation. The GmMYB176 binding site (TAGT[T/A][A/T]) and putative binding site for bZIP (ACGT core) are underlined.

3.5 **Conclusion**

A total of 716 GmMYB176 interacting proteins were identified in soybean hairy roots by a Co-IP using GmMYB176 as the bait. The interactions of GmMYB176 with six TFs were verified *in planta*. GmMYB176 interacts with Gm04bZIP and Gm05bZIP TFs in the nucleus irrespective of its phosphorylation state. However, a relatively stronger interaction was observed when GmMYB176 was in an unphosphorylated state. Similarly, GmMYB176's DNA binding ability with the *GmCHS8* promoter was not dependent on phosphorylation. Since both GmMYB176 and bZIP TF binding motifs are present in the critical sequence that determines *GmCHS8* gene activity, I conclude that GmMYB176 works in combination with Gm04bZIP and/or Gm05bZIP for *GmCHS8* gene activation in soybean. GmMYB176S29A also interacts with bZIPs in the nucleus and showed DNA binding activity to the *GmCHS8* promoter. Hence, an unphosphorylated form of GmMYB176 interacts with Gm04bZIP/Gm05bZIP and may regulate *GmCHS8* gene expression in soybean. The information generated from these studies can lead to several paths for future research in isoflavonoid biosynthesis pathway. Understanding the role of GmMYB176 interactome will give us deeper knowledge of the proteins involved in the regulation of isoflavonoid biosynthesis.

3.6 **Literature cited**

Boddu, J., Jiang, C., Sangar, V., Olson, T., Peterson, T., and Chopra, S. (2006). Comparative structural and functional characterization of sorghum and maize duplications containing orthologous MYB transcription regulators of 3-deoxyflavonoid biosynthesis. Plant Molecular Biology **60,** 185-199.

Bracha-Drori, K., Shichrur, K., Katz, A., Oliva, M., Angelovici, R., Yalovsky, S., and Ohad, N. (2004). Detection of protein-protein interactions in plants using bimolecular fluorescence complementation. The Plant Journal **40,** 419-427.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry **72,** 248-254.

Bunney, T.D., van den Wijngaard, P.W., and de Boer, A.H. (2002). 14-3-3 protein regulation of proton pumps and ion channels. Plant Molecular Biology **50,** 1041-1051.

Catala, R., Lopez-Cobollo, R., Mar Castellano, M., Angosto, T., Alonso, J.M., Ecker, J.R., and Salinas, J. (2014). The *Arabidopsis* 14-3-3 protein RARE COLD INDUCIBLE 1A links low-temperature response and ethylene biosynthesis to regulate freezing tolerance and cold acclimation. Plant Cell **26,** 3326-3342.

Chevalier, D., Morris, E.R., and Walker, J.C. (2009). 14-3-3 and FHA domains mediate phosphoprotein interaction. Annual Review of Plant Biology **60,** 67-91.

Craig, R.W., and Bloch, A. (1984). Early decline in c-myb oncogene expression in the differentiation of human myeloblastic leukemia (ML-1) cells induced with 12-O-tetradecanoylphorbol-13-acetate. Cancer Research**44,** 442-446.

Dastmalchi, M., and Dhaubhadel, S. (2015a). Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. Planta **241,** 507-523.

Dastmalchi, M., and Dhaubhadel, S. (2015b). Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. Planta **241,** 507-523.

Dastmalchi, M., Bernards, M.A., and Dhaubhadel, S. (2016). Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. The Plant Journal **85,** 689-706.

Denison, F.C., Paul, A.-L., Zupanska, A.K., and Ferl, R.J. (2011). 14-3-3 proteins in plant physiology. Seminars in Cell and Developmental Biology **22,** 720-727.

Dhaubhadel, S., Farhangkhoee, M., and Chapman, R. (2008). Identification and characterization of isoflavonoid specific glycosyltransferase and malonyltransferase from soybean seeds. Journal of Experimental Botany **59,** 981-994.

Dhaubhadel, S., Gijzen, M., Moy, P., and Farhangkhoee, M. (2007). Transcriptome analysis reveals a critical role of *CHS7* and *CHS8* genes for isoflavonoid synthesis in soybean seeds. Plant Physiology **143,** 326-338.

Du, H., Zhang, L., Liu, L., Tang, X.F., Yang, W.J., Wu, Y.M., Huang, Y.B., and Tang, Y.X. (2009). Biochemical and molecular characterization of plant MYB transcription factor family. Biochemistry (Mosc) **74,** 1-11.

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

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. The Plant Journal **45,** 616-629.

Fernández-Marcos, M., Desvoyes, B., Manzano, C., Liberman, L.M., Benfey, P.N., del Pozo, J.C., and Gutierrez, C. (2016). Control of *Arabidopsis* lateral root primordium boundaries by MYB36. New Phytologist**213,** 105-112.

Goff, S.A., Cone, K.C., and Chandler, V.L. (1992). Functional analysis of the transcriptional activator encoded by the maize *B* gene: evidence for a direct functional interaction between two classes of regulatory proteins. Genes and Development **6,** 864-875.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research**40,** D1178-1186.

Hahlbrock, K., and Scheel, D. (1989). Physiology and molecular biology of phenylpropanoid metabolism. Annual Review of Plant Physiology and Plant Molecular Biology **40,** 347-369.

Hancock, K.R., Collette, V., Fraser, K., Greig, M., Xue, H., Richardson, K., Jones, C., and Rasmussen, S. (2012). Expression of the R2R3-MYB transcription factor TaMYB14 from *Trifolium arvense* activates proanthocyanidin biosynthesis in the legumes *Trifolium repens* and *Medicago sativa*. Plant Physiology **159,** 1204-1220.

Hou, C.J., and Yang, C.H. (2016). Comparative analysis of the pteridophyte *Adiantum* MFT ortholog reveals the specificity of combined FT/MFT C and N terminal interaction with FD for the regulation of the downstream gene *AP1*. Plant Molecular Biology **91,** 563-579.

Huang, Y., Li, M.Y., Wu, P., Xu, Z.S., Que, F., Wang, F., and Xiong, A.S. (2016). Members of WRKY Group III transcription factors are important in TYLCV defense signaling pathway in tomato (*Solanum lycopersicum*). BMC Genomics **17,** 788.

Huang, Y.F., Vialet, S., Guiraud, J.L., Torregrosa, L., Bertrand, Y., Cheynier, V., This, P., and Terrier, N. (2014). A negative MYB regulator of proanthocyanidin accumulation, identified through expression quantitative locus mapping in the grape berry. New Phytologist**201,** 795-809.

Izawa, T., Foster, R., and Chua, N.H. (1993). Plant bZIP protein DNA binding specificity. Journal of Molecular Biology**230,** 1131-1144.

Jakoby, M., Weisshaar, B., Droge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., Parcy, F., and b, Z.I.P.R.G. (2002). bZIP transcription factors in *Arabidopsis*. Trends in Plant Science **7,** 106-111.

Jeffery, C.J. (2014). An introduction to protein moonlighting. Biochemical Society Transactions**42,** 1679-1683.

Jiang, Y., Duan, Y., Yin, J., Ye, S., Zhu, J., Zhang, F., Lu, W., Fan, D., and Luo, K. (2014). Genome-wide identification and characterization of the *Populus* WRKY transcription factor family and analysis of their expression in response to biotic and abiotic stresses. Journal of Experimental Botany **65,** 6629-6644.

Jin, H., and Martin, C. (1999). Multifunctionality and diversity within the plant MYB-gene family. Plant Molecular Biology **41,** 577-585.

Kalde, M., Barth, M., Somssich, I.E., and Lippok, B. (2003). Members of the *Arabidopsis* WRKY group III transcription factors are part of different plant defense signaling pathways. Molecular Plant-Microbe Interactions **16,** 295-305.

Kim, T.W., Guan, S., Sun, Y., Deng, Z., Tang, W., Shang, J.X., Sun, Y., Burlingame, A.L., and Wang, Z.Y. (2009). Brassinosteroid signal transduction from cell-surface receptor kinases to nuclear transcription factors. Nature Cell Biology**11,** 1254-1260.

Li, C., Lin, H., and Dubcovsky, J. (2015). Factorial combinations of protein interactions generate a multiplicity of florigen activation complexes in wheat and barley. The Plant Journal **84,** 70-82.

Li, X., Chen, L., and Dhaubhadel, S. (2012). 14-3-3 proteins regulate the intracellular localization of the transcriptional activator GmMYB176 and affect isoflavonoid synthesis in soybean. The Plant Journal **71,** 239-250.

Li, Z., Zhao, J., Tikhanovich, I., Kuravi, S., Helzberg, J., Dorko, K., Roberts, B., Kumer, S., and Weinman, S.A. (2016). Serine 574 phosphorylation alters transcriptional programming of FOXO3 by selectively enhancing apoptotic gene expression. Cell Death and Differentiation**23,** 583-595.

Lu, Q., Tang, X., Tian, G., Wang, F., Liu, K., Nguyen, V., Kohalmi, S.E., Keller, W.A., Tsang, E.W., Harada, J.J., Rothstein, S.J., and Cui, Y. (2010). *Arabidopsis* homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin. The Plant Journal **61,** 259-270.

Mainali, H.R. (2014). Identification of Cyclophilin gene family in soybean and characterization of Gm*CYP1* (London, Ont.: School of Graduate and Postdoctoral Studies, University of Western Ontario).

Ogata, K., Kanei-Ishii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., Nakamura, H., Nishimura, Y., Ishii, S., and Sarai, A. (1996). The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation. Nature Structural Biology **3,** 178-187.

Park, E.J., Kim, B.J., Kim, S.H., Kim, S.Y., Sung, T.S., Chae, H.G., Kim, S.J., Kim, J., Park, H.H., So, I., and Jeon, J.H. (2009). Altered biochemical properties of transient receptor potential vanilloid 6 calcium channel by peptide tags. Biological and Pharmaceutical Bulletin **32,** 1790-1794.

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A., and Saedler, H. (1987). The regulatory C1 locus of *Zea mays* encodes a protein with homology to MYB proto-oncogene products and with structural similarities to transcriptional activators. EMBO Journal **6,** 3553-3558.

Perron-Savard, P., De Crescenzo, G., and Le Moual, H. (2005). Dimerization and DNA binding of the *Salmonella enterica* PhoP response regulator are phosphorylation independent. Microbiology **151,** 3979-3987.

Routzahn, K.M., and Waugh, D.S. (2002). Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. Journal of Structural and Functional Genomics **2,** 83-92.

Rubio-Somoza, I., Martinez, M., Abraham, Z., Diaz, I., and Carbonero, P. (2006). Ternary complex formation between HvMYBS3 and other factors involved in transcriptional control in barley seeds. The Plant Journal **47,** 269-281.

Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. Trends in Plant Science **15,** 247-258.

Saslowsky, D.E., Warek, U., and Winkel, B.S. (2005). Nuclear localization of flavonoid enzymes in *Arabidopsis*. Journal of Biological Chemistry **280,** 23735-23740.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., and Jackson, S.A. (2010). Genome sequence of the palaeopolyploid soybean. Nature **463,** 178-183.

Shahmuradov, I.A., and Solovyev, V.V. (2015). Nsite, NsiteH and NsiteM computer tools for studying transcription regulatory elements. Bioinformatics **31,** 3544-3545.

Solovyev, V.V., Shahmuradov, I.A., and Salamov, A.A. (2010). Identification of promoter regions and regulatory sites. Methods in Molecular Biology**674,** 57-83.

Song, Y.H., Yoo, C.M., Hong, A.P., Kim, S.H., Jeong, H.J., Shin, S.Y., Kim, H.J., Yun, D.J., Lim, C.O., Bahk, J.D., Lee, S.Y., Nagao, R.T., Key, J.L., and Hong, J.C. (2008). DNA-binding study identifies C-box and hybrid C/G-box or C/A-box motifs as high-affinity binding sites for STF1 and LONG HYPOCOTYL5 proteins. Plant Physiology **146,** 1862-1877.

Sparkes, I.A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nature Protocol**1,** 2019-2025.

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. Current Opinion in Plant Biology **4,** 447 - 456.

Subramanian, S., Graham, M.A., Yu, O., and Graham, T.L. (2005). RNA interference of soybean isoflavone synthase genes leads to silencing in tissues distal to the transformation site and to enhanced susceptibility to *Phytophthora sojae*. Plant Physiology **137,** 1345-1353.

Takahashi, Y., Ebisu, Y., Kinoshita, T., Doi, M., Okuma, E., Murata, Y., and Shimazaki, K. (2013). bHLH transcription factors that facilitate K(+) uptake during stomatal opening are repressed by abscisic acid through phosphorylation. Science Signaling**6,** ra48.

Tuteja, J.H., and Vodkin, L.Q. (2008). Structural features of the endogenous *CHS* silencing and target loci in the soybean genome. Crop Science **48,** S49-S68.

Xu, F., Ning, Y., Zhang, W., Liao, Y., Li, L., Cheng, H., and Cheng, S. (2014). An R2R3-MYB transcription factor as a negative regulator of the flavonoid biosynthesis pathway in *Ginkgo biloba*. Functional and Integrative Genomics **14,** 177-189.

Yan, J., Wang, B., Zhong, Y., Yao, L., Cheng, L., and Wu, T. (2015). The soybean R2R3 MYB transcription factor GmMYB100 negatively regulates plant flavonoid biosynthesis. Plant Molecular Biology.

Yi, J., Derynck, M.R., Li, X., Telmer, P., Marsolais, F., and Dhaubhadel, S. (2010). A single-repeat MYB transcription factor, GmMYB176, regulates *CHS8* gene expression and affects isoflavonoid biosynthesis in soybean. The Plant Journal **62,** 1019-1034.

Zhang, Y., Yu, H., Yang, X., Li, Q., Ling, J., Wang, H., Gu, X., Huang, S., and Jiang, W. (2016). CsWRKY46, a WRKY transcription factor from cucumber, confers cold resistance in transgenic-plant by regulating a set of cold-stress responsive genes in an ABA-dependent manner. Plant Physiology and Biochemistry **108,** 478-487.

Zhou, Y., Zhang, Z.T., Li, M., Wei, X.Z., Li, X.J., Li, B.Y., and Li, X.B. (2015). Cotton (*Gossypium hirsutum*) 14-3-3 proteins participate in regulation of fibre initiation and elongation by modulating brassinosteroid signalling. Plant Biotechnology Journal **13,** 269-280.

4 REGULATION OF ISOFLAVONOID BIOSYNTHETIC GENES' EXPRESSION AND METABOLITE PRODUCTION BY GmMYB176 IN SOYBEAN

4.1 Introduction

Plant metabolism has been classically divided into two types: primary and secondary metabolism (Dudareva et al., 2010; Pichersky and Lewinsohn, 2011). Metabolic processes that are essential for all plants for their growth and development are classified as primary metabolism. Other metabolic processes that are distributed discretely among plant species which are not essential for their primary growth, but are required for physiochemical phenotypic differences or in defense mechanisms are called secondary or specialized metabolism. Plants produce a wide variety of specialized metabolites and many of them are biologically active.

Isoflavonoids are specialized metabolites that are synthesized through a legume-specific branch of phenylpropanoid pathway. Phenylpropanoid biosynthesis is regulated by coordinated expression of several structural genes. These structural genes are often regulated by multiple transcription factors (Vom Endt et al., 2002). Soybean is a paleopolyploid and most of the phenylpropanoid pathway enzymes are encoded by multigene families that resulted from two whole genome duplication events. Studying expression levels of gene family members can be challenging due to their high sequence identity (Blanc and Wolfe, 2004; Shoemaker et al., 2006; Panchy et al., 2016). In the phenylpropanoid pathway, phenylalanine ammonia lyase (PAL) is the first enzymeinvolved in the conversion of phenylalanineto cinnamic acid. PAL catalyzes the elimination of ammonia from phenylalanine to generate the C6-C3 precursor cinnamic acid. Some of the early downstream enzymes of the phenylpropanoid pathway, after PAL, are cinnamate 4-hydroxylase (C4H), 4 coumoryl-CoA ligase (4CL) and chalcone synthase (CHS) (Fig 1.2), with the latter being a key step in directing C6-C3 carbon skeletons into the flavonoid/isoflavonoid branch of the phenylpropanoid pathway. GmCHS produces naringenin chalcone which is then utilized in isoflavonoid and

flavonoid biosynthesis. A total of 9 *CHS* genes (*GmCHS1-GmCHS9*) have been reported in soybean (Tuteja and Vodkin, 2008). More recently, the availability of whole genome sequence of soybean identified additional 5 putative *GmCHS* genes in soybean genome (Section 2.3.1). Previously it has been shown that *GmCHS7* and *GmCHS8* are critical for isoflavonoid biosynthesis and accumulation in soybean seeds (Dhaubhadel et al., 2007). A subcellular localization study revealed GmCHS7 is localized to the cytoplasm, whereas GmCHS8 is localized to the nucleus and cytoplasm (Dastmalchi et al., 2016). Thus, members of the same gene family were differentially expressed and have shown functional divergence. Other multigene families encoding enzymes in this pathway, such as chalcone isomerase (*GmCHI*s) (Dastmalchi and Dhaubhadel, 2015), chalcone reductase (*GmCHR*s) (Graham et al., 2007; Liu, 2009; Wu et al., 2014) and isoflavonoid transferases (Dhaubhadel et al., 2008), exhibit genetic and functional variations.

The majority of phenylpropanoid biosynthetic genes are regulated by MYB transcription factor (TF) family members (Jin and Martin, 1999; Boddu et al., 2006; Albert et al., 2014; Lin-Wang et al., 2014). In *Arabidopsis*, an MYB transcription factor known as anthocyanin pigment 1 (PAP1) co-regulates PAL, CHS, and bifunctional dihydroflavonol 4-reductase (DFR) and affects anthocyanin biosynthesis (Borevitz et al., 2000). In maize, PAL, CHS, flavanone 3-hydroxylase (F3H) and DFR genes involved in anthocyanin biosynthesis are theorized to be regulated coordinately by the bHLH protein-encoding gene R and the MYB gene C1 (Grotewold et al., 1998; Bruce et al., 2000; Kong et al., 2012). In soybean, phenylpropanoid pathway genes were activated by heterologous expression of the maize C1 and R transcription factors, and showed decreased genistein and increased daidzein levels with a small overall increase in total isoflavone levels (Yu et al., 2003). Recently, a study in gene regulatory network for phenylpropanoid biosynthesis in maize showed 11 TFs to be involved in the coregulation of pathway genes by recognizing 10 or more promoters (Yang et al., 2016).

Manipulation of any structural gene involved in the phenylpropanoid biosynthesis could potentially affect the production of one or more downstream metabolite(s). A previous study from the Dhaubhadel laboratory showed that down regulation of *GmCHS8* reduced isoflavonoid levels in soybean hairy roots (Yi et al., 2010). This was achieved by RNAi

silencing of *GmMYB176* in soybean and monitoring its effect on *GmCHS8* expression and metabolite accumulation. GmMYB176 regulates multiple genes in the isoflavonoid biosynthetic pathway, which may, in turn, affect the production of metabolites upstream or downstream of isoflavone aglycones. Gene regulation is a dynamic process and changes in transcript level may not necessarily correlate with the protein level or the metabolite level. Since transcriptomic studies predict changes in gene expression, whereas metabolomic studies investigate the actual products resulting from the gene action and the transcript patterns and metabolite profiles are temporally and spatially regulated, it is essential to investigate simultaneous transcriptional and metabolic changes in response to treatment or environmental changes. Therefore, the integration of these two approaches should lead to a better understanding of isoflavonoid biosynthesis in soybean.

Here using transcriptomic and metabolomic analysis of *GmMYB176* RNAi silenced "GmMYB176-Si" and *GmMYB176* overexpressed "GmMYB176-OE" hairy roots, I identified a total of 75 differentially expressed genes (DEGs) and 995 differentially produced metabolites in GmMYB176-Si hairy roots, and 7099 DEGs and 149 differentially produced metabolites in GmMYB176-OE hairy roots. A targeted analysis of the isoflavonoid pathway identified 11 isoflavonoid biosynthetic genes and 6 metabolites as differentially regulated in GmMYB176-OE and GmMYB176-Si soybean hairy roots.

4.2 Materials and methods

4.2.1 Plasmids construction

The coding region of *GmMYB176* was amplified from soybean root cDNA and recombined into the Gateway cloning entry vector pDONR/Zeo (Invitrogen, USA) as described in section 3.2.2. The resulting 'entry clone', was used in an LR recombination reaction with the destination vectors, pK7GWIWG2D (II) and pK7WG2D (Karimi et al., 2002), to obtain silencing (pGmMYB176-Si) and overexpression (pGmMYB176-OE) constructs, respectively. The LR reaction products were transformed into *E. coli* DH5α

by electroporation, plasmid extracted and sequenced, and then transformed into *Agrobacterium rhizogenes* K599 by electroporation.

4.2.2 Plant material, growth condition, and generation of soybean hairy roots

To obtain soybean cotyledons for hairy root transformation, soybean (*Glycine max* L. Merr.) cv. Harosoy63 seeds were washed with 70% ethanol (v/v) containing 3% H_2O_2 (v/v) for 2 minutes and then rinsed with water prior to planting. Seeds were planted in vermiculite and grown for 6 days in a growth chamber with a 16 h light cycle at 25ºC, and an 8 h dark cycle at 20ºC, with 60-70% relative humidity.

To generate hairy roots, six-day-old soybean cotyledons were harvested and transformed with *A. rhizogenes* K599 carrying silencing (pGmMYB176-Si) or overexpressing (pGmMYB176-OE) constructs, according to the method from Subramanian et al. (2005). Transgenic hairy roots were selected using a GFP marker 20-30 days post-infiltration, under a Leica MZ FL III fluorescence stereomicroscope with a GFP filter. Control soybean hairy roots were also generated using *A. rhizogenes* K599 and collected 20-30 days post-infiltration. The root samples were flash frozen, and stored at -80ºC until use.

4.2.3 RNA extraction, transcriptome sequencing and read pre-processing

RNA was extracted from four pooled replicates (100 mg each) of GmMYB176-OE, GmMYB176-Si and control soybean hairy root samples using RNeasy plant mini kit (Qiagen, USA). The RNA quality and concentration were checked in Agilent Bioanalyzer (Agilent, USA). RNA was diluted to 100 ng/µL and 4 µg of RNA was used for the RNA sequencing library preparation. TruSeq RNA sequencing libraries were constructed following the standard preparation guide (Illumina, USA). All12 samples were multiplexed in a lane of a flow cell, and paired-end sequencing (125 cycles) was completed using the Illumina HiSeq 2500 platform at the National Research Council of Canada (Saskatoon).

Table 4.1 Sequence of oligonucleotides used for qPCR

Gene	Primer name	Sequence (5' to 3')	Amplicon size (bp)
GmMYB176	MYB176F	GAACAGGTCGAGAATCAAGACA	103
	MYB176R	GATACACTGGCATCGCTGGAAA	
GmCHS8	CHS8/qRTf1	GCTCCCATTTAATTGATTTCTGAA	245
	CHS78/2EXNr	GACTTGTCACACATGCGCTGGAA	
GmIFS1	IFS1qF	AAACTAGTGTCTTGCCATCGC	158
	IFS1qR	GGGCCGGGAAAAAAAAATTGTC	
Gm20C4H	qGm20C4H-F	ACTACTCTGCCCCCAGGTCC	168
	qGm20C4H-R	AAGCTCGGGGTCAGAGACCA	
Gm05F3'H	qGm05F3'HF	GAGAACGGGACTGACCCAAA	230
	qGm05F3'HR	TCAGGGTCAAAAAGAAAAGAACA	
GmPT10a	PT10a-QF1	AATCGCTTCATCGTGGACGG	157
	PT10a-QR1	TGGCCTCTTCAACACAACGG	
CON4	CON4F	GATCAGCAATTATGCACAACG	106
	CON4R	CCGCCACCATTCAGATTATGT	

Before read mapping and expression quantification, all reads were filtered using Trimmomatic version 0.36 (Bolger et al., 2014) by (i) removing adapter sequences, (ii) trimming leading and trailing low quality sequences (below 15), (iii) removing sequences when the average quality per base dropped below 15 within a 4-base wide sliding window, and (iv) keeping only those pairs where both reads were longer than 75 bp.

4.2.4 Read mapping and differential expression profiling

Pair-end sequencing reads were aligned end-to-end to the *G.max* reference genome (Gmax_275_v2.0; softmasked sequence) with STARv2.5.2b (Dobin et al., 2013), allowing a maximum of two mismatches in an alignment. Read alignments were reported only if both reads in a pair were mapped concordantly. The maximum and minimum intron lengths were set at 5000 and 20 bp, respectively. Alignments that contained non-canonical junctions were filtered out. For all other parameters, default settings were used.

Gene-level raw read counts were obtained using the htseq-count tool from the HTSeq *python* library (Anders et al., 2015), using the "intersection_nonempty" mode and Gmax_275_wm82.a2.v1. gene_exonannotation. Raw gene counts are loaded, sample-normalized and differential expression between controls with GmMYB176-OE or GmMYB176-Si samples was evaluated using DESeq2 v1.14.0 (Love et al., 2014).

GO annotations categories and pathway enrichment analysis were carried out using Glyma identifiers of DEGs in PhytoMine (http://phytozome.jgi.doe.gov/phytomine/). Data from KEGG and PlantCyc (Plant Metabolic Network (PMN), www.plantcyc.org) resources were used to conduct pathway enrichment analyses, with the *Glycine max* database selected as the reference. The Bonferroni statistical test was used to adjust the *p* values for multiple hypotheses.

4.2.5 Quantitative PCR

For qPCR studies, 1 µg of total RNA was used for reverse transcription using the ThermoScript^{TM} RT-PCR System (Invitrogen, USA). SsoFast EvaGreen Supermix kit (BioRad, USA) and primers described in Table 4.1were used for qPCR. Reactions were analyzed in a Bio-Rad C1000 Thermal Cycler with the CFX96TM Real-Time PCR

System. All reactions were performed in triplicates. A reaction performed in the absence of template was used as a negative control, and the expression was normalized to the reference gene, *CON4* (Libault et al., 2008).The data were analyzed using CFX manager (BioRad, USA).

4.2.6 Metabolite extraction

Extraction of polyphenols from soybean hairy roots was carried out by the following method. Frozen root samples were ground with liquid nitrogen and 60 mg of ground samples were extracted in 5 mL of methanol: water (80:20, v/v) containing 1.0% hydrochloric acid. The samples were sonicated first at ambient temperature for 30 min, and then at 40ºC for additional 30 min followed by incubation in a boiling water bath for 30 min with intermittent vortexing. The extracts were cooled to room temperature, shaken at 300 rpm for 4 h and centrifuged at 1000x *g* for 20 min at ambient temperature. The supernatant was filtered through a 0.45 μm syringe filter (Millipore, USA) and then dried to a pellet at 50ºC under nitrogen gas for ~30 min (until volume ≈ 1 mL). The dried pellet was washed with 1 mL of ethyl acetate twice and then dried under nitrogen gas. The dried metabolite extracts were dissolved in 50% methanol (400 μL) and filtered through a 0.2 μm filter (Millipore, USA) into a 2 mL amber glass HPLC vial.

4.2.7 LC-MS/MS data acquisition and analysis

A Q-Exactive Quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific, USA) coupled to an Agilent 1290 HPLC was used for LC-MS/MS analysis. The HPLC system was equipped with a Zorbax Eclipse Plus RRHD C18 column (2.1 x 50 mm, 1.8 μm) maintained at 35°C. Samples (5 μL each) were injected and run with a flow rate of 0.3 mL min^{-1}. Water with 0.1% formic acid and acetonitrile with 0.1% formic acid was used as mobile phases A and B, respectively. Mobile phase B was held at 0% for 0.5 min and increased to 100% over 3.0 min. It was held at 100% for 2 minutes before returning to 0% over 30 seconds. Heated electrospray ionization (HESI) conditions used are as follows; spray voltage, 3.9 kV (ESI+), 3.7 kV (ESI-); capillary temperature, 400 °C; probe heater temperature, 450 °C; sheath gas, 17 arbitrary units; auxiliary gas, 8 arbitrary units; and S-Lens RF level, 45%. Compounds were detected and monitored using a top 5

non-targeted and data-dependent MS^2. Top 5 $ddMS^2$ experiments were performed in both positive and negative ionization modes by first acquiring a full MS spectrum between m/z 100-1200 at 35,000 resolution, automatic gain control (AGC) target of 3 e6, maximum injection time (IT) of 128 ms and intensity threshold of 8.0 e5. MS/MS spectra were collected at 17,500 resolution, AGC target 1e6, maximum IT 60ms and isolation window of 1 m/z. Normalized collision energy (NCE) of 30 was used fortheddMS2 method. Data analysis was performed using the manufacturer's proprietary Xcalibur™ software (Thermo Fisher Scientific, USA). Full MS peak areas were normalized with a [$^{13}C_6$] phenylalanine internal standard and compared across samples. Compounds were identified using commercial standards when possible. Thermo .raw files were converted to .mzml format using ProteoWizard (Kessner et al., 2008) with peak filter was applied, and imported into R using the XCMS package (Smith et al., 2006). Features were detected with centWave (ppm tolerance of 1.0) (Tautenhahn et al., 2008). Prefilter was 6 scans with minimum 5000 intensity, signal to noise threshold was 5, and the noise was set to 3×10^6 and 1×10^6 for positive and negative mode respectively. Retention time correction was conducted using the obiwarp method (Prince and Marcotte, 2006). Grouping included features present in at least 25% of all samples, allowable retention time deviation was 10 seconds, and *m/z* width set to 0.015. Default settings on "fillPeaks" function were used with remaining zeros imputed with two-thirds the minimum value on a per mass basis.

4.3 Results

4.3.1 Quantitative expression analysis to confirm GmMYB*176* silencing and overexpression in hairy roots

To investigate the global effect of GmMYB176 on gene expression and metabolite production, soybean hairy roots were generated where *GmMYB176* was either silenced (GmMYB176-Si) or overexpressed (GmMYB176-OE). These transgenic hairy roots, also co-expressed GFP proteins as the plasmids have GFP marker as a separate cassette. The roots with strong GFP expressions were collected and pooled. Each hairy root with GFP represents an individual transgenic line. RNA extracted from 10 biological replicates of

GmMYB176-Si and control hairy roots, and 9 replicates of GmMYB176-OE hairy roots were used for qPCR analysis to determine the expression levels of *GmMYB176* in GmMYB176-Si and GmMYB176-OE hairy roots (Appendix 4). The *GmMYB176* expression was significantly altered in GmMYB176-Si and GmMYB176-OE tissues and a difference in the level of silencing/overexpressing among the replicates was also observed (Appendix 4). Four replicates from each sample category were chosen for RNAseq experiment (Fig 4.1). In these replicates, the transcript level of *GmMYB176* was either significantly (2-times) reduced in GmMYB176-Si tissues compared to controls at $p<0.05$, or significantly (1.8-times) higher in GmMYB176-OE tissues compared to controls.

4.3.2 Data quality and coverage of soybean transcriptome

The soybean hairy root transcriptomes of GmMYB176-Si, GmMYB176-OE, and the control were studied. Hairy root mRNA from four biological replicates for each of the three samples were sequenced using 100 bp paired-end reads. The resulting data were then mapped against the soybean transcriptome 5 v2.0 (Wm82.a2.v1). The mapping and quality information including total reads per RNA sequencing library, mapped reads, uniquely mapped reads and gene coverage (%) are shown in Table 4.2. The coverage of the 56,044 protein-coding loci ranged from 74.74% to 77.04%. The principal component analysis of RNAseq data showed that there is a shift in gene expressions in GmMYB176-OE and GmMYB176-Si samples (Fig 4.2). In addition, a large variation amongthe biologicalreplicates, especially in GmMYB176-Si samples and controls, were observed, indicating the heterogeneity of biological samples.

4.3.3 Functional and structural annotation of differentially expressed genes

A total of 7099 genes were differentially expressed in GmMYB176-OE tissues compared to the control, whereas, only 75 genes were differentially expressed in GmMYB176-Si tissues compared to the control at $p<0.05$ (Fig 4.3). The “Glyma IDs” of DEGs were used for GO annotation analysis in PhytoMine (https://phytozome.jgi.doe.gov/phytomine/begin.do) and were categorized based on the

following GO annotations: "Biological process", "Molecular function", and "Cellular component" (Fig 4.4, Fig 4.5). DEGs in GmMYB176-OE tissues in comparison to controls were grouped into 29 GO annotation “Biological process” categories, 17 “Cellular component” categories, and 33 “Molecular function” categories (Fig 4.4). A high number of DEGs (618) were annotated as being involved in oxidation-reduction processes (Fig 4.4). There were 375 DEGs annotated as involved in the regulation of RNA biosynthetic process or RNA metabolic process (Fig 4.4). Seventy-nine DEGs were annotated as genes involved in photosynthesis (Fig 4.4). Among the “Cellular process” categories, 626 DEGs were annotated as part of the membrane (Fig 4.4). A total of 39 DEGs was annotated as part of photosystem (Fig 4.4). From the “Molecular function” categories, 2254 DEGs were annotated to have catalytic activity whereas, 858 and 619 DEGs were annotated as transferase activity and oxidoreductase activity, respectively (Fig 4.4). In addition, there were 253 DEGs annotated as transcription factor activity (Fig 4.4). DEGs in GmMYB176-Si tissues relative to controls were grouped into 11 categories by GO annotation “Biological process” and 2 categories by “Molecular function” (Fig 4.5). There were 12 DEGs annotated to be involved in the organonitrogen compound metabolic process, 10 were annotated as peptide biosynthetic/metabolic process, 5 were related to photosynthesis (Fig 4.5). In order to retrieve differentially expressed phenylpropanoid pathway gene candidates from GmMYB176-Si and GmMYB176-OE samples, a pathway enrichment analysis were performed (Fig 4.6). Among the 7099 DEGs in GmMYB176-OE samples, 104 phenylpropanoid pathway genes were identified, but PhytoMine categorized several peroxidases in this category leaving just one trans-cinnamate 4-monooxygenase (Glyma.20G114200), and two methyltransferases (Glyma.01G187700 and Glyma.11G054500) that are related to the phenylpropanoid pathway. Hence, a targeted approach was carried out to find isoflavonoid biosynthetic pathway specific genes that were differentially expressed in GmMYB176-Si and GmMYB176-OE samples compared to controls (Table 4.3). Out of 75 DEGs in GmMYB176-Si tissues, 5 in “Photosynthesis - antenna proteins” and 6 in “ribosome” categories were enriched (Fig 4.6).

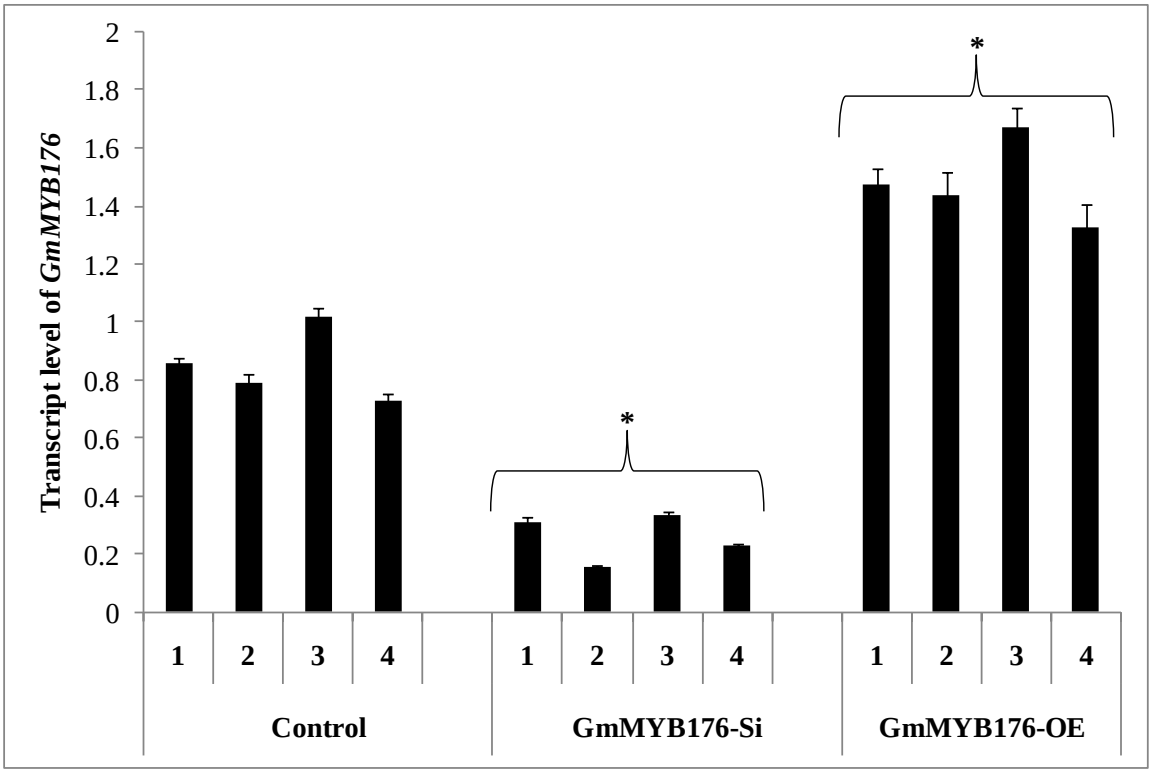


Figure 4.1 Expression levels of *GmMYB176* in the samples used for RNAseq analysis. GmMYB176-Si, GmMYB176-OE, and control samples were chosen from 10 biological replicates (Appendix 2). The cDNA samples from four replicate for each of three samples was used for RNAseq experiment. Error bars indicate SEM. Values were normalized against the reference gene, *CON4*. Asterisks represent the significant differences with the control at $p<0.05$.

Table 4.2 RNA sequencing reads and coverage of soybean transcriptome

Sample	Replicate	Total raw reads	Clean reads	Total mapped reads	Uniquely mapped reads	Unmapped reads	Gene coverage (%)
Control	1	13389565	10726040	9727667	9602622	998373	74.74
	2	22454365	19061738	17349519	17130185	1712219	76.02
	3	15761275	13593008	12261354	12112510	1331654	75.31
	4	18983435	16253950	14871395	14681029	1382555	76.76
GmMYB176-Si	1	13997019	12693897	11418283	11266836	1275614	74.74
	2	19032775	15750216	14305367	14107464	1444849	75.39
	3	17811575	16126313	14871462	14683603	1254851	76.43
	4	20314253	17692384	16269587	16058374	1422797	77.04%
GmMYB176-OE	1	19556135	16808681	15581413	15371028	1227268	76.76
	2	15777590	13872484	12904965	12736516	967519	76.22
	3	19070654	14875590	13694580	13503216	1181010	76.18
	4	15257415	11649723	10742872	10596737	906851	75.49

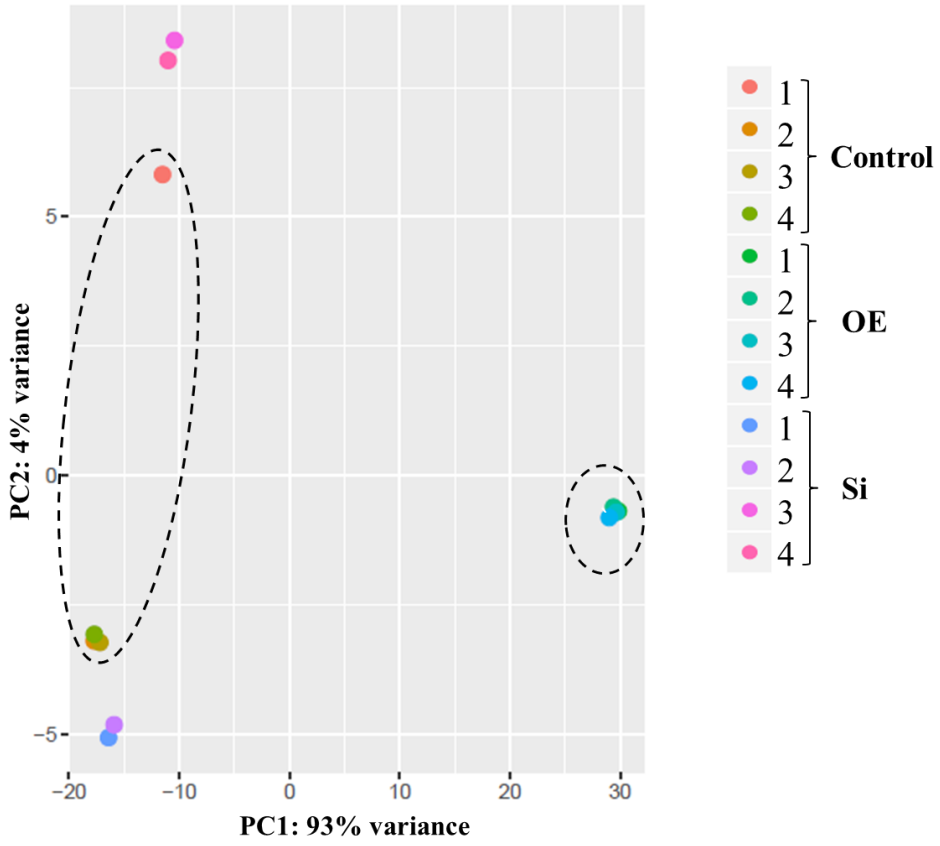
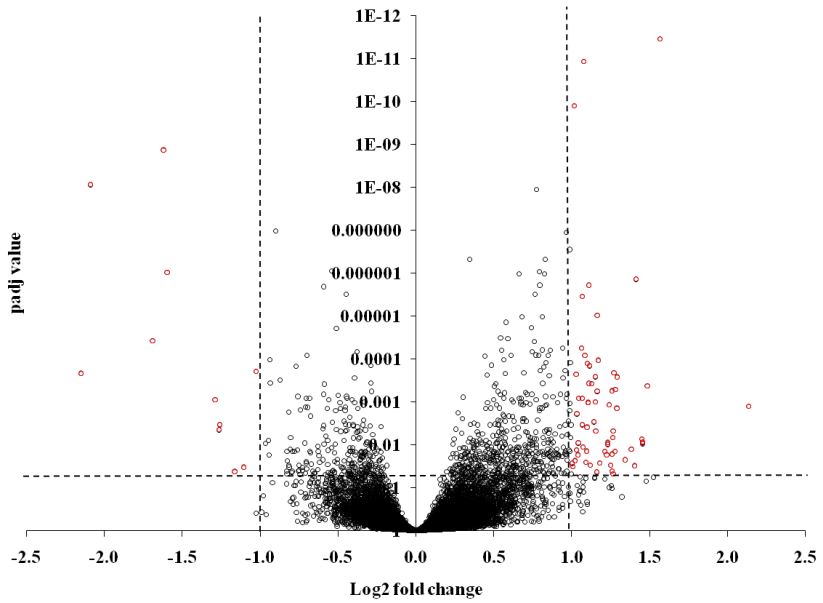


Figure 4.2 Principal component analysis of RNAseq experiment. PCA plot reveals there is a shift in gene expression profile among samples. OE, GmMYB176-OE;Si, GmMYB176-Si.The number indicates replicate a number of the sample.

a)


padj value
1E-12
1E-11
1E-10
1E-09
1E-08
0.000000
0.000001
0.00001
0.0001
0.001
0.01
1
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
Log2 fold change

b)

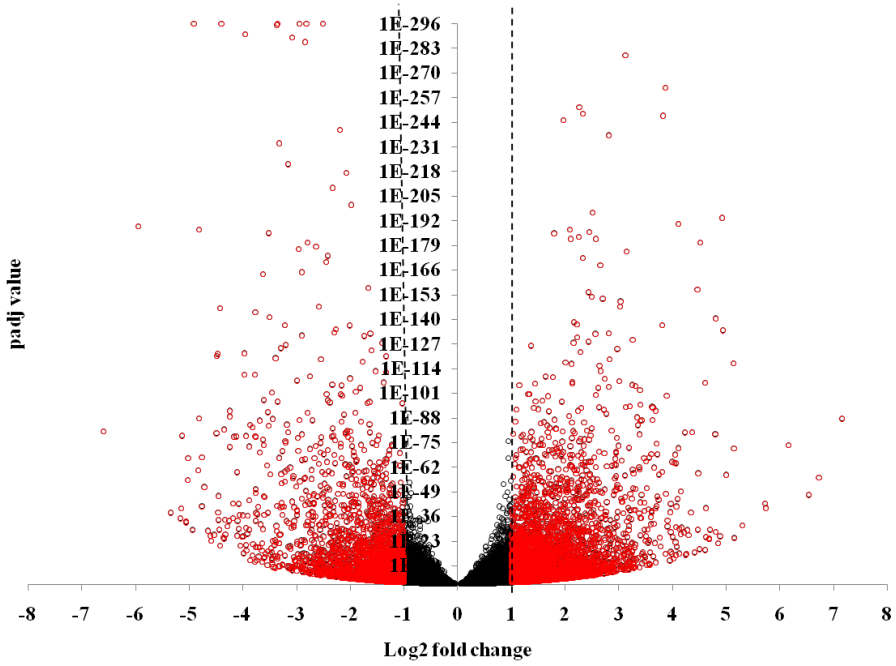

padj value
1E-296
1E-283
1E-270
1E-257
1E-244
1E-231
1E-218
1E-205
1E-192
1E-179
1E-166
1E-153
1E-140
1E-127
1E-114
1E-101
1E-88
1E-75
1E-62
1E-49
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Log2 fold change

Figure 4.3 Volcano plots of DEGs in GmMYB176-Si and GmMYB176-OE samples relative to controls. a) GmMYB176-Si vs control; b) GmMYB176-OE vs control. Red circles indicate the significantly DEGs at $p<0.05$.

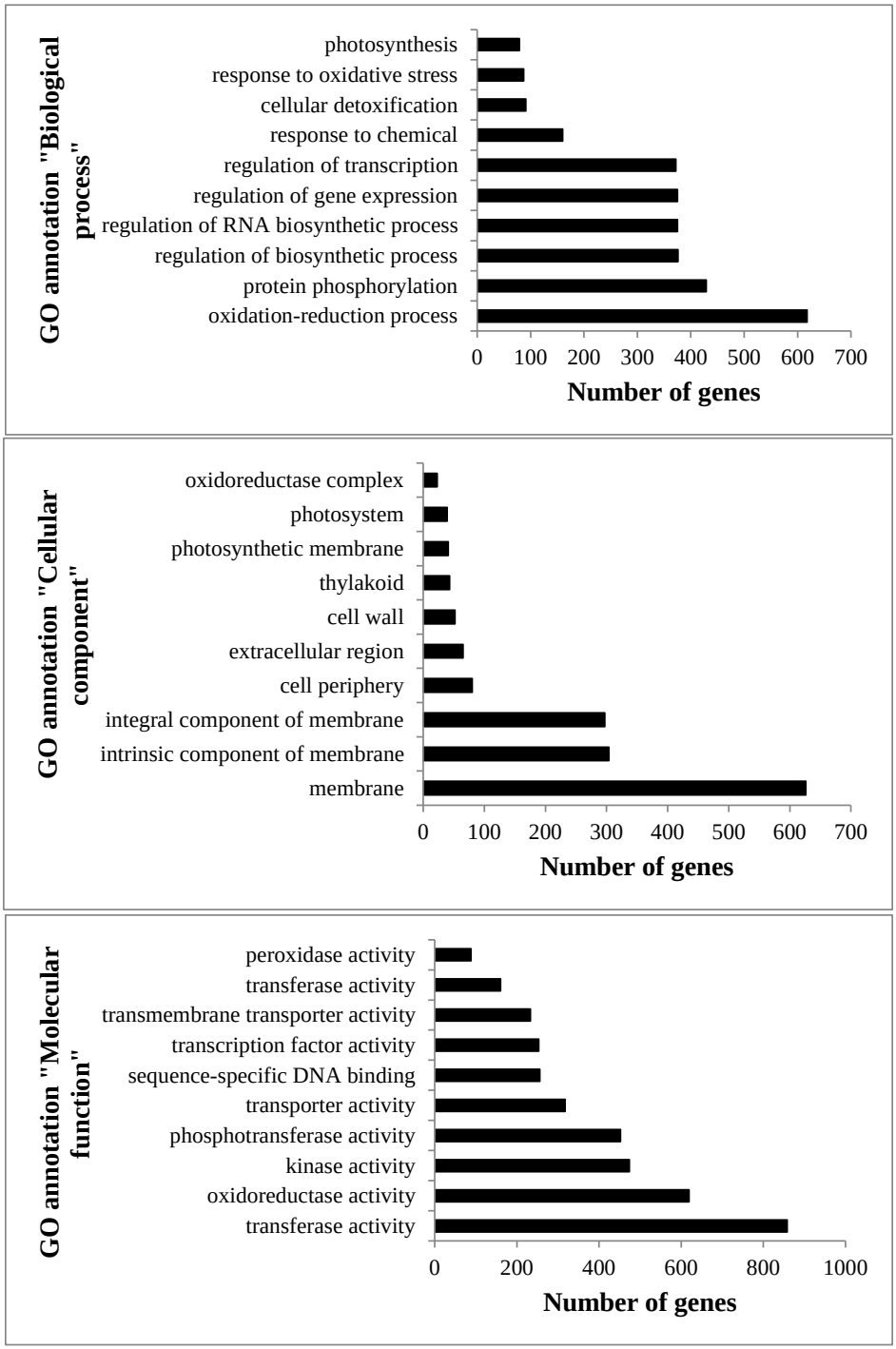

GO annotation "Biological process"
photosynthesis
response to oxidative stress
cellular detoxification
response to chemical
regulation of transcription
regulation of gene expression
regulation of RNA biosynthetic process
regulation of biosynthetic process
protein phosphorylation
oxidation-reduction process
0 100 200 300 400 500 600 700
Number of genes
GO annotation "Cellular component"
oxidoreductase complex
photosystem
photosynthetic membrane
thylakoid
cell wall
extracellular region
cell periphery
integral component of membrane
intrinsic component of membrane
membrane
0 100 200 300 400 500 600 700
Number of genes
GO annotation "Molecular function"
peroxidase activity
transferase activity
transmembrane transporter activity
transcription factor activity
sequence-specific DNA binding
transporter activity
phosphotransferase activity
kinase activity
oxidoreductase activity
transferase activity
0 200 400 600 800 1000
Number of genes

Figure 4.4 GO annotations of 7099 DEGs in GmMYB176-OE compared to control roots in soybean. (a) biological processes; (b) cellular component; (c) molecular function, using the Gene Ontology Database annotations of Glyma identifiers in PhytoMine (https://phytozome.jgi.doe.gov/phytomine/begin.do).

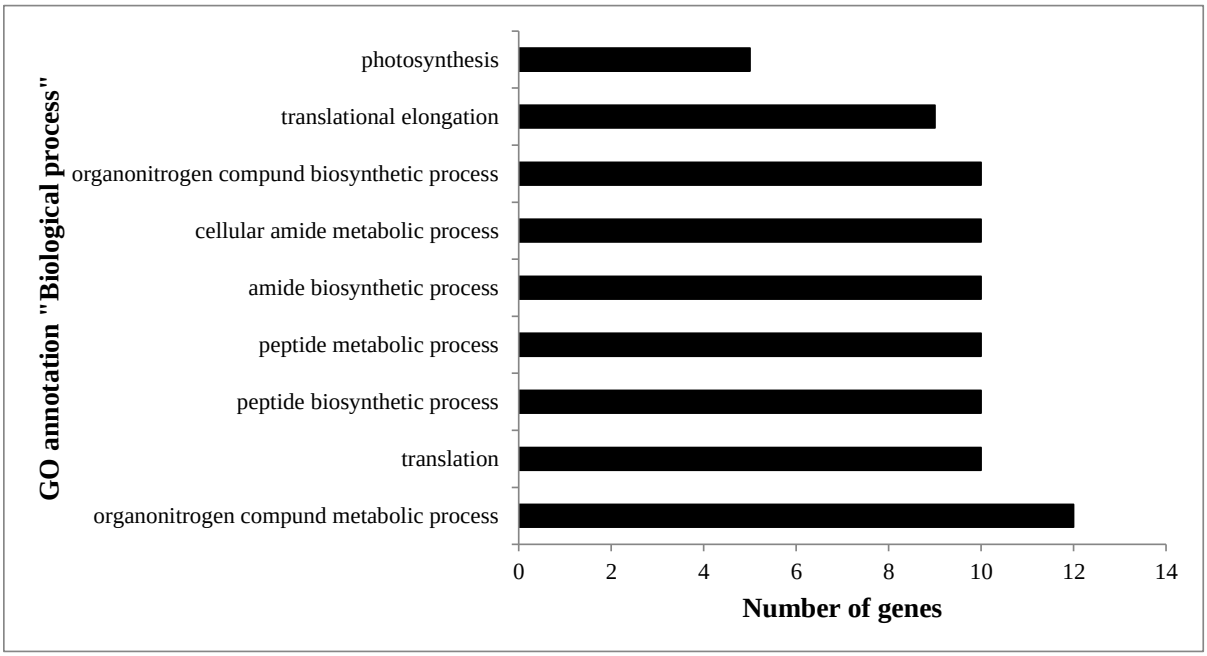


Figure 4.5 GO annotations of 75 DEGs in GmMYB176-Si compared to control hairy roots in soybean. DEGs categorized based on biological processes; Two molecular function categories such as a structural constituent of ribosome (8 genes) and structural molecule activity (8 genes) (not shown) were identified with the Gene Ontology Database annotations of Glyma identifiers in PhytoMine (https://phytozome.jgi.doe.gov/phytomine/begin.do).

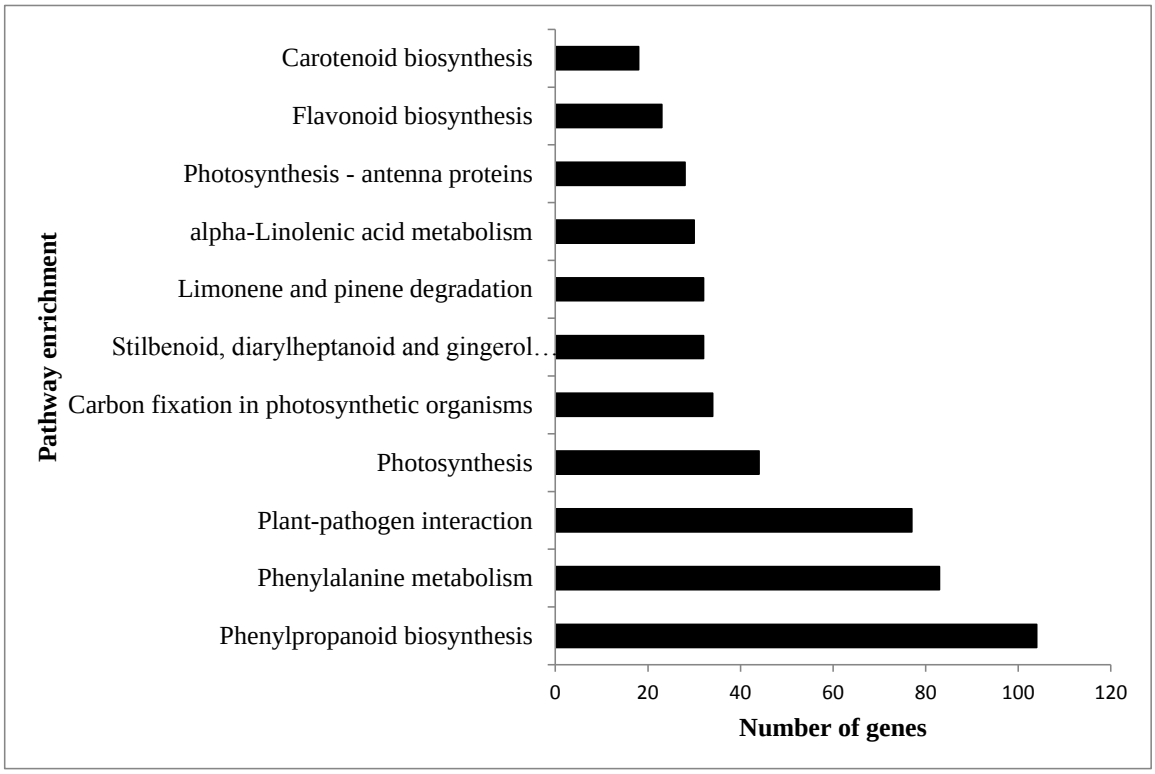


Figure 4.6 Pathway enrichment of DEGs in GmMYB176-OE compared to control hairy roots. Enrichment analysis was carried out using Glyma identifiers of DEGs in PhytoMine (https://phytozome.jgi.doe.gov/phytomine/begin.do).

Table 4.3 Effect of silencing and overexpression of *GmMYB176* on the expression of isoflavonoid biosynthetic genes.

	GmMYB176-OE vs Control		**GmMYB176-Si vs Control**		
Locus ID	**log2 Fold Change**	**p-value**	**log2 Fold Change**	**p-value**	**Annotation**
Glyma.20G114200	-4.11	1.90E-26	-0.29	0.84	*Trans-cinnamate 4-monooxygenase (Gm20C4H)*
Glyma.19G105100	1.29	0.015	0.65	0.63	*Naringenin-chalcone synthase (GmCHS13)*
Glyma.08G109500	1.17	1.24E-29	0.26	0.19	*Naringenin-chalcone synthase (GmCHS9)*
Glyma.08G109400	1.17	2.08E-10	0.52	0.08	*Naringenin-chalcone synthase (GmCHS1)*
Glyma.11G011500	-1.01	4.75E-20	0.11	0.75	*Naringenin-chalcone synthase (GmCHS8)*
Glyma.02G130400	-1.33	0.0006	0.16	0.92	*Naringenin-chalcone synthase (GmCHS10)*
Glyma.01G091400	-1.35	0.0001	-	-	*Naringenin-chalcone synthase (GmCHS11)*
Glyma.09G075200	-1.53	1.25E-05	0.32	0.76	*Naringenin-chalcone synthase (GmCHS6)*
Glyma.10G292200	-1.83	3.29E-38	-0.07	0.89	*Chalcone-flavonone isomerase 1 (GmCHI1B2)*
Glyma.07G202300	-1.49	4.02E-62	0.04	0.92	*2-hydroxyisoflavanone synthase (GmIFS1)*
Glyma.10G295300	-1.66	4.06E-45	-0.24	0.31	*putative prenyltransferase (GmPT10a)*
Glyma.05G021800	-2.17	8.96E-21	0.12	0.85	*Flavonoid 3'-monooxygenase (Gm05F3'H)*
Glyma.05G032200	1.21	8.47E-18	0.42	0.06	SF26 - F25A4.19 protein *(GmMYB176)*

4.3.4 Differentially expressed phenylpropanoid pathway genes

All DEGs were screened for phenylpropanoid biosynthetic pathway genes using pathway gene names as keywords against the annotation list of 7099 DEGs in GmMYB176-OE and 75 DEGs in GmMYB176-Si. There were 11 DEGs identified as isoflavonoid biosynthetic pathway genes in GmMYB176-OE tissues and their log2 fold change and p value from GmMYB176-Si tissues were also compiled (Table 4.3). *CHS* genes (*GmCHS1, GmCHS6, GmCHS8, GmCHS9, GmCHS10, GmCHS11,* and *GmCHS13), GmCHI1B2, GmPT10a, GmIFS1*, *Gm20C4H*, and *Gm05F3'H* were more than 2-times differentially regulated in GmMYB176-OE in comparison to the controls. However, these candidates were not differentially regulated in GmMYB176-Si hairy roots compared to controls (Table 4.3). Members of the same gene families were found to be regulated differentially. For example, four of *GmCHS*s (*GmCHS6, GmCHS8, GmCHS10,* and *GmCHS11*) were down regulated and three of Gm*CHS*s (Gm*CHS1,* Gm*CHS9,* and *GmCHS13*) were up regulated in GmMYB176-OE tissues compared to the controls. One of the flavonoid biosynthetic pathway genes (*Gm05F3'H*) was also added in Table4.3 for qPCR validation along with four other genes (*Gm20C4H*, *GmCHS8*, *GmIFS1*, and *GmPT10a*).

4.3.5 Quantitative expression analysis to confirm differentially expressed candidates

To confirm the RNA sequencing results, a qRT-PCR was performed. The same RNA samples utilized in the preparation of the RNAseq libraries were used for qPCR analysis. Primer sequences were designed to amplify unique regions within target genes (Table 4.1). The transcript levels of *GmMYB176* in GmMYB176-Si and GmMYB176-OEtissues respectively was confirmed by qPCR (Fig 4.7). The expression level of five genes (*Gm20C4H, GmCHS8, GmIFS1, GmPT10a,* and *Gm05F3'H*) wasanalyzed by qPCR (Fig 4.7). The expression of *GmIFS1* and *GmPT10a* was significantly lower in GmMYB176-OE compared to the controls in qPCR analysis and their expression level was not significantly changed in GmMYB176-Si tissues compared to the controls (Fig 4.7). The expression of *Gm20C4H* was ~38 times higher in GmMYB176-OE tissues whereas, there

was no change in its expression in GmMYB176-Si tissues compared to the controls. The expression of *GmCHS8* was significantly lower in GmMYB176-Si tissues compared to the controls, but there was no difference in its expression level between GmMYB176-OE and the controls (Fig 4.7). Transcript abundance of *GmCHS8* in GmMYB176-Si and GmMYB176-OE tissues by qPCR analysis were consistent with the previous work (Yi et al., 2010) but did not match with the RNAseq data. The expression level of *Gm05F3'H* was significantly lower in GmMYB176-OE than the controls. However, there was no change in its expression level in GmMYB176-Si tissues relative to the controls. The expression profiles of *GmIFS1*, *GmPT10a*, and *Gm05F3'H* genes were consistent with the patterns apparent in the RNAseq analysis.

4.3.6 Differentially produced metabolites in soybean hairy roots

The PCA plot of metabolomics data showed a metabolic shift when *GmMYB176* transcript level was altered (Fig 4.8). PCA plot revealed that one replicate from each of the three samples were outliers in the plot. The outliers were included in the analysis since they indicate the heterogeneity of biological samples. Total metabolite changes in the samples were analyzed. Figure 4.9 shows a representative chromatogram of metabolites for GmMYB176-Si, GmMYB176-OE and control samples at both positive and negative MS run demonstrating the impact of *GmMYB176* silencing and overexpression on total metabolites in hairyroot samples. A large impact on metabolite production was observed in GmMYB176-Si and GmMYB176-OE hairy root tissues compared to controls. i.e. many metabolites were differentially produced in GmMYB176-Si and GmMYB176-OE hairy root tissues compared to controls (Fig 4.10). Volcano plots showed an overview picture of differentially produced metabolites with their p-value and log2 fold change information in GmMYB176-Si or GmMYB176-OE compared to controls in both positive and negative ESI-MS run. There were 995 and 149 differentially produced metabolites identified in GmMYB176-Si and GmMYB176-OE tissues, respectively (Fig 4.10, Appendix 5). Common contaminants from LC-MS/MS were removed from the list of metabolites (Appendix 5) according to the procedure described by Keller et al. (2008). The number of significantly differentially produced metabolites in GmMYB176-Si tissues was higher than GmMYB176-OE tissues relative

to controls (Fig 4.11). A higher number of significantly differentially produced metabolites were identified in the positive mode of ESI than in the negative mode of ESI in both GmMYB176-Si and GmMYB176-OE samples compared to controls. A total of 1144 metabolites were found differentially regulated in (both ESI+ and ESI- mode) GmMYB176-Si and GmMYB176-OE samples combined (Appendix 5). Identification of all of them is grueling and extremely time-consuming.

4.3.7 **A targeted approach to identify differentially produced isoflavonoid biosynthetic pathway metabolites**

The objective was to study the effect of silencing and overexpression of *GmMYB176* on metabolites accumulation, especially with regard to isoflavonoids. A targeted approach was used to identify the isoflavonoids levels that were affected by the alteration of *GmMYB176* expression in soybean hairy roots. Isoflavonoids were identified by using m/z and comparing their retention time with authentic standards (Table 4.4). Phenylalanine, liquiritigenin, glycitein, daidzin, genistin, and glyceollin were identified as differentially accumulated in GmMYB176-Si tissues compared to control. Liquiritigenin level was reduced 5-times in GmMYB176-Si tissues and elevated 2.5-times in GmMYB176-OE tissues, compared to the control. Since astandard forglyceollinwasnot available commercially, its identification was based on m/z only by comparing fragment ions of the compound as determined by Quadri et al. (2013). Figure 4.12 shows strip charts of the peak area of all the metabolites identified by the targeted approach in GmMYB176-Si, GmMYB176-OE, and controls. Phenylalanine levels were dramatically reduced when *GmMYB176* was silenced. The differential accumulation of liquiritigenin in GmMYB176-Si and GmMYB176-OE tissues can be seen in the strip chart representing each replicate of the samples. There was one outlier replicate observed in each sample for all the metabolites.

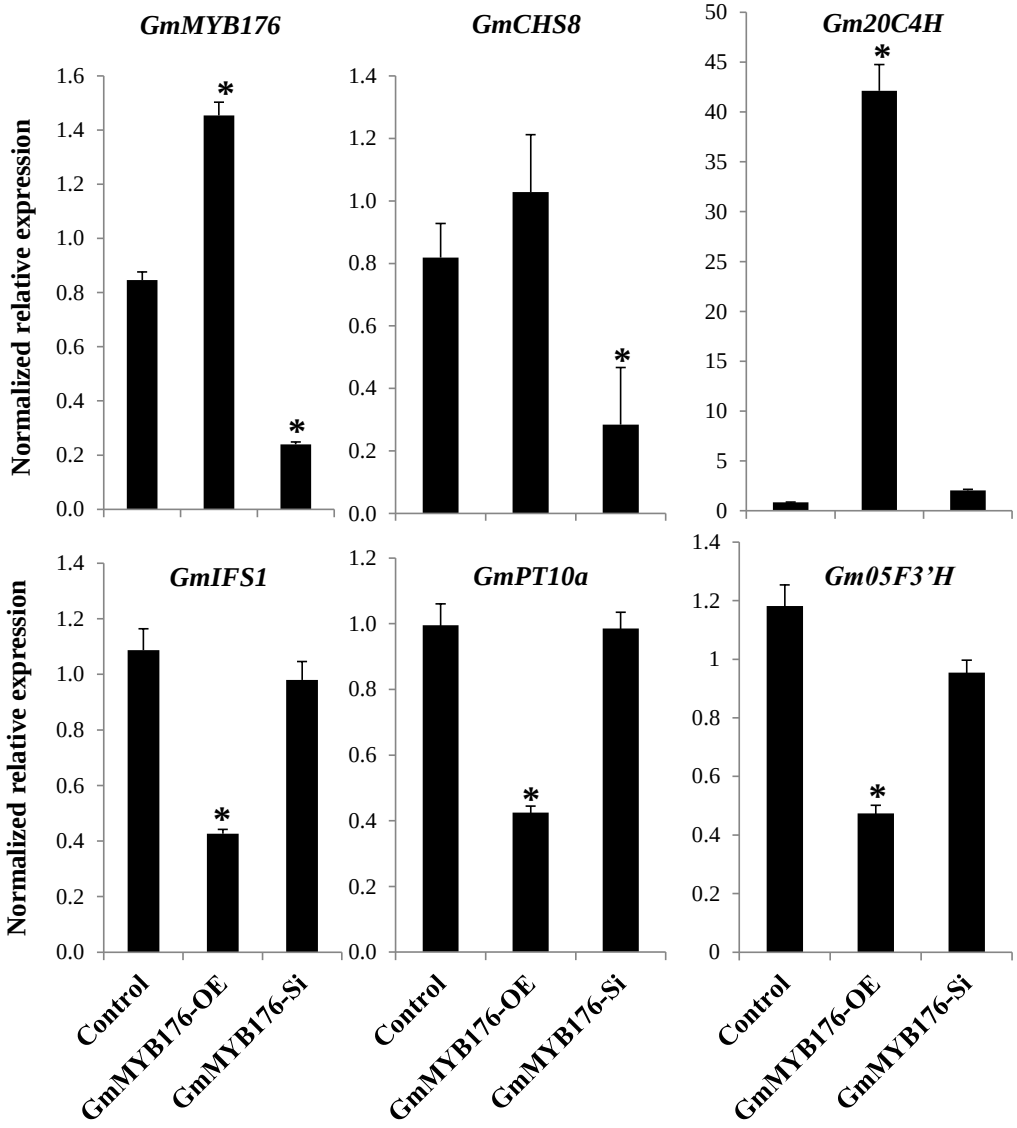


Figure 4.7 Expression analysis of candidate phenylpropanoid genes in GmMYB176-Si, GmMYB176-OE and control tissues in qPCR. The RNA samples used in RNAseq library preparation were also subjected to qPCR analysis of (iso)flavonoid genes using the gene-specific primers. Relative expression corresponds to mean gene expression in four biological replicates, with three technical triplicates. Error bars indicate SEM. Values were normalized against the reference gene *CONS4*. Asterisk indicates the significant difference in transcript level at $p<0.05$ compared to control.

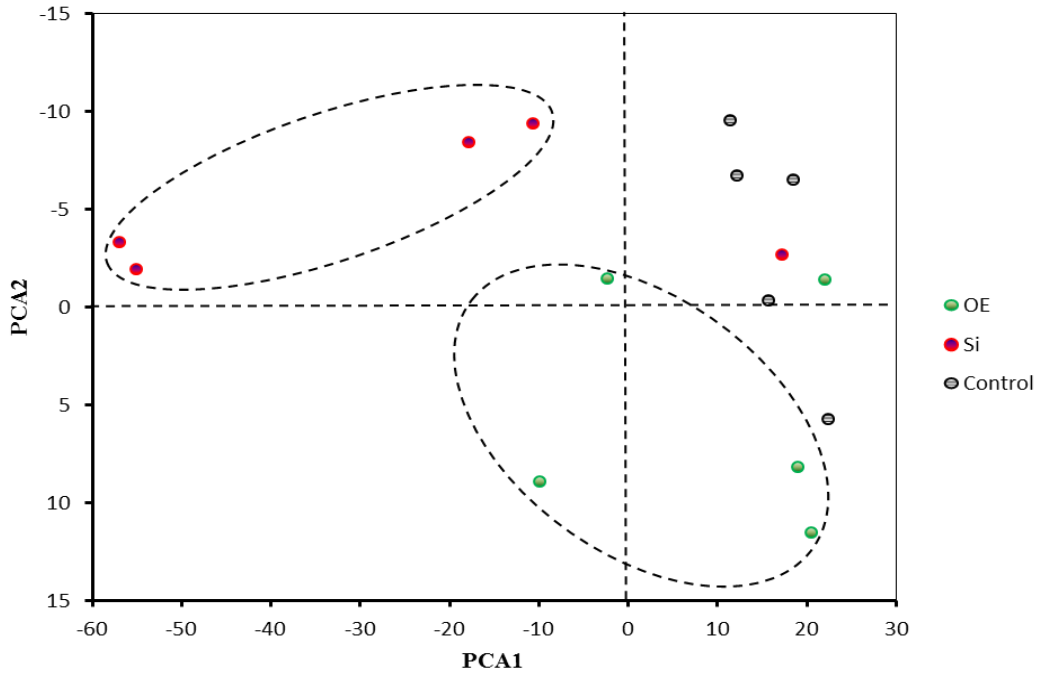


Figure 4.8 Principal component analysis (PCA) of metabolite profiles in *GmMYB176* transgenic and control soybean hairy roots. PCA reveals metabolic shifts when *GmMYB176* transcripts level is altered. Each dot represents one replicate and there were 5 biological replicates for each of the samples: GmMYB176-Si, GmMYB176-OE, and controls used for the metabolite analysis. Red circles represent Si, GmMYB176-Si; green circles represent OE, GmMYB176-OE; black circles represent control.

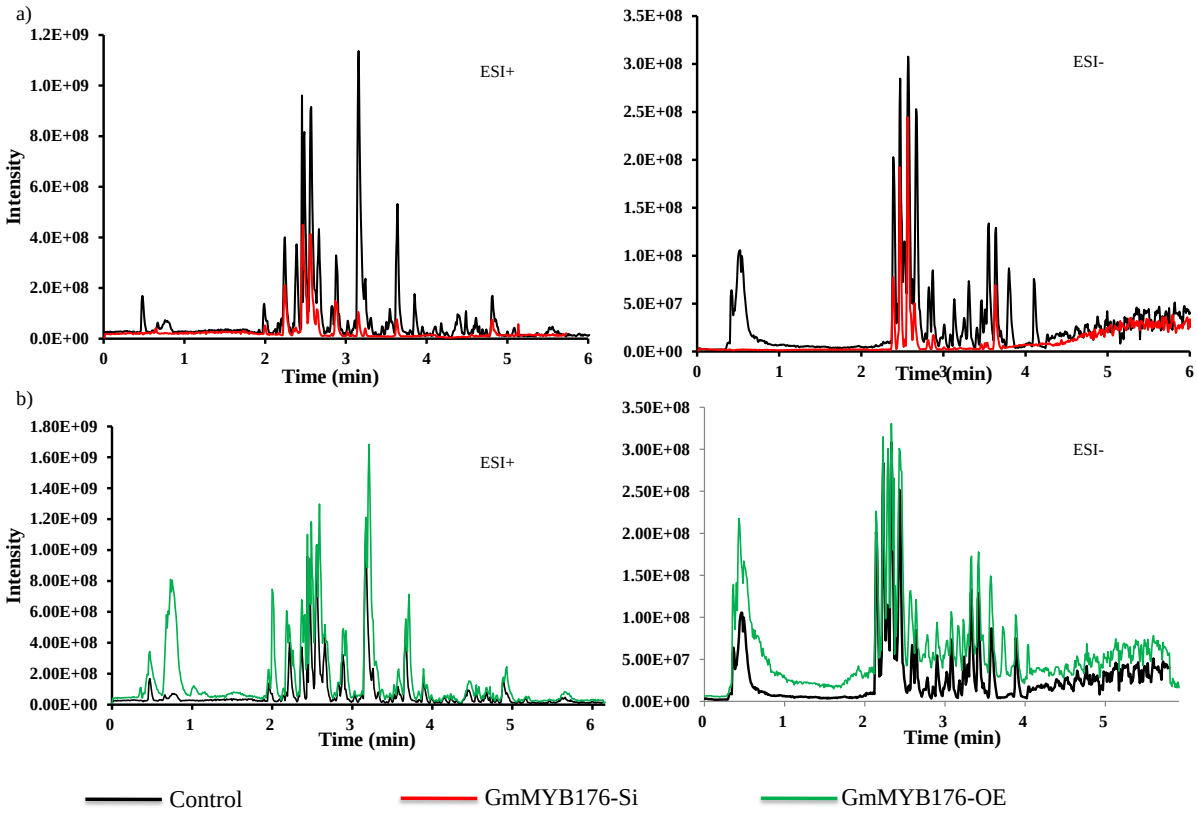


Figure 4.9 LC-MS chromatograms of GmMYB176-Si, GmMYB176-OE and control samples. a) Control and GmMYB176-Si in positive and negative mode. b) Control and GmMYB176-OE in positive and negative mode. The black, red, and green peaks represent control, GmMYB176-Si, and GmMYB176-OE, respectively.

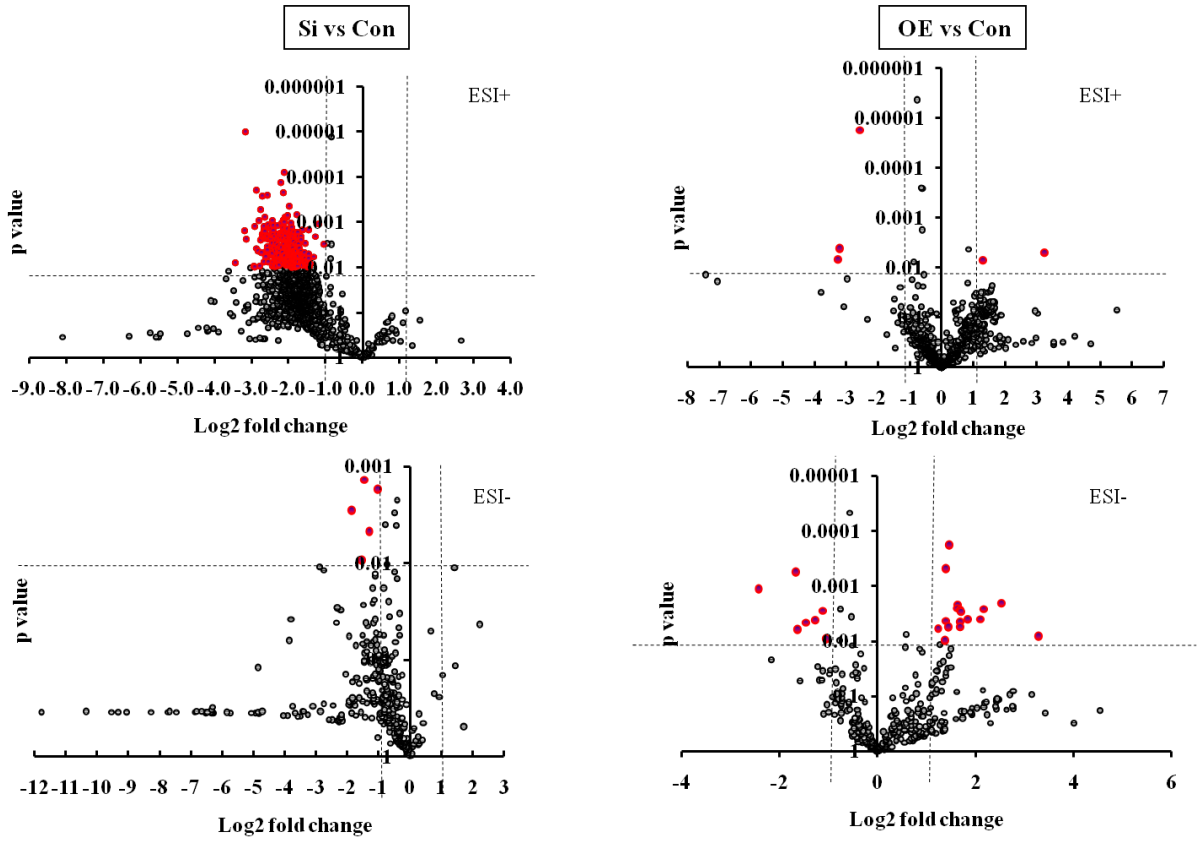


Figure 4.10 Volcano plots of differentially produced metabolites in *GmMYB176* transgenic soybean hairy roots. Red dots represent metabolites that were more than 2-times differentially produced at p<0.01. Si, GmMYB176-Si; OE, GmMYB176-OE; Con, Control.

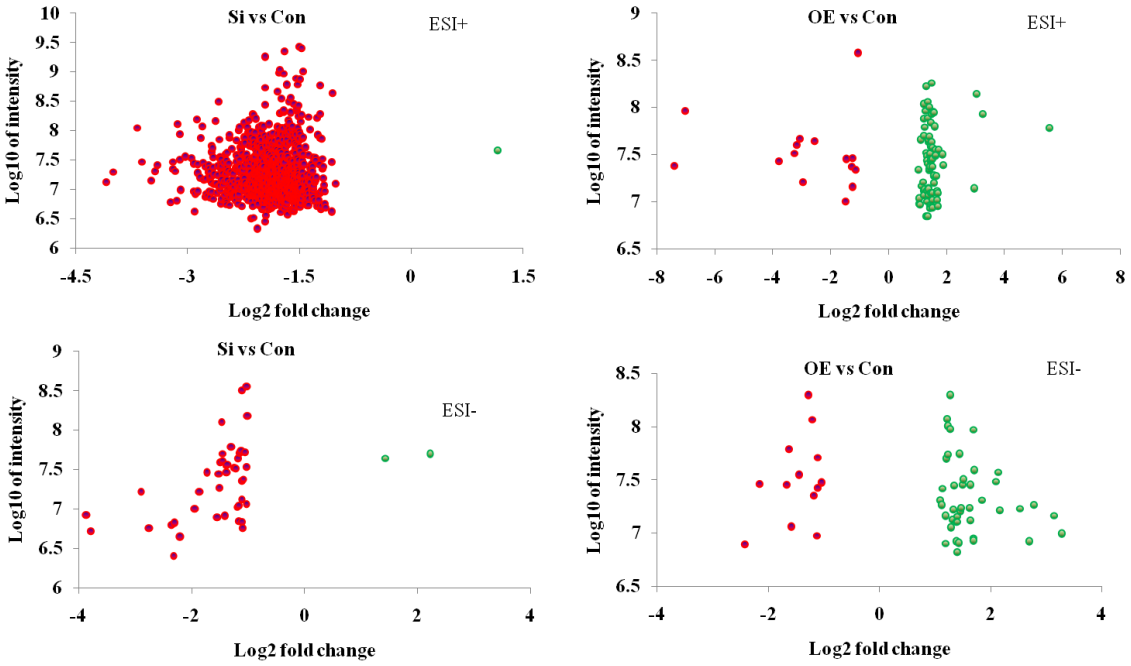


Figure 4.11 Summed intensity plot. GmMYB176-Si vs control and GmMYB176-OE vs control in both positive mode and negative mode ESI. Red and green dots represent metabolites level that arereduced and increased respectively. Si, GmMYB176-Si; OE, GmMYB176-OE; Con, Control.

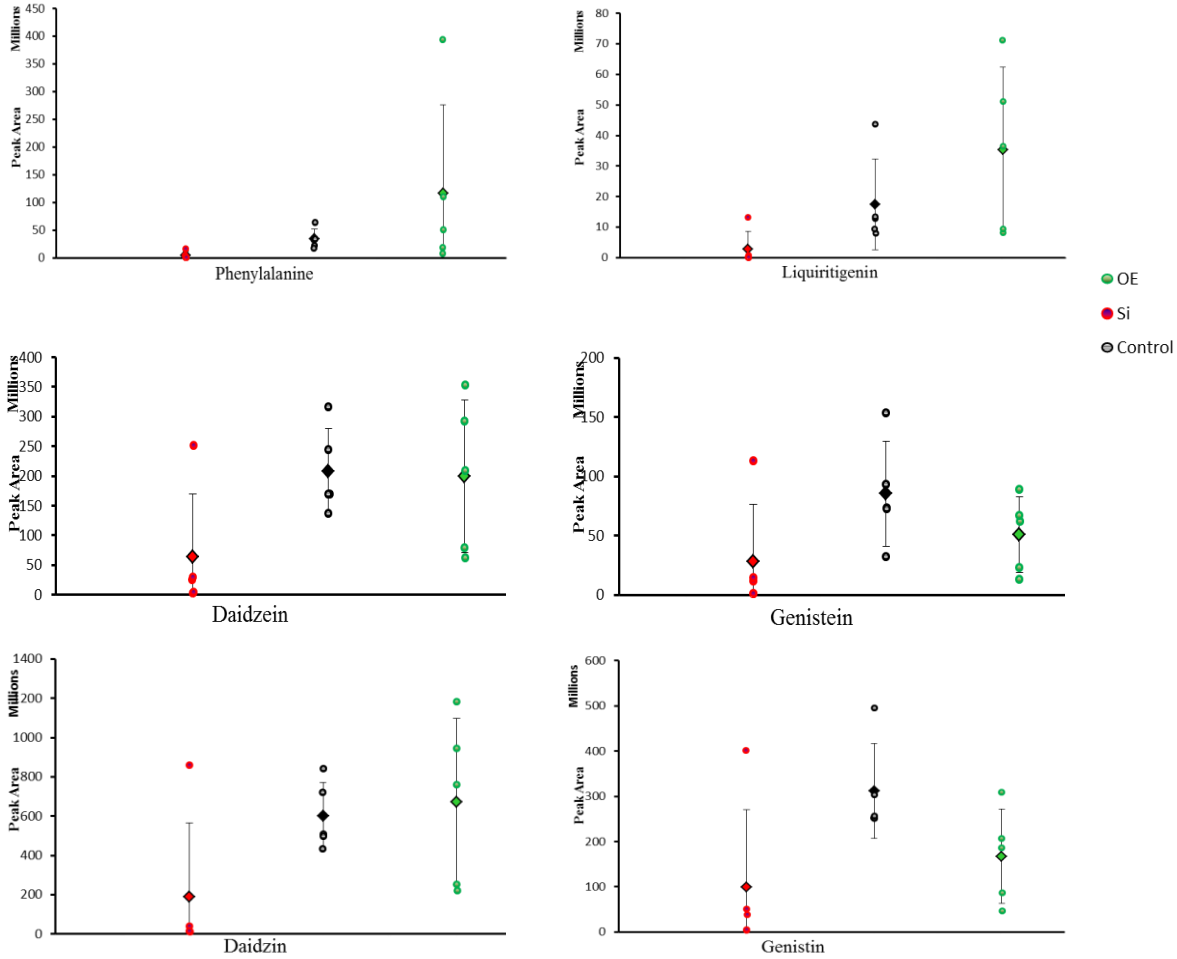


Figure 4.12 Strip charts of metabolites identified by targeted approach. Red circles represent GmMYB176-Si; Black circles represent control; Green circles represent GmMYB176-OE. Each circle represents one biological replicate of the samples.

Table 4.4 Differentially produced metabolites identified by targeted approach and confirmed by standards.

Mz	rt	P value	Log2 FC	FC	ESI	Sample (vs Con)	Metabolite
255.0643	2.39	0.060	-1.72	-3.43	+	GmMYB176-Si	Daidzein (in source fragment)
271.0592	2.54	0.091	-1.58	-3.17	+	GmMYB176-Si	Genistein (in source fragment)
417.1167	2.39	0.074	-1.67	-3.34	+	GmMYB176-Si	Daidzin
433.1117	2.54	0.067	-1.64	-3.29	+	GmMYB176-Si	Genistin
257.0842	2.83	0.084	-2.59	-5.19	+	GmMYB176-Si	Liquiritigenin
285.0750	2.90	0.076	-1.78	-3.57	+	GmMYB176-Si	Glycitein
166.0858	2.02	0.032	-2.95	-5.90	+	GmMYB176-Si	Phenylalanine
339.2882	4.42	0.081	-2.23	-4.46	+	GmMYB176-Si	Glyceollin
129.0181	1.39	0.034	-2.89	-5.79	-	GmMYB176-Si	Acetyl pyruvate related
255.0699	2.84	0.036	1.28	2.57	-	GmMYB176-OE	Liquiritigenin

mz - mass-to-charge ratio; rt - retention time; FC -fold change; ESI - electrospray ionization.

4.4 Discussion

MYB TFs coordinate the expression of multiple genes in the phenylpropanoid pathway, thus affecting the metabolite levels in plants (Vom Endt et al., 2002). An R2R3 MYB TF coordinates the expressions of *PAL*, *CHS,* and *DFR* and alters the production of flavonoids in maize (Grotewold et al., 1998; Bruce et al., 2000). GmMYB176 regulates *GmCHS8* transcript levels and affects isoflavonoid accumulation in soybean hairy roots (Yi et al., 2010). It is possible that GmMYB176 affects isoflavonoid biosynthesis by regulating multiple genes in the pathway. This study presents an overview of the regulation of isoflavonoid biosynthesis by GmMYB176 by combining the global transcripts and metabolites that are affected by the alteration of the *GmMYB176* level. Soybean hairy root system was used for silencing and overexpression of *GmMYB176*. Each soybean transgenic hairy root represents an independent transformation event and each replicate included multiple transgenic hairy roots. Hence, high sample heterogeneity was observed in the PCA plot (Fig 4.2). A total of 75 and 7099 DEGs were identified in GmMYB176-Si and GmMYB176-OE tissues respectively, as compared to controls. Soybean genome is predicted to generate ~88,000 transcripts. The transcriptome analysis of hairy roots with *GmMYB176* levels altered revealed that 0.08% of the soybean transcriptome were differentially expressed in GmMYB176-Si and 8% in GmMYB176-OE in comparison to controls. Thousands of plant genes are differentially expressed in response to stress conditions. For example, 4,729 genes and 151 metabolites were reported as differently regulated in rice when attacked by *Chilo suppressalis* larvae (Liu et al., 2016). In that study, 2% of DEGs were stress related and/or phenylpropanoid pathway related. Similarly, 1.4% of DEGs were phenylpropanoid pathway related in soybean when *GmMYB176* expression was altered (Fig 4.6). Altering the *GmMYB176* transcript level affected the expression of genes in multiple pathways in soybean.

Using a targeted approach, 12 DEGs were identified from the iso(flavonoid) biosynthetic pathway (Table 4.3) in hairy roots with modified *GmMYB176* expression, with members of *GmCHS*s family regulated differently. Table 4.3 shows three of the *GmCHS* gene family members (*GmCHS1, GmCHS9,* and *GmCHS13*) were ~2.3 times up regulated in GmMYB176-OE tissues whereas 4 *GmCHS* genes such as *GmCHS6* (2.8 times),

GmCHS8 (2 times), *GmCHS10* (2.5 times), and *GmCHS11* (2.55 times) were down regulated in the same tissues compared to controls. As compared to the control tissues, the expression of *GmCHS11* was 2.55 times down regulated in GmMYB176-OE tissues but its expression was undetectable in GmMYB176-Si tissues. *GmCHS11* might be completely silenced or the expression level was too low to be detected. A drawback in RNAseq analysis is that low abundant transcripts cannot be detected and thus may affect the interpretation of the data (Mehta et al., 2016). There are also genomic limitations to RNAseq experiments especially with plant species that have multiple gene families like soybean (Hirsch et al., 2015). In the study, it was shown that the expected count values can deviate in the case of polyploidy species. For example, the expected count values were >20% deviated for more than 25% of genes in maize. In addition, it was also shown that the transcripts of less than 600 bp have lower average expression than larger size transcripts. The reasons for variations could be due to differences in genome structures and multigene families, genes encoding short transcripts, as well as gene family size.

Figure 4.7 shows the expression level of selected, differentially expressed isoflavonoid biosynthetic genes that were validated by qPCR (Table 4.3). The expression profiles of *GmIFS1*, *GmPT10a*, and *Gm05F3'H* tested by qPCR were consistent with those analyzed by RNA-seq. The expression level of *GmIFS1* and *GmPT10a* were 2.5 times down regulated in GmMYB176-OE tissues as revealed by both qPCR and RNAseq analysis. The reduction in the level of *GmIFS1* and *GmPT10a* in GmMYB176-OE tissues may limit the production of isoflavonoids in the tissue. The expression of *Gm05F3'H* was 4.5 times lower in RNAseq analysis, but only 2 times lower in GmMYB176-OE tissues compared to controls in the qPCR analysis. *F3'H* is involved in flavonoid biosynthesis and reduction in their expression level may affect the flavonol synthesis but, may not divert the substrate flux into isoflavonoid biosynthesis. The qPCR expression level of *GmMYB176* in GmMYB176-Si tissues was different than that of RNAseq reads in GmMYB176-Si tissues (Fig 4.7, Table 4.3). The number of reads for *GmMYB176* in GmMYB176-Si tissues was close to the number of reads observed in control tissues (Table 4.3). This could be due to RNA sequencing of dsRNA produced by the RNAi construct that was integrated into soybean genome in GmMYB176-Si hairy roots. The

expression level of *GmCHS8* determined by RNAseq and qPCR did not match in both GmMYB176-Si and GmMYB176-OE tissues. The differences could be due to technical pitfalls of the experiments: 2 mismatches were allowed in RNAseq read mapping which could make a fragment unique to another member of a gene family instead of the real one; primer specificity should be a concern and a problematic one while amplifying multigene family members. Since gene-specificqPCR primers were used to amplify a unique gene product, qPCR results can be reliable in this case. Similarly, the expression data from RNAseq and qPCR did not match for the gene, *Gm20C4H* (16 times lower in RNAseq, but 38 times higher in the qPCR analysis in GmMYB176-OE tissues) (Fig 4.7, Table 4.3). Similar issues were observed in a recent study where transcriptome analysis of rice plants attacked by rice stem borer *Chilo suppressalis* showed not all gene expression results of RNAseq were correlated with qPCR (Liu et al., 2016). In this study, the expression of *zeaxanthin epoxidase*, *ZEP* was 2 times lower in RNAseq analysis at 72h post rice stem borer attack but, the expression level was not changed in the qPCR analysis (Liu et al., 2016).

Apart from phenylpropanoid pathway genes, alteration of *GmMYB176* level affected the expression of several genes involved in a wide range of biological processes. There were 79 DEGs annotated as photosynthesis-related and 36 DEGs were annotated as part of photosystem in GmMYB176-OE tissues compared to the control (Fig 4.4). GmMYB176 could be involved in the regulation of photosynthesis. Recently, the role of an MYB TF in the regulation of photosynthesis has been studied in birch. In the study, net photosynthetic and growth rates were significantly increased in *BpMYB106* overexpressing plants compared to the wild-type birch. RNAseq of *BpMYB106* overexpressed samples also showed up regulation of photosynthesis-related and oxidative phosphorylation-related genes in transgenic plants (Zhou and Li, 2016). Besides, 317 DEGs that were annotated to have transporter activity showed that *GmMYB176* probably regulate multiple biological processes and could have a wide range of target genes in soybean.

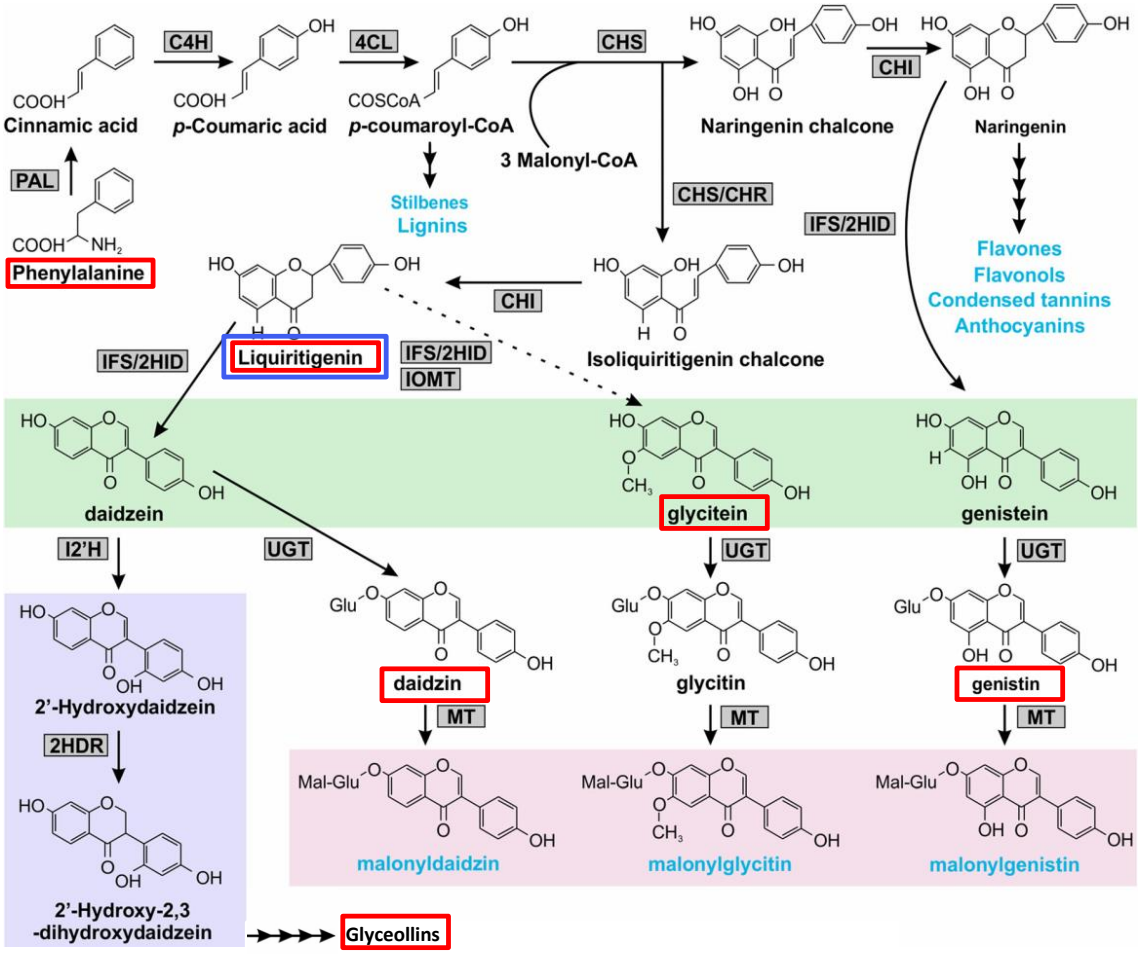
Members of the same gene families, such as *GmCHS*s, were differentially expressed when transcript level of *GmMYB176* was altered in soybean hairy roots. Three of

GmCHS genes (*GmCHS1, GmCHS9,* and *GmCHS13*) were up regulated and the others (*GmCHS6, GmCHS8, GmCHS10,* and *GmCHS11*) were down regulated in GmMYB176-OE tissues (Table 4.3). A gene expression is regulated at multiple levels. A recent emerging area of gene expression studies involves long noncoding RNAs (lincRNA), which play roles in many biological processes by gene expression control and epigenetic control (Deniz and Erman, 2016). The gene expression and function of a different isoform of the gene families could correlate with metabolite content. However, regulation of biosynthetic pathways is dynamic and has multiple layers. Hence, the level of gene expressions may not be directly correlated with levels of their ultimate end products.

Plants synthesize specialized metabolites that define the biochemical phenotype of a cell or tissue and can be viewed as the end products of gene expression (Sumner et al., 2003). A comprehensive study of cellular metabolites provides an elaborate view of the biochemical state of an organism that could be used to assess gene function (Fiehn et al., 2000). In addition, metabolic profiling can contribute significantly in identifying different compounds such as by-products of defense mechanism, abiotic and biotic stress response, molecules that are part of plant acclimation process, plant-environment interactions etc. (Weckwerth, 2003; Larrainzar et al., 2009; Guo et al., 2016). A non-targeted approach identified 995 and 149 differentially produced metabolites in GmMYB176-Si and GmMYB176-OE tissues, respectively. Identification of a large number of metabolites is laborious and time-consuming. Targeted approach has been widely used to follow the dynamics of acertain number of metabolites predicted to be responsive to a particular treatment. Additionally, comparative metabolite profiling can also be done for a large number of known metabolites (Ramalingam et al., 2015). Recently, a targeted metabolomics approach was used to understand the differences in flavonoid biosynthesis in yellow and red raspberries (Carvalho et al., 2013). Similarly in this study, from phenylalanine to glyceollins, 6 metabolites were identified by targeted approach as differentially produced in GmMYB176-Si and GmMYB176-OE samples, and their identifications were confirmed by corresponding standards. Two of isoflavone aglycones, daidzein, and genistein identified were insourced fragments of daidzin and genistin, respectively (Table 4.4). Phenylalanine is a precursor of many specialized metabolites

produced in plants and it is the first substrate that is utilized into the phenylpropanoid biosynthesis pathway, leading to the production of flavonoids, isoflavonoids, tannins, and lignin. A small change in phenylalanine metabolic level could affect the production of many downstream metabolites. Figure 4.13 illustrates the isoflavonoid biosynthesis pathway with differentially produced metabolites in GmMYB176-Si and GmMYB176-OE tissues are highlighted. Phenylalanine level was 5 times lower in GmMYB176-Sitissues relative to the control (Table 4.4) which may result in reduced production of downstream metabolites (Fig 4.13). However, cinnamic acid, the next immediate product synthesized from phenylalanine in the pathway, was not reduced in GmMYB176-Si tissues. Cinnamic acid is utilized in the further sequential steps in the pathway until naringenin chalcone or naringenin, after which the product is shared by flavonoid biosynthesis and isoflavonoid biosynthesis branches in the phenylpropanoid pathway. Interestingly, metabolites downstream of phenylalanine and cursory to liquiritigenin were not affected in GmMYB176-Si or GmMYB176-OE tissues in compared to controls.

The level of liquiritigenin was altered in both GmMYB176-Si and GmMYB176-OE tissues showing that overexpression of *GmMYB176* also has an impact on liquiritigenin level which is the only metabolite that was increased in GmMYB176-OE tissues (Fig 4.12, Fig 4.13, and Table 4.4). Liquiritigenin is the precursors for the synthesis of the isoflavone aglycones, daidzein, and glycitein. The level of glycitein was reduced in GmMYB176-Si tissues which could be due to the lower availability of liquiritigenin; interestingly the level of daidzein was unaffected in the same tissues. Daidzin and genistin, two glycosylated isoflavones were 3.3 times lesser produced in GmMYB176-Sitissues than the controls (Table 4.4). The level of daidzein in GmMYB176-Si tissues was not affected in comparisonto controls which might be due to the daidzein substrate was not being used for the production of glyceollins or daidzin (Fig 4.13). Glyceollins are phytoalexins produced from daidzein upon stress in soybean and it is the final step in induced isoflavonoid synthesis pathway (Welle and Grisebach, 1988). Glyceollins are induced metabolites and are produced by soybean upon exposure to biotic and abiotic stress (Graham et al., 1990; Graham and Graham, 1991). It is possible that cutting the root samples during sample collection induces stress response thereby leading to



Increased levels of metabolites in GmMYB176-OE tissues compared to the controls

Reduced levels of metabolites in GmMYB176-Si tissues compared to the controls

Figure 4.13 Differentially accumulated isoflavonoid biosynthetic pathway metabolites in GmMYB176-Si and GmMYB176-OE hairy roots compared to the controls. The multiple *arrows* indicate multiple steps in the pathway and the *dotted arrow* indicates speculated steps. The major metabolites synthesized from the phenylpropanoid pathway are shown in *blue* text. The three isoflavoneaglycones, daidzein, glycitein, and genistein are highlighted in *green* and malonylisoflavones are highlighted in *pink*. The induced isoflavonoid phytoalexin synthesis branch is highlighted in purple (light). PAL, phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylate; 4CL, 4-coumarate-CoA-ligase; CHS, chalconesynthase; CHR, chalcone reductase; CHI, chalcone isomerase; IFS, 2-hydroxyisoflavanone synthase; 2HID, 2-hydroxyisoflavanone dehydratase; IOMT, isoflavone *O*-methyltransferase; UGT- uridine diphosphate glycosyltransferase; MT, malonyltransferase; I2′H, Isoflavone 2′-hydroxylase; 2HDR, 2′-hydroxydaidzein reductase.[Modified from Anguraj Vadivel et al. (2015)]

glyceollin production. Reduction in the level of glyceollins in GmMYB176-Si tissues implies *GmMYB176* regulates stress responses by regulating the production of phytoalexins. There was no significant increase in the level of isoflavonoids in GmMYB176-OE tissues. Liquiritigenin is the only metabolite in the pathway to increase in concentration in GmMYB176-OE tissues. The increase in the expression level of genes (*GmCHS1*, *GmCHS9*, and *GmCHS13*) upstream of liquiritigenin may lead to increase in liquiritigenin level (Table 4.3; Fig 4.13). The increase in the level of one metabolite may not necessarily increase the accumulation of downstream metabolites in the pathway. Coordinate expression of pathway genes, substrate flux, product threshold level contribute to dynamic of pathway regulation.

4.5 **Conclusion**

GmMYB176 regulates multiple genes and metabolite production in soybean. The production of isoflavonoids was affected by altering *GmMYB176* transcript levels in soybean hairy roots. No good correlation was found between transcriptomics and metabolomics results. This explains the complexity of GmMYB176 mediated gene regulation. Based on the RNAseq results obtained, it appears that GmMYB176 alone regulates expression of some genes where a direct correlation in gene expression (such as for *GmCHS1*, *GmCHS9*, and *GmCHS13*) was observed in GmMYB176-OE and GmMYB176-Si tissues. GmMYB176 may interact with its cofactor (s) (same or different) to regulate expression of other genes such as *Gm20C4H*, *GmCHS6*, *GmCHS8*, *GmCHS10*, *GmCHS11*, *GmCHI1B2*, *GmIFS1*, *GmPT10a*, and *Gm05F3'H*. In the later case, the availability of required cofactor may define its effect on target gene expression. Phenylalanine, liquiritigenin, glycitein, daidzin, genistin, and glyceollin were differentially produced in soybean hairy roots when *GmMYB176* expression level was altered. The level of isoflavonoids (daidzin, genistin, glycitein, and glyceollin), which were reduced in GmMYB176-Si tissues, were not increased in GmMYB176-OE tissues. Transcripts levels do not necessarily correlate with protein and/or metabolite levels. Gene expression is a dynamic process, that involves multiple regulatory steps and so does the associated pathways. GmMYB176 regulates the accumulation of at least six metabolites in the isoflavonoid biosynthetic pathway in soybean.

4.6 Literature cited

Albert, N.W., Davies, K.M., Lewis, D.H., Zhang, H., Montefiori, M., Brendolise, C., Boase, M.R., Ngo, H., Jameson, P.E., and Schwinn, K.E. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. Plant Cell **26,** 962-980.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **31,** 166-169.

Anguraj Vadivel, A.K., Sukumaran, A., Li, X., and Dhaubhadel, S. (2015). Soybean isoflavonoids: role of GmMYB176 interactome and 14-3-3 proteins. Phytochemistry Reviews **15,** 391-403.

Blanc, G., and Wolfe, K.H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell **16,** 1667-1678.

Boddu, J., Jiang, C., Sangar, V., Olson, T., Peterson, T., and Chopra, S. (2006). Comparative structural and functional characterization of sorghum and maize duplications containing orthologous MYB transcription regulators of 3-deoxyflavonoid biosynthesis. Plant Molecular Biology **60,** 185-199.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30,** 2114-2120.

Borevitz, J.O., Xia, Y., Blount, J., Dixon, R.A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell **12,** 2383-2394.

Bruce, W., Folkerts, O., Garnaat, C., Crasta, O., Roth, B., and Bowen, B. (2000). Expression profiling of the maize flavonoid pathway genes controlled by estradiol-inducible transcription factors CRC and P. Plant Cell **12,** 65-80.

Carvalho, E., Franceschi, P., Feller, A., Palmieri, L., Wehrens, R., and Martens, S. (2013). A targeted metabolomics approach to understand differences in flavonoid biosynthesis in red and yellow raspberries. Plant Physiology and Biochemistry **72,** 79-86.

Dastmalchi, M., and Dhaubhadel, S. (2015). Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. Planta **241,** 507-523.

Dastmalchi, M., Bernards, M.A., and Dhaubhadel, S. (2016). Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. The Plant Journal **85,** 689-706.

Deniz, E., and Erman, B. (2016). Long noncoding RNA (lincRNA), a new paradiGm in gene expression control. Functional and Integrative Genomics **17,** 135-143.

Dhaubhadel, S., Farhangkhoee, M., and Chapman, R. (2008). Identification and characterization of isoflavonoid specific glycosyltransferase and malonyltransferase from soybean seeds. Journal of Experimental Botany**59,** 981-994.

Dhaubhadel, S., Gijzen, M., Moy, P., and Farhangkhoee, M. (2007). Transcriptome analysis reveals a critical role of *CHS7* and *CHS8* genes for isoflavonoid synthesis in soybean seeds. Plant Physiology **143,** 326-338.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics **29,** 15-21.

Dudareva, N., Pichersky, E., Werck-Reichhart, D., and Lewinsohn, E. (2010). Plant metabolism. Editorial. Molecular Plant **3,** 1.

Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R.N., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. Nature Biotechnology **18,** 1157-1161.

Graham, T.L., and Graham, M.Y. (1991). Glyceollin elicitors induce major but distinctly different shifts in isoflavonoid metabolism in proximal and distal soybean cell populations. Molecular Plant-Microbe Interactions **4,** 60-68.

Graham, T.L., Kim, J.E., and Graham, M.Y. (1990). Role of constitutive isoflavone conjugates in the accumulation of glyceollin in soybean infected with *Phytophthora-Megasperma*. Molecular Plant-Microbe Interactions **3,** 157-166.

Graham, T.L., Graham, M.Y., Subramanian, S., and Yu, O. (2007). RNAi silencing of genes for elicitation or biosynthesis of 5-deoxyisoflavonoids suppresses race-specific resistance and hypersensitive cell death in *Phytophthora sojae* infected tissues. Plant Physiology**144,** 728-740.

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

Guo, R., Shi, L., Yang, C., Yan, C., Zhong, X., Liu, Q., Xia, X., and Li, H. (2016). Comparison of Ionomic and Metabolites Response under Alkali Stress in Old and Young Leaves of Cotton (*Gossypium hirsutum* L.) Seedlings. Frontiers in Plant Science**7,** 1785.

Hirsch, C.D., Springer, N.M., and Hirsch, C.N. (2015). Genomic limitations to RNA sequencing expression profiling. The Plant Journal **84,** 491-503.

Jin, H., and Martin, C. (1999). Multifunctionality and diversity within the plant MYB-gene family. Plant Molecular Biology **41,** 577-585.

Karimi, M., Inze, D., and Depicker, A. (2002). Gateway vectors for *Agrobacterium*-mediated plant transformation. Trends in Plant Science **7,** 193-195.

Keller, B.O., Sui, J., Young, A.B., and Whittal, R.M. (2008). Interferences and contaminants encountered in modern mass spectrometry. Analytica Chimica Acta **627,** 71-81.

Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics **24,** 2534-2536.

Kong, Q., Pattanaik, S., Feller, A., Werkman, J.R., Chai, C., Wang, Y., Grotewold, E., and Yuan, L. (2012). Regulatory switch enforced by basic helix-loop-helix and ACT-domain mediated dimerizations of the maize transcription factor R. Proceedings of the National Academy of Sciences of the United States of America **109,** E2091-2097.

Larrainzar, E., Wienkoop, S., Scherling, C., Kempa, S., Ladrera, R., Arrese-Igor, C., Weckwerth, W., and Gonzalez, E.M. (2009). Carbon metabolism and bacteroid functioning are involved in the regulation of nitrogen fixation in *Medicago truncatula* under drought and recovery. Molecular Plant-Microbe Interactions **22,** 1565-1576.

Libault, M., Thibivilliers, S., Bilgin, D.D., Radwan, O., Benitez, M., Clough, S.J., and Stacey, G. (2008). Identification of four soybean reference genes for gene expression normalization. Plant Genome **1,** 44-54.

Lin-Wang, K., McGhie, T.K., Wang, M., Liu, Y., Warren, B., Storey, R., Espley, R.V., and Allan, A.C. (2014). Engineering the anthocyanin regulatory complex of strawberry (*Fragaria vesca*). Frontiers in Plant Science**5,** 651.

Liu, G.Y. (2009). Isolation, sequence identification and tissue expression profile of two novel soybean (*glycine max*) genes-vestitone reductase and chalcone reductase. Molecular Biology Reports**36,** 1991-1994.

Liu, Q., Wang, X., Tzin, V., Romeis, J., Peng, Y., and Li, Y. (2016). Combined transcriptome and metabolome analyses to understand the dynamic responses of rice plants to attack by the rice stem borer *Chilo suppressalis* (Lepidoptera: Crambidae). BMC Plant Biology**16,** 259.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology**15,** 550.

Mehta, S., Tsai, P., Lasham, A., Campbell, H., Reddel, R., Braithwaite, A., and Print, C. (2016). A study of TP53 RNA splicing illustrates pitfalls of RNA-seq methodology. Cancer Research**76,** 7151-7159.

Panchy, N., Lehti-Shiu, M., and Shiu, S.H. (2016). Evolution of gene duplication in plants. Plant Physiology**171,** 2294-2316.

Pichersky, E., and Lewinsohn, E. (2011). Convergent evolution in plant specialized metabolism. Annual Review of Plant Biology **62,** 549-566.

Prince, J.T., and Marcotte, E.M. (2006). Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. Analytical Chemistry**78,** 6140-6152.

Ramalingam, A., Kudapa, H., Pazhamala, L.T., Weckwerth, W., and Varshney, R.K. (2015). Proteomics and metabolomics: two emerging areas for legume improvement. Frontiers in Plant Science**6,** 1116.

Shoemaker, R.C., Schlueter, J., and Doyle, J.J. (2006). Paleopolyploidy and gene duplication in soybean and other legumes. Current Opinion in Plant Biology**9,** 104-109.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical Chemistry**78,** 779-787.

Subramanian, S., Graham, M.A., Yu, O., and Graham, T.L. (2005). RNA interference of soybean isoflavone synthase genes leads to silencing in tissues distal to the transformation site and to enhanced susceptibility to *Phytophthora sojae*. Plant Physiology **137,** 1345-1353.

Sumner, L.W., Mendes, P., and Dixon, R.A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry **62,** 817-836.

Tautenhahn, R., Bottcher, C., and Neumann, S. (2008). Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics **9,** 504.

Tuteja, J.H., and Vodkin, L.Q. (2008). Structural features of the endogenous *CHS* silencing and target loci in the soybean genome. Crop Science **48,** S49-S68.

Vom Endt, D., Kijne, J.W., and Memelink, J. (2002). Transcription factors controlling plant secondary metabolism: what regulates the regulators? Phytochemistry **61,** 107-114.

Weckwerth, W. (2003). Metabolomics in systems biology. Annual Review of Plant Biology **54,** 669-689.

Welle, R., and Grisebach, H. (1988). Induction of phytoalexin synthesis in soybean: Enzymatic cyclization of prenylated pterocarpans to glyceollin isomers. Archives of Biochemistry and Biophysics **263,** 191-198.

Wu, N., Wang, P., Li, D., Dai, L., Zheng, C., Lu, S., Cai, Y., Zhang, Z., Qu, J., and Xia, H. (2014). Function of chalcone reductase gene *CHR1* in soybean. Yi Chuan **36,** 707-712.

Yang, F., Li, W., Jiang, N., Yu, H., Morohashi, K., Ouma, W.Z., Morales-Mantilla, D.E., Gomez-Cano, F., Mukundi, E., Prada-Salcedo, L.D., Velazquez, R.A., Valentin, J., Mejia-Guerra, M.K., Gray, J., Doseff, A.I., and Grotewold, E. (2016). A maize gene regulatory network for phenolic metabolism. Molecular Plant **10,** 498-515.

Yi, J., Derynck, M.R., Li, X., Telmer, P., Marsolais, F., and Dhaubhadel, S. (2010). A single-repeat MYB transcription factor, GmMYB176, regulates *CHS8* gene expression and affects isoflavonoid biosynthesis in soybean. The Plant Journal **62,** 1019-1034.

Yu, O., Shi, J., Hession, A.O., Maxwell, C.A., McGonigle, B., and Odell, J.T. (2003). Metabolic engineering to increase isoflavone biosynthesis in soybean seed. Phytochemistry **63,** 753-763.

Zhou, C., and Li, C. (2016). A novel R2R3-MYB transcription factor BpMYB106 of Birch (*Betula platyphylla*) confers increased photosynthesis and growth rate through up-regulating photosynthetic gene expression. Frontiers in Plant Science**7,** 315.

5 GENERAL DISCUSSION

5.1 GmMYB176 interactome and coregulation of multiple genes in isoflavonoid biosynthesis in soybean

Soybean is a paleopolyploid and has undergone two whole genome duplication events (Shoemaker et al., 1996; Blanc and Wolfe, 2004; Gill et al., 2009). Genome sequence data indicated duplicated gene frequencies of approximately 75% in soybean (Schmutz et al., 2010). It is believed that the duplicated genes retained post-whole genome duplication could have contributed to soybean domestication traits (Schmutz et al., 2010). Duplication results in multiple gene families in the genome (Panchy et al., 2016). Most of the phenylpropanoid pathway enzymes belong to multigene families; hence studying expressions of gene family members can be challenging (Blanc and Wolfe, 2004; Shoemaker et al., 2006; Panchy et al., 2016).

Several multigene families such as *SGF14* (14-3-3 proteins), *GmCHR*s, and *GmCHI*s have been characterized in soybean. The soybean CHI family contains twelve members that are grouped into four subfamilies: GmCHI1, GmCHI2, GmCHI3, and GmCHI4. The expression of *GmCHI1A* and *GmCHI1B2* transcripts correlates with isoflavonoid accumulation in soybean seeds (Dastmalchi and Dhaubhadel, 2015). Biosynthesis and accumulation of isoflavonoids is a complex trait that involves multidimensional regulation in soybean. The complexity of the phenylpropanoid pathways emphasizes the importance of understanding the multigene families that regulate the isoflavonoid biosynthesis. CHS is one of the key enzymes in the isoflavonoid pathway. With the advantage of genome sequence availability, I have identified 5 additional *CHS* genes in the soybean genome that were not previously known. *GmCHS*s have shown different expression patterns within different soybean tissues similar to other multigene family members like *SGF14* and *GmCHI*s. Likewise, the maize genome has 14 *CHS* genes and they have shown different tissue-specific expressions (Han et al., 2016). Multigene family members usually have tissue-specific roles, where some could be inducible during stress or as part of a plant defense mechanism. The difference in their promoter sequences with different *cis*-acting elements leads to their tissue-specific or stress-dependent regulation.

Hence, not all members of the multigene families show similar spatio-temporal expression. The *GmCHS7* and *GmCHS8* are critical for isoflavonoid accumulation in soybean seeds (Dhaubhadel et al., 2007). The expression of *GmCHS8*, which is regulated by an R1 MYB TF GmMYB176, correlates with isoflavonoid accumulation level in soybean hairy roots.

Several plant structural genes are regulated by the MYB TF family. Transcription is often regulated by multiple TFs. Transcription complexes are made up of transcriptional activators/repressors that bind to the *cis*-acting elements in the gene promoter for the regulation. MYB TFs have been reported to play a key role in the regulation of the phenylpropanoid pathway (Jin and Martin, 1999; Boddu et al., 2006). Several studies have reported that MYB TFs are involved in a ternary complex with bHLH and WD40 proteins to regulate phenylpropanoid gene expression (Payne et al., 2000; Wang and Chen, 2008; Ben-Simhon et al., 2011; Shangguan et al., 2016). In this thesis, it was hypothesized that GmMYB176 works in combination with other factor(s) for the regulation of *GmCHS8* gene expression and isoflavonoid biosynthesis. Here I have identified a total of 716 GmMYB176 interacting proteins by Co-IP experiment. GO annotation of GmMYB176 interactome revealed that candidate proteins belong to a wide range of biological processes and have diverse molecular functions. The objective of the study was to identify GmMYB176 interacting TFs and co-factors that regulate *GmCHS8* gene expression. The *in planta* interaction of GmMYB176 with two bZIP TFs (Gm04bZIP and Gm05bZIP) and their DNA binding activities with *GmCHS8* gene promoter is strong evidence towards the hypothesis since the binding motifs of both GmMYB176 and bZIP TF are present in the critical region that determines *GmCHS8* gene activity. The protein-protein interaction was observed in the nucleus irrespective of the phosphorylation state of GmMYB176 (Fig 3.6). The DNA binding ability of GmMYB176 with the *GmCHS8* promoter is also not dependent on the phosphorylation state of GmMYB176 (Fig 3.7a). Transcriptional complex consisting of an MYB and a bZIP TFs is not as well studied as MBW complexes in plants. In barley, an R1 MYB HvMYBS3 and a bZIP TFs function as transcriptional complex for the regulation of a developing endosperm-specific gene *ltr1* (Rubio-Somoza et al., 2006). In soybean,

GmMYB176 (unphosphorylated) forms a complex with Gm04bZIP/Gm05bZIP, and that this transcriptional complex may regulate *GmCHS8* gene expression (Fig 3.8).

GmMYB176 interacting TFs that were not involved in the regulation of *GmCHS8* gene expression could be regulating some other genes in the pathway or another biological process. GmMYB176 and Gm04bZIP/Gm05bZIP complex could also regulate other genes in the pathway. Silencing of GmMYB176 interacting bZIP TF is required to illustrate its role in *GmCHS8* gene expression and isoflavonoid levels in soybean. This should provide strong, supportive evidence for the role of Gm04bZIP/Gm05bZIP in the GmMYB176 transcription complex for the *GmCHS8* gene regulation in soybean. In addition, the same transcriptional complex could be involved in the regulation of other members in the *GmCHS* gene family or other genes in the isoflavonoid biosynthetic pathway as GmMYB176 was found to be regulating the expression of other *GmCHS* family members (Table 4.3).

GmMYB176 could interact with its cofactor (s) (same or different) to regulate expression of other genes in the isoflavonoid biosynthesis pathway. An MYB TF called anthocyanin pigment 1 (PAP1) co-regulates *PAL*, *CHS*, and *DFR* in anthocyanin biosynthesis in *Arabidopsis* (Borevitz et al., 2000). The structural genes *PAL*, *CHS*, *F3H*, *DFR* that are involved in anthocyanin biosynthesis are believed to be regulated coordinately by a combinatorial action of the bHLH protein-encoding gene *R* and the MYB gene *C1* in maize (Grotewold et al., 1998; Bruce et al., 2000; Kong et al., 2012). Similarly, in soybean, the phenylpropanoid pathway genes *PAL*, *C4H*, *CHS*, *CHI*, and *CHR* were activated by heterologous expression of the maize C1 and R TFs, resulting in an overall increase in total isoflavone levels (Yu et al., 2003). In petunia, co-regulation of structural genes, *F3H*, *F3'5'H*, *DFR*, *ANS*, *3RT*, *5GT*, *GST* in the flavonoid biosynthesis pathway by an MYB TF, MYB27, has been demonstrated and the coordinated expression of these genes regulating the anthocyanin accumulation level was also shown (Albert et al., 2014).

As soybean is a paleopolyploid, with a majority of proteins being encoded by multigene families, their differential expression and their effect on metabolite production should be addressed. Each member of a gene family corresponds to an enzymatic isoform often

specific to certain spatio-temporal dimensions and/or responsive to environmental interactions. The isoflavonoid biosynthetic pathway is a legume-specific branch of phenylpropanoid pathway which produces specialized metabolites. Phenylpropanoid biosynthesis is regulated by coordinated expression of structural genes that are regulated by multiple transcription factors (Vom Endt et al., 2002).Silencing of *GmMYB176* in soybean hairy roots reduced the transcript level of *GmCHS8* and *GmIFS* (Yi et al., 2010). Hence, it was hypothesized that GmMYB176 may regulate multiple genes in the pathway. Here I identified 7129 genes that GmMYB176 might regulates and also demonstrate that the accumulation of several metabolites including isoflavonoids is affected when transcription of *GmMYB176* was altered in soybean hairy roots. The effect that silencing and overexpression of *GmMYB176* have on metabolite production may be due to the changes in transcript level of the structural genes, whose products are involved in the synthesis of the respective metabolites. For example, *GmCHSs*, *GmIFS1*, and *GmPT10a* were differentially expressed in GmMYB176-OE tissues when compared to the controls (Table 4.3; Fig 4.7). Although there was no direct correlation between structural genes' expression and their corresponding metabolite level, the overall effect of altering the transcripts level of *GmMYB176* were associated with the changes in isoflavonoid biosynthetic pathway metabolites such as liquiritigenin (Fig 4.13). Transcripts level may not necessarily correlate with protein level or metabolite level. Pathway regulation is dynamic and has multiple levels of regulations, synchronized to produce specific metabolites at a given time and space. A multi-branched pathway can be controlled at multiple levels, depending on the plants' needs, including the production of phytoalexins during pathogen attack.

Based on the results shown in Fig 4.5, GmMYB176 may also regulate several other processes such as photosynthesis and translation. It is speculated that it may exert its function by partnering with Gm04bZIP/Gm05bZIP as identified in this study or with different co-factors yet to be identified. Silencing of *Gm04bZIP* and/or *Gm05bZIP* in soybean hairy roots will provide more information on whether it affects the expression of genes that are altered by silencing *GmMYB176*. If the transcript levels of DEGs in GmMYB176-Si tissues correlate with DEGs in *Gm04bZIP*/*Gm05bZIP*-silenced tissues,

then the genes could be regulated by the same transcriptional complexes containing GmMYB176 and bZIPs. Co-ordination of multiple gene regulations in a pathway by R2R3 MYB TF has been studied before but not by R1 MYB TF. The role of GmMYB176 and its transcriptional complexes in other biological process is yet to be explored.

5.2 Metabolic engineering of isoflavonoids in plants

Identifying GmMYB176 interacting proteins in soybean and understanding their role in isoflavonoid biosynthesis will allow genetic manipulation of isoflavonoid level in soybean and also will help in metabolic engineering of the isoflavonoid biosynthetic pathway in non-legume plants. The introduction of a TF on its own or in combination with a structural gene, in transgenic plants to improve the accumulation of specialized metabolites, has been studied in plants. An R2R3 MYB TF from *Populus trichocarpa*, PtrMYB119 enhanced the production of anthocyanin in *Arabidopsis* and Poplar (Cho et al., 2016). Metabolic engineering of phenylpropanoids in transgenic plants with a MYB TF, AtPAP1, enhanced the expression of six phenylpropanoid genes (He et al., 2016). Likewise, GmMYB176 will be one of the significant candidates to be considered for metabolic engineering of isoflavonoid biosynthesis in soybean and also other plants such as tomato and lettuce.

The current study will not only help in metabolic engineering of isoflavonoid biosynthetic pathway in non-legume plants but also for the production of isoflavonoid in an engineered cell or ‘synthetic cell’ (Gibson et al., 2010). Synthetic biology is an emerging field that combines science and engineering to construct synthetic cell with desired functions. It holds the promise of a wide range of applications such as the production of desired metabolites, vaccines, and medicines. A flavonoid naringenin was produced *de novo* in engineered *Saccharomyces cerevisiae* (Koopman et al., 2012). Similarly, the knowledge gained in this research on isoflavonoid biosynthetic pathway will be useful in future for designing genetic circuits in a plant or synthetic cell to produce isoflavonoids for therapeutic and health benefits.

5.3 **Literature cited**

Albert, N.W., Davies, K.M., Lewis, D.H., Zhang, H., Montefiori, M., Brendolise, C., Boase, M.R., Ngo, H., Jameson, P.E., and Schwinn, K.E. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. Plant Cell **26,** 962-980.

Ben-Simhon, Z., Judeinstein, S., Nadler-Hassar, T., Trainin, T., Bar-Ya'akov, I., Borochov-Neori, H., and Holland, D. (2011). A pomegranate (*Punica granatum* L.) WD40-repeat gene is a functional homologue of *Arabidopsis* TTG1 and is involved in the regulation of anthocyanin biosynthesis during pomegranate fruit development. Planta **234,** 865-881.

Blanc, G., and Wolfe, K.H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell **16,** 1667-1678.

Boddu, J., Jiang, C., Sangar, V., Olson, T., Peterson, T., and Chopra, S. (2006). Comparative structural and functional characterization of sorghum and maize duplications containing orthologous MYB transcription regulators of 3-deoxyflavonoid biosynthesis. Plant Molecular Biology **60,** 185-199.

Borevitz, J.O., Xia, Y., Blount, J., Dixon, R.A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell **12,** 2383-2394.

Bruce, W., Folkerts, O., Garnaat, C., Crasta, O., Roth, B., and Bowen, B. (2000). Expression profiling of the maize flavonoid pathway genes controlled by estradiol-inducible transcription factors CRC and P. Plant Cell **12,** 65-80.

Cho, J.S., Nguyen, V.P., Jeon, H.W., Kim, M.H., Eom, S.H., Lim, Y.J., Kim, W.C., Park, E.J., Choi, Y.I., and Ko, J.H. (2016). Overexpression of PtrMYB119, a R2R3-MYB transcription factor from *Populus trichocarpa*, promotes anthocyanin production in hybrid poplar. Tree Physiology**36,** 1162-1176.

Dastmalchi, M., and Dhaubhadel, S. (2015). Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. Planta **241,** 507-523.

Dhaubhadel, S., Gijzen, M., Moy, P., and Farhangkhoee, M. (2007). Transcriptome analysis reveals a critical role of *CHS7* and *CHS8* genes for isoflavonoid synthesis in soybean seeds. Plant Physiology **143,** 326-338.

Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C.,

Denisova, E.A., Young, L., Qi, Z.Q., Segall-Shapiro, T.H., Calvey, C.H., Parmar, P.P., Hutchison, C.A., 3rd, Smith, H.O., and Venter, J.C. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. Science **329,** 52-56.

Gill, N., Findley, S., Walling, J.G., Hans, C., Ma, J., Doyle, J., Stacey, G., and Jackson, S.A. (2009). Molecular and chromosomal evidence for allopolyploidy in soybean. Plant Physiology **151,** 1167-1174.

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

Han, Y., Ding, T., Su, B., and Jiang, H. (2016). Genome-wide identification, characterization and expression analysis of the *Chalcone synthase* family in maize. International Journal of Molecular Science**17,** 161.

He, X., Li, Y., Lawson, D., and Xie, D.Y. (2016). Metabolic engineering of anthocyanins in dark tobacco varieties. Physiologica Plantarum**159,** 2-12.

Jin, H., and Martin, C. (1999). Multifunctionality and diversity within the plant MYB-gene family. Plant Molecular Biology **41,** 577-585.

Kong, Q., Pattanaik, S., Feller, A., Werkman, J.R., Chai, C., Wang, Y., Grotewold, E., and Yuan, L. (2012). Regulatory switch enforced by basic helix-loop-helix and ACT-domain mediated dimerizations of the maize transcription factor R. Proceedings of the National Academy of Sciences of the United States of America **109,** E2091-2097.

Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A., Hall, R.D., Bosch, D., van Maris, A.J., Pronk, J.T., and Daran, J.M. (2012). De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. Microbial Cell Factories **11,** 155.

Panchy, N., Lehti-Shiu, M.D., and Shiu, S.-H. (2016). Evolution of gene duplication in plants. Plant Physiology **171,** 2294-2316.

Payne, C.T., Zhang, F., and Lloyd, A.M. (2000). *GL3* encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with GL1 and TTG1. Genetics **156,** 1349-1362.

Rubio-Somoza, I., Martinez, M., Abraham, Z., Diaz, I., and Carbonero, P. (2006). Ternary complex formation between HvMYBS3 and other factors involved in transcriptional control in barley seeds. The Plant Journal **47,** 269-281.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y.,

Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., and Jackson, S.A. (2010). Genome sequence of the palaeopolyploid soybean. Nature **463,** 178-183.

Shangguan, X.X., Yang, C.Q., Zhang, X.F., and Wang, L.J. (2016). Functional characterization of a basic helix-loop-helix (bHLH) transcription factor GhDEL65 from cotton (*Gossypium hirsutum*). Physiologica Plantarum**158,** 200-212.

Shoemaker, R.C., Schlueter, J., and Doyle, J.J. (2006). Paleopolyploidy and gene duplication in soybean and other legumes. Current Opinion in Plant Biology**9,** 104-109.

Shoemaker, R.C., Polzin, K., Labate, J., Specht, J., Brummer, E.C., Olson, T., Young, N., Concibido, V., Wilcox, J., Tamulonis, J.P., Kochert, G., and Boerma, H.R. (1996). Genome duplication in soybean (*Glycine* subgenus *soja*). Genetics **144,** 329-338.

Vom Endt, D., Kijne, J.W., and Memelink, J. (2002). Transcription factors controlling plant secondary metabolism: what regulates the regulators? Phytochemistry **61,** 107-114.

Wang, S., and Chen, J.G. (2008). *Arabidopsis* transient expression analysis reveals that activation of GLABRA2 may require concurrent binding of GLABRA1 and GLABRA3 to the promoter of GLABRA2. Plant and Cell Physiology **49,** 1792-1804.

Yi, J., Derynck, M.R., Li, X., Telmer, P., Marsolais, F., and Dhaubhadel, S. (2010). A single-repeat MYB transcription factor, GmMYB176, regulates *CHS8* gene expression and affects isoflavonoid biosynthesis in soybean. The Plant Journal **62,** 1019-1034.

Yu, O., Shi, J., Hession, A.O., Maxwell, C.A., McGonigle, B., and Odell, J.T. (2003). Metabolic engineering to increase isoflavone biosynthesis in soybean seed. Phytochemistry **63,** 753-763.

Appendices

Appendix 1 The list of 716 candidates identified in Co-IP using GmMYB176-YFP and YFP-GmMYB176 as baits.

GO annotation 'Biological process' information was retrieved from Phytomine using candidates protein IDs. 'w/ both the tags' - Candidates identified with both the baits GmMYB176-YFP and YFP-GmMYB176. 'NoP' - number of unique peptides matched for a protein.

Bait	Glyma. ID	Mascot score	Mass (KDa)	NoP	Phytozome annotation	Biological process
GmMYB176-YFP only	Glyma.07G110800.1.p	24	24.2	1	No functional domain available	biosynthetic process
GmMYB176-YFP only	Glyma.12G163700.1.p	25	122.4	1	methionine S-methyltransferase	biosynthetic process
GmMYB176-YFP only	Glyma.05G026300.1.p	23	43.9	1	malate dehydrogenase	carbohydrate metabolic process
GmMYB176-YFP only	Glyma.12G159300.1.p	685	36.3	22	Lactate/malate dehydrogenase family protein	carbohydrate metabolic process
GmMYB176-YFP only	Glyma.01G137700.1.p	22	57	1	Plant invertase/pectin methylesterase inhibitor superfamily	cell wall modification
GmMYB176-YFP only	Glyma.16G038300.2.p	35	78.2	3	Cobalamin-independent synthase family protein	cellular amino acid biosynthetic process
GmMYB176-YFP only	Glyma.08G249900.1.p	32	40.7	1	reversibly glycosylated polypeptide 2	cellulose biosynthetic process
GmMYB176-YFP only	Glyma.12G237000.1.p	24	121.8	1	Cellulose synthase family protein	cellulose biosynthetic process
GmMYB176-YFP only	Glyma.05G165600.1.p	24	114.8	1	NB-ARC domain-containing disease resistance protein	defense response
GmMYB176-YFP only	Glyma.16G147400.1.p	28	84.2	1	Disease resistance protein (TIR-NBS-LRR class) family	defense response
GmMYB176-YFP only	Glyma.20G042700.1.p	38	107.4	2	NB-ARC domain-containing disease resistance protein	defense response
GmMYB176-YFP only	Glyma.03G144600.1.p	39	58.8	2	DNA polymerase delta small subunit	DNA replication

GmMYB176-YFP only	Glyma.14G140800.1.p	23	239.1	1	helicases;ATP-dependenthelicases;nucleic acid	DNA replication
GmMYB176-YFP only	Glyma.06G155600.1.p	29	105.1	1	topoisomerase 3alpha	DNA topological change
GmMYB176-YFP only	Glyma.10G037400.1.p	23	41.8	1	Rho termination factor	DNA-dependent transcription, termination
GmMYB176-YFP only	Glyma.06G143000.1.p	226	23.5	12	Chalcone-flavanone isomerase family protein	flavonoid biosynthetic process
GmMYB176-YFP only	Glyma.20G241500.1.p	63	23.3	3	Chalcone-flavanone isomerase family protein	flavonoid biosynthetic process
GmMYB176-YFP only	Glyma.09G270500.1.p	51	61.1	2	Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	glucose catabolic process
GmMYB176-YFP only	Glyma.02G074200.1.p	45	54.3	1	glutathione-disulfide reductase	glutathione metabolic process
GmMYB176-YFP only	Glyma.04G247900.1.p	40	59	2	PDI-like 1-2	glycerol ether metabolic process
GmMYB176-YFP only	Glyma.05G178300.1.p	26	23.3	1	RAB GTPase homolog H1E	GTP catabolic process
GmMYB176-YFP only	Glyma.07G039600.1.p	30	79.9	2	RELA/SPOT homolog 3	guanosine tetraphosphate metabolic process
GmMYB176-YFP only	Glyma.05G038100.1.p	22	17.7	1	SufE/NifU family protein	iron-sulfur cluster assembly
GmMYB176-YFP only	Glyma.07G056800.1.p	24	75.6	1	alpha/beta-Hydrolases superfamily protein	lipid metabolic process
GmMYB176-YFP only	Glyma.09G158100.1.p	33	26.9	1	Translation initiation factor IF6	mature ribosome assembly
GmMYB176-YFP only	Glyma.04G109000.1.p	23	55.1	1	UDP-glucosyl transferase 85A7	metabolic process
GmMYB176-YFP only	Glyma.05G231800.1.p	54	54.7	2	aldehyde dehydrogenase 2C4	metabolic process
GmMYB176-YFP only	Glyma.06G118500.1.p	32	42.2	1	Chalcone and stilbene synthase family protein	metabolic process
GmMYB176-YFP only	Glyma.07G193400.1.p	27	63.5	1	TYROSINASE	metabolic process
GmMYB176-YFP only	Glyma.07G239700.1.p	195	66.6	7	ATP citrate lyase subunit B 2	metabolic process

GmMYB176-YFP only	Glyma.07G261500.1.p	23	52.2	1	APS reductase 3	metabolic process
GmMYB176-YFP only	Glyma.13G243400.1.p	116	25.8	6	3-hydroxyacyl-CoA dehydratase 1	metabolic process
GmMYB176-YFP only	Glyma.15G040100.1.p	23	48.9	1	ACT domain-containing protein	metabolic process
GmMYB176-YFP only	Glyma.17G256100.1.p	23	109.2	1	isopropylmalateisomerase large subunit 1	metabolic process
GmMYB176-YFP only	Glyma.19G035800.1.p	25	54.9	2	UDP-glucosyl transferase 85A7	metabolic process
GmMYB176-YFP only	Glyma.05G208100.1.p	33	121.6	1	P-loop nucleoside triphosphate hydrolases superfamily protein with CH	microtubule-based movement
GmMYB176-YFP only	Glyma.05G233900.1.p	22	140.7	1	P-loop containing nucleoside triphosphate hydrolases superfamily	microtubule-based movement
GmMYB176-YFP only	Glyma.09G194900.1.p	40	144	1	kinesin-like calmodulin-binding protein (ZWICHEL)	microtubule-based movement
GmMYB176-YFP only	Glyma.02G224100.1.p	47	11.4	2	Histone superfamily protein	nucleosome assembly
GmMYB176-YFP only	Glyma.08G189900.1.p	23	30.8	1	Homeodomain-like/winged-helix DNA-binding family protein	nucleosome assembly
GmMYB176-YFP only	Glyma.01G102600.1.p	22	35	1	NAD(P)-linked oxidoreductase superfamily protein	oxidation-reduction process
GmMYB176-YFP only	Glyma.05G056100.1.p	22	67.2	1	Cupredoxin superfamily protein	oxidation-reduction process
GmMYB176-YFP only	Glyma.08G058000.1.p	39	61.5	2	alternative NAD(P)H dehydrogenase 2	oxidation-reduction process
GmMYB176-YFP only	Glyma.08G364300.1.p	23	39.8	1	Flavodoxin family protein	oxidation-reduction process
GmMYB176-YFP only	Glyma.10G092500.1.p	22	58.5	1	cytochrome P450, family 93, subfamily D, polypeptide 1	oxidation-reduction process
GmMYB176-YFP only	Glyma.14G005700.1.p	43	34	2	NAD(P)-linked oxidoreductase superfamily protein	oxidation-reduction process
GmMYB176-YFP only	Glyma.16G073600.1.p	230	34.6	6	Oxidoreductase, zinc-binding dehydrogenase family protein	oxidation-reduction process
GmMYB176-YFP only	Glyma.18G285800.1.p	23	35.3	1	NAD(P)-linked oxidoreductase superfamily protein	oxidation-reduction process

GmMYB176-YFP only	Glyma.03G223900.1.p	48	42.3	4	Ca2+ activated outward rectifying K+ channel 5	potassium ion transmembrane transport
GmMYB176-YFP only	Glyma.06G074600.1.p	33	68.3	1	potassium channel in Arabidopsis thaliana 3	potassium ion transport
GmMYB176-YFP only	Glyma.13G042100.1.p	28	39.5	1	Protein phosphatase 2C family protein	protein dephosphorylation
GmMYB176-YFP only	Glyma.02G305600.1.p	34	88	2	HEAT SHOCK PROTEIN 89.1	protein folding
GmMYB176-YFP only	Glyma.09G155200.1.p	31	76.6	1	ATP-dependent Clp protease	protein folding
GmMYB176-YFP only	Glyma.10G147600.1.p	26	50.1	1	calreticulin 1b	protein folding
GmMYB176-YFP only	Glyma.01G099600.1.p	24	81.2	1	Protein kinase superfamily protein	protein phosphorylation
GmMYB176-YFP only	Glyma.02G186300.1.p	22	40.1	1	MAP kinase kinase 6	protein phosphorylation
GmMYB176-YFP only	Glyma.08G174500.1.p	24	124.5	1	Leucine-rich receptor-like protein kinase family protein	protein phosphorylation
GmMYB176-YFP only	Glyma.08G240700.1.p	30	76.2	1	CDC2C	protein phosphorylation
GmMYB176-YFP only	Glyma.11G214400.1.p	25	99	1	transmembrane kinase 1	protein phosphorylation
GmMYB176-YFP only	Glyma.14G107600.1.p	23	77.4	1	protein kinase family protein	protein phosphorylation
GmMYB176-YFP only	Glyma.03G257900.1.p	35	50.8	2	Eukaryotic aspartyl protease family protein	proteolysis
GmMYB176-YFP only	Glyma.04G194100.1.p	30	53.7	1	Zn-dependent exopeptidases superfamily protein	proteolysis
GmMYB176-YFP only	Glyma.15G123800.1.p	24	50.8	1	Eukaryotic aspartyl protease family protein	proteolysis
GmMYB176-YFP only	Glyma.06G116100.1.p	24	54.4	1	Cyclin A1;1	regulation of cyclin-dependent protein serine/threonine kinase
GmMYB176-YFP only	Glyma.04G078600.1.p	27	40	1	NAC domain containing protein 38	regulation of transcription, DNA-dependent
GmMYB176-YFP only	Glyma.04G085000.1.p	42	93.9	2	homeobox gene 8	regulation of transcription, DNA-dependent

GmMYB176-YFP only	Glyma.05G122400.1.p	24	24.7	1	Basic-leucine zipper (bZIP) transcription factor family protein	regulation of transcription, DNA-dependent
GmMYB176-YFP only	Glyma.06G148400.1.p	25	35	1	Integrase-type DNA-binding superfamily protein	regulation of transcription, DNA-dependent
GmMYB176-YFP only	Glyma.06G248900.1.p	24	38.2	1	NAC (No Apical Meristem) domain transcriptional regulator superfamily	regulation of transcription, DNA-dependent
GmMYB176-YFP only	Glyma.08G018300.1.p	23	33.3	1	WRKY family transcription factor	regulation of transcription, DNA-dependent
GmMYB176-YFP only	Glyma.13G100500.1.p	23	80.2	1	pathogenesis related homeodomain protein A	regulation of transcription, DNA-dependent
GmMYB176-YFP only	Glyma.15G079200.1.p	28	29.9	1	Integrase-type DNA-binding superfamily protein	regulation of transcription, DNA-dependent
GmMYB176-YFP only	Glyma.08G051700.1.p	22	71	1	alpha dioxygenase	response to oxidative stress
GmMYB176-YFP only	Glyma.02G002800.1.p	100	29.7	4	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein	ribosome biogenesis
GmMYB176-YFP only	Glyma.16G036100.1.p	25	119.5	1	polyribonucleotide nucleotidyltransferase, putative	RNA processing
GmMYB176-YFP only	Glyma.06G007300.1.p	25	24.4	1	ribosome-binding factor A family protein	rRNA processing
GmMYB176-YFP only	Glyma.14G128900.1.p	30	56.7	1	Protein phosphatase 2A regulatory B subunit family protein	signal transduction
GmMYB176-YFP only	Glyma.09G080200.1.p	35	167.5	1	PIF1 helicase	telomere maintenance
GmMYB176-YFP only	Glyma.18G210300.1.p	23	151.6	1	HEAT repeat ;WD domain, G-beta repeat protein protein	TOR signaling cascade
GmMYB176-YFP only	Glyma.11G026700.1.p	24	165.4	1	nuclear RNA polymerase D1A	transcription, DNA-dependent
GmMYB176-YFP only	Glyma.01G029200.1.p	23	17.9	1	Ribosomal protein S27a / Ubiquitin family protein	translation
GmMYB176-YFP only	Glyma.01G124300.1.p	27	30.1	1	Ribosomal protein S4 (RPS4A) family protein	translation
GmMYB176-YFP only	Glyma.01G236600.1.p	120	29.8	4	Ribosomal protein S3Ae	translation
GmMYB176-YFP only	Glyma.02G085300.1.p	188	30.4	10	Ribosomal protein S5 family protein	translation

GmMYB176-YFP only	Glyma.02G167900.1.p	25	22.9	1	ribosomal protein 5B	translation
GmMYB176-YFP only	Glyma.02G191200.1.p	81	20.6	4	ribosomal protein S6	translation
GmMYB176-YFP only	Glyma.02G239700.1.p	38	15.3	2	Ribosomal protein S26e family protein	translation
GmMYB176-YFP only	Glyma.03G090700.1.p	34	125.6	2	ribosomal protein S6	translation
GmMYB176-YFP only	Glyma.03G142200.1.p	38	58.2	1	Ribosomal protein S10p/S20e family protein	translation
GmMYB176-YFP only	Glyma.03G207500.1.p	24	18.7	1	Translation protein SH3-like family protein	translation
GmMYB176-YFP only	Glyma.04G171400.1.p	24	23.1	1	Ribosomal protein S4	translation
GmMYB176-YFP only	Glyma.09G000700.1.p	72	44.8	5	ribosomal protein 1	translation
GmMYB176-YFP only	Glyma.09G024000.1.p	25	23.8	1	Ribosomal protein L13 family protein	translation
GmMYB176-YFP only	Glyma.U009000.1.p	24	10.9	1	ribosomal protein L23.1	translation
GmMYB176-YFP only	Glyma.07G273500.1.p	28	48.1	1	Translation elongation factor EF1B, gamma chain	translational elongation
GmMYB176-YFP only	Glyma.01G200100.1.p	23	67	1	nitrate transporter 1.5	Transport
GmMYB176-YFP only	Glyma.12G172500.1.p	275	31.1	12	plasma membrane intrinsic protein 2	transport
GmMYB176-YFP only	Glyma.14G061500.1.p	33	30.8	1	plasma membrane intrinsic protein 1;4	transport
GmMYB176-YFP only	Glyma.17G259100.1.p	23	136.1	1	P-glycoprotein 13	transport
GmMYB176-YFP only	Glyma.20G090100.1.p	22	83.6	1	Glutamyl/glutaminyl-tRNAsynthetase, class Ic	tRNAaminoacylation for protein translation
GmMYB176-YFP only	Glyma.01G124900.1.p	465	27.5	15	20S proteasome alpha subunit C1	ubiquitin-dependent protein catabolic process
GmMYB176-YFP only	Glyma.02G210300.1.p	27	24.1	1	No functional domain available	ubiquitin-dependent protein catabolic process

GmMYB176-YFP only	Glyma.06G254200.1.p	43	27.5	2	proteasome alpha subunit A1	ubiquitin-dependent protein catabolic process
GmMYB176-YFP only	Glyma.07G155700.1.p	328	27.5	13	20S proteasome alpha subunit C1	ubiquitin-dependent protein catabolic process
GmMYB176-YFP only	Glyma.08G167600.1.p	182	25.6	8	proteasome subunit PAB1	ubiquitin-dependent protein catabolic process
GmMYB176-YFP only	Glyma.10G279900.1.p	236	26.2	9	20S proteasome alpha subunit E2	ubiquitin-dependent protein catabolic process
GmMYB176-YFP only	Glyma.16G126200.2.p	25	22.1	1	proteasome alpha subunit D2	ubiquitin-dependent protein catabolic process
GmMYB176-YFP only	Glyma.19G072100.1.p	121	27.6	4	20S proteasome alpha subunit G1	ubiquitin-dependent protein catabolic process
GmMYB176-YFP only	Glyma.01G024900.1.p	59	20.9	3	No functional domain available	
GmMYB176-YFP only	Glyma.01G026500.1.p	22	36.6	1	F-box family protein	
GmMYB176-YFP only	Glyma.01G029900.1.p	30	60.4	1	No functional domain available	
GmMYB176-YFP only	Glyma.01G044300.1.p	25	53.9	1	myb domain protein 88	
GmMYB176-YFP only	Glyma.01G057600.1.p	24	58	1	Pentatricopeptide repeat (PPR) superfamily protein	
GmMYB176-YFP only	Glyma.01G106000.1.p	85	25.1	2	glutathione S-transferase TAU 8	
GmMYB176-YFP only	Glyma.01G114500.1.p	31	39.6	1	glucose-6-phosphate/phosphate translocator 2	
GmMYB176-YFP only	Glyma.01G129300.1.p	24	22.2	1	Polynucleotidyl transferase, ribonuclease H-like superfamily protein	
GmMYB176-YFP only	Glyma.01G210600.1.p	29	89.2	1	OSBP(oxysterol binding protein)-related protein 2A	
GmMYB176-YFP only	Glyma.01G212500.1.p	23	28.6	1	No functional domain available	
GmMYB176-YFP only	Glyma.02G008600.1.p	31	113.5	1	squamosa promoter binding protein-like 1	
GmMYB176-YFP only	Glyma.02G012900.1.p	23	194.2	1	WD40/YVTN repeat-like-containing domain;Bromodomain	

GmMYB176-YFP only	Glyma.02G078100.1.p	29	69.4	1	RH39
GmMYB176-YFP only	Glyma.02G118900.1.p	24	43	1	Nucleic acid-binding, OB-fold-like protein
GmMYB176-YFP only	Glyma.02G160100.1.p	23	82.4	1	Enhancer of polycomb-like transcription factor protein
GmMYB176-YFP only	Glyma.02G162500.1.p	22	40.9	1	Single-stranded nucleic acid binding R3H protein
GmMYB176-YFP only	Glyma.02G191400.1.p	22	54.2	1	BED zinc finger ;hAT family dimerisation domain
GmMYB176-YFP only	Glyma.02G294200.1.p	30	39	1	EPS15 homology domain 1
GmMYB176-YFP only	Glyma.02G299400.1.p	24	24.3	1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
GmMYB176-YFP only	Glyma.03G009900.1.p	31	75.4	1	SKP1 interacting partner 4
GmMYB176-YFP only	Glyma.03G015700.1.p	26	200.2	1	RAP
GmMYB176-YFP only	Glyma.03G041600.1.p	111	24.4	5	SCAR homolog 2
GmMYB176-YFP only	Glyma.03G190000.1.p	23	65.3	1	Protein of unknown function (DUF668)
GmMYB176-YFP only	Glyma.03G201700.1.p	34	22.2	2	No functional domain available
GmMYB176-YFP only	Glyma.03G253000.1.p	22	145.9	1	transcription activators
GmMYB176-YFP only	Glyma.04G005000.1.p	33	70.6	1	armadillo repeat only 1
GmMYB176-YFP only	Glyma.04G018300.1.p	27	66.1	2	O-fucosyltransferase family protein
GmMYB176-YFP only	Glyma.04G032700.1.p	24	197.6	1	GYF domain-containing protein
GmMYB176-YFP only	Glyma.04G048800.1.p	26	85.7	1	lipases;hydrolases, acting on ester bonds
GmMYB176-YFP only	Glyma.04G069800.1.p	27	164.7	1	pleiotropic drug resistance 12

GmMYB176-YFP only	Glyma.04G107400.2.p	23	54.8	1	translocase inner membrane subunit 44-2
GmMYB176-YFP only	Glyma.04G114100.1.p	25	83.9	1	P-loop containing nucleoside triphosphate hydrolases superfamily
GmMYB176-YFP only	Glyma.04G225000.1.p	25	120	1	No functional domain available
GmMYB176-YFP only	Glyma.05G062600.1.p	25	13.4	1	No functional domain available
GmMYB176-YFP only	Glyma.05G102300.1.p	23	72.4	1	ARM-repeat/Tetratricopeptide repeat (TPR)-like protein
GmMYB176-YFP only	Glyma.05G123700.1.p	23	37.6	1	polygalacturonase inhibiting protein 1
GmMYB176-YFP only	Glyma.06G017200.1.p	48	7.7	4	No functional domain available
GmMYB176-YFP only	Glyma.06G098100.1.p	31	44	1	F-box/RNI-like superfamily protein
GmMYB176-YFP only	Glyma.06G123100.1.p	22	55.8	1	Protein of unknown function (DUF300)
GmMYB176-YFP only	Glyma.06G126700.1.p	31	40.4	3	spermidine synthase 3
GmMYB176-YFP only	Glyma.06G304200.1.p	22	58.5	1	O-fucosyltransferase family protein
GmMYB176-YFP only	Glyma.07G031500.1.p	24	44.4	1	O-fucosyltransferase family protein
GmMYB176-YFP only	Glyma.07G032500.1.p	23	69	1	WD40 repeat-containing protein
GmMYB176-YFP only	Glyma.07G061300.1.p	24	74.6	1	No functional domain available
GmMYB176-YFP only	Glyma.07G140400.1.p	143	25.9	7	glutathione S-transferase TAU 8
GmMYB176-YFP only	Glyma.07G158200.1.p	23	59.5	1	indeterminate(ID)-domain 4
GmMYB176-YFP only	Glyma.07G227800.1.p	24	40.1	1	No functional domain available
GmMYB176-YFP only	Glyma.07G231500.1.p	37	34.1	1	No functional domain available

GmMYB176-YFP only	Glyma.07G264300.1.p	40	85.1	1	Octicosapeptide/Phox/Bem1p (PB1) domain-containing protein /
GmMYB176-YFP only	Glyma.08G006100.7.p	25	49.2	1	No functional domain available
GmMYB176-YFP only	Glyma.08G040300.1.p	23	48.2	1	actin-related protein 4
GmMYB176-YFP only	Glyma.08G041700.1.p	30	83	2	CRS1 / YhbY (CRM) domain-containing protein
GmMYB176-YFP only	Glyma.08G072900.1.p	34	20.2	2	Pentatricopeptide repeat (PPR) superfamily protein
GmMYB176-YFP only	Glyma.08G173500.1.p	29	248.3	1	U5 small nuclear ribonucleoprotein helicase, putative
GmMYB176-YFP only	Glyma.08G291300.1.p	23	39.6	1	RNase H family protein
GmMYB176-YFP only	Glyma.08G344300.1.p	25	11.4	2	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin
GmMYB176-YFP only	Glyma.09G067400.1.p	24	81.9	1	Aminotransferase-like, plant mobile domain family protein
GmMYB176-YFP only	Glyma.09G089600.1.p	24	36.2	1	Mitochondrial transcription termination factor family protein
GmMYB176-YFP only	Glyma.09G091000.1.p	24	17.3	1	LYR family of Fe/S cluster biogenesis protein
GmMYB176-YFP only	Glyma.09G166600.1.p	23	44.2	1	Nucleotide-diphospho-sugar transferase family protein
GmMYB176-YFP only	Glyma.09G167100.1.p	24	117.9	1	Stabilizer of iron transporter SufD / Polynucleotidyl transferase
GmMYB176-YFP only	Glyma.09G214700.1.p	24	56.2	1	Ribonuclease P protein subunit P38-related
GmMYB176-YFP only	Glyma.09G267700.1.p	29	42.5	1	FASCICLIN-like arabinogalactan-protein 10
GmMYB176-YFP only	Glyma.10G070800.1.p	28	6	1	No functional domain available
GmMYB176-YFP only	Glyma.10G106000.1.p	23	44.6	1	IQ-domain 12
GmMYB176-YFP only	Glyma.10G149600.1.p	24	46.5	1	Protein phosphatase 2C family protein

GmMYB176-YFP only	Glyma.10G158400.1.p	25	60.7	1	vamp/synaptobrevin-associated protein 27-2
GmMYB176-YFP only	Glyma.10G266500.1.p	34	55.6	2	Plant protein of unknown function (DUF247)
GmMYB176-YFP only	Glyma.11G008600.1.p	30	164.9	2	RING/FYVE/PHD zinc finger superfamily protein
GmMYB176-YFP only	Glyma.11G017400.1.p	31	12.1	1	No functional domain available
GmMYB176-YFP only	Glyma.11G039900.1.p	39	488.6	1	pleckstrin homology (PH) domain-containing protein
GmMYB176-YFP only	Glyma.11G071000.1.p	23	157.7	1	auxin-like 1 protein
GmMYB176-YFP only	Glyma.11G071100.1.p	33	13.3	1	No functional domain available
GmMYB176-YFP only	Glyma.11G083100.1.p	28	32.7	2	Arabidopsis protein of unknown function (DUF241)
GmMYB176-YFP only	Glyma.11G087400.1.p	24	17.6	1	No functional domain available
GmMYB176-YFP only	Glyma.11G090500.1.p	23	14.5	1	SWIB/MDM2 domain superfamily protein
GmMYB176-YFP only	Glyma.11G096400.1.p	23	153.4	1	E3 ubiquitin ligase involved in syntaxin degradation
GmMYB176-YFP only	Glyma.11G110600.1.p	34	101.8	4	pentatricopeptide (PPR) repeat-containing protein
GmMYB176-YFP only	Glyma.11G114300.1.p	30	61	1	NOP56-like pre RNA processing ribonucleoprotein
GmMYB176-YFP only	Glyma.11G190900.1.p	30	100.4	1	Argonaute family protein
GmMYB176-YFP only	Glyma.11G238700.1.p	24	21.3	1	Protein of unknown function (DUF581)
GmMYB176-YFP only	Glyma.11G244500.1.p	32	93.4	2	ARM repeat superfamily protein
GmMYB176-YFP only	Glyma.12G032400.3.p	24	49	1	Calcium-dependent lipid-binding (CaLB domain) family protein
GmMYB176-YFP only	Glyma.12G046100.1.p	25	46.1	1	NAD(P)-binding Rossmann-fold superfamily protein

GmMYB176-YFP only	Glyma.12G192200.2.p	38	26.4	1	alpha/beta-Hydrolases superfamily protein
GmMYB176-YFP only	Glyma.12G206000.1.p	31	44.8	1	Reticulon family protein
GmMYB176-YFP only	Glyma.12G212800.1.p	30	37.7	1	Homeobox-leucine zipper family protein / lipid-binding START domain-
GmMYB176-YFP only	Glyma.12G222700.1.p	23	64.4	1	Leucine-rich repeat (LRR) family protein
GmMYB176-YFP only	Glyma.12G228700.1.p	23	148.5	1	Acyl-CoA N-acyltransferase with RING/FYVE/PHD-type zinc finger
GmMYB176-YFP only	Glyma.12G234700.1.p	26	21.7	1	Kunitz family trypsin and protease inhibitor protein
GmMYB176-YFP only	Glyma.13G047300.1.p	27	27.9	1	allene oxide cyclase 3
GmMYB176-YFP only	Glyma.13G071600.1.p	28	46	1	RING/U-box superfamily protein
GmMYB176-YFP only	Glyma.13G075300.1.p	24	104.9	1	RNA recognition motif (RRM)-containing protein
GmMYB176-YFP only	Glyma.13G160800.1.p	22	81.9	1	MAK10 homologue
GmMYB176-YFP only	Glyma.13G368500.1.p	30	33.3	1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
GmMYB176-YFP only	Glyma.14G031000.1.p	80	24.9	4	glutathione S-transferase PHI 9
GmMYB176-YFP only	Glyma.14G080200.1.p	26	32.7	1	Protein of unknown function (DUF1645)
GmMYB176-YFP only	Glyma.14G212800.1.p	24	51.4	1	NAD(P)-binding Rossmann-fold superfamily protein
GmMYB176-YFP only	Glyma.14G214400.1.p	22	73.8	1	No functional domain available
GmMYB176-YFP only	Glyma.15G076800.1.p	23	100.4	1	Zinc finger C-x8-C-x5-C-x3-H type family protein
GmMYB176-YFP only	Glyma.15G083500.1.p	24	74.8	1	Protein of unknown function (DUF688)
GmMYB176-YFP only	Glyma.15G092600.1.p	27	42.3	1	F-box family protein

GmMYB176-YFP only	Glyma.15G122400.1.p	31	153.1	1	varicose-related
GmMYB176-YFP only	Glyma.15G183400.1.p	22	77.3	1	DnaJ / Sec63 Brl domains-containing protein
GmMYB176-YFP only	Glyma.15G248300.1.p	22	26.9	1	SOUL heme-binding family protein
GmMYB176-YFP only	Glyma.16G019700.1.p	32	61.7	1	IQ-domain 17
GmMYB176-YFP only	Glyma.16G111200.1.p	39	144.2	1	kinesin-like calmodulin-binding protein (ZWICHEL)
GmMYB176-YFP only	Glyma.16G112900.1.p	32	133.5	1	FOIE GRAS
GmMYB176-YFP only	Glyma.17G002000.1.p	30	136.1	1	cullin-associated and neddylation dissociated
GmMYB176-YFP only	Glyma.17G029800.1.p	33	69.4	1	transducin family protein / WD-40 repeat family protein
GmMYB176-YFP only	Glyma.17G081100.1.p	26	53.2	1	nonsense-mediated mRNA decay NMD3 family protein
GmMYB176-YFP only	Glyma.17G094200.1.p	28	59.2	1	No functional domain available
GmMYB176-YFP only	Glyma.17G161500.1.p	25	70.4	1	Phototropic-responsive NPH3 family protein
GmMYB176-YFP only	Glyma.17G254300.1.p	22	27.6	1	LisH and RanBPM domains containing protein
GmMYB176-YFP only	Glyma.18G133600.1.p	23	9	1	No functional domain available
GmMYB176-YFP only	Glyma.18G190500.1.p	139	25.7	6	glutathione S-transferase TAU 8
GmMYB176-YFP only	Glyma.18G262600.1.p	33	81.9	1	hypothetical protein 1
GmMYB176-YFP only	Glyma.18G276300.1.p	32	81.1	1	SMAD/FHA domain-containing protein
GmMYB176-YFP only	Glyma.18G279800.1.p	23	51.9	1	TRICHOME BIREFRINGENCE-LIKE 19
GmMYB176-YFP only	Glyma.19G113400.1.p	25	48.8	1	No functional domain available

GmMYB176-YFP only	Glyma.19G136600.2.p	28	10.3	1	No functional domain available	
GmMYB176-YFP only	Glyma.19G137700.1.p	27	126.9	1	ABC-2 type transporter family protein	
GmMYB176-YFP only	Glyma.19G204200.1.p	24	154.8	1	Cleavage and polyadenylation specificity factor (CPSF) A subunit	
GmMYB176-YFP only	Glyma.19G234600.1.p	34	53.1	1	homolog of yeast sucrose nonfermenting 4	
GmMYB176-YFP only	Glyma.20G047800.1.p	28	97	1	RINT-1 / TIP-1 family	
GmMYB176-YFP only	Glyma.20G219700.1.p	27	37.6	1	R3H domain	
GmMYB176-YFP only	Glyma.20G240300.1.p	299	23.5	10	dehydroascorbate reductase 2	
GmMYB176-YFP only	Glyma.20G243000.1.p	27	110	1	No functional domain available	
GmMYB176-YFP[2]	Glyma.06G061500.1.p	33	58.9	1	PHOSPHOLIPASE D - RELATED	metabolic process
GmMYB176-YFP[2]	Glyma.03G143700.1.p	44	48.9	2	cytochrome P450, family 93, subfamily D, polypeptide 1	oxidation-reduction process
GmMYB176-YFP[2]	Glyma.04G164400.1.p	32	32.1	1	soluble N-ethylmaleimide-sensitive factor adaptor protein 33	
GmMYB176-YFP[2]	Glyma.09G048200.1.p	35	105.7	2	INHIBITOR OF APOPTOSIS 1, DIAP1	
GmMYB176-YFP[2]	Glyma.10G072900.1.p	38	46.9	1	UNCHARACTERIZED	
GmMYB176-YFP[2]	Glyma.19G141900.1.p	32	133.8	1	ARM repeat superfamily protein	
w/ both the tags	Glyma.07G255000.1.p	26	26.5	1	Copper amine oxidase family protein	amine metabolic process
w/ both the tags	Glyma.06G231500.1.p	313	36.2	13	Lactate/malate dehydrogenase family protein	carbohydrate metabolic process
w/ both the tags	Glyma.07G103000.1.p	35	149	2	ketose-bisphosphate aldolase class-II family protein	carbohydrate metabolic process
w/ both the tags	Glyma.15G142400.1.p	94	41	4	Glycosyl hydrolase superfamily protein	carbohydrate metabolic process

w/ both the tags	Glyma.03G086000.1.p	25	123.3	1	heavy metal atpase 5	cation transport
w/ both the tags	Glyma.13G186000.1.p	35	82.2	3	Plant invertase/pectin methylesterase inhibitor superfamily	cell wall modification
w/ both the tags	Glyma.16G041200.1.p	319	44.8	10	glutamate dehydrogenase 1	cellular amino acid metabolic process
w/ both the tags	Glyma.12G019800.1.p	102	38.1	4	NAD(P)-binding Rossmann-fold superfamily protein	cellular metabolic process
w/ both the tags	Glyma.04G066400.1.p	36	71.6	3	epsin N-terminal homology (ENTH) domain-containing protein / clathrin	clathrin coat assembly
w/ both the tags	Glyma.01G039000.1.p	62	158.8	4	transmembrane receptors;ATP binding	defense response
w/ both the tags	Glyma.16G211400.1.p	55	148.3	5	Disease resistance protein (TIR-NBS-LRR class) family	defense response
w/ both the tags	Glyma.10G268500.1.p	105	42.8	5	Aldolase superfamily protein	glycolysis
w/ both the tags	Glyma.01G204900.1.p	28	121.4	1	SU(VAR)3-9 homolog 6	histone lysine methylation
w/ both the tags	Glyma.05G237400.1.p	87	109.4	12	delta-adaptin	intracellular protein transport
w/ both the tags	Glyma.19G144800.1.p	23	39.1	1	geranylgeranyl pyrophosphate synthase 1	isoprenoid biosynthetic process
w/ both the tags	Glyma.10G130800.1.p	24	97.3	1	triglyceride lipases;triglyceride lipases	lipid metabolic process
w/ both the tags	Glyma.01G008500.1.p	285	65.3	10	NADP-malic enzyme 4	malate metabolic process
w/ both the tags	Glyma.05G033500.1.p	206	53.5	4	aldehyde dehydrogenase 10A8	metabolic process
w/ both the tags	Glyma.06G186300.1.p	544	55.4	15	aldehyde dehydrogenase 10A8	metabolic process
w/ both the tags	Glyma.08G302600.1.p	33	44.5	1	alanine:glyoxylate aminotransferase	metabolic process
w/ both the tags	Glyma.09G010000.1.p	49	46.1	3	fumarylacetoacetase, putative	metabolic process
w/ both the tags	Glyma.10G155300.1.p	27	32.2	1	NAD(P)-binding Rossmann-fold superfamily protein	metabolic process

w/ both the tags	Glyma.13G336200.1.p	879	24.8	23	triosephosphate isomerase	metabolic process
w/ both the tags	Glyma.15G038100.1.p	1057	27.4	31	triosephosphate isomerase	metabolic process
w/ both the tags	Glyma.05G152000.1.p	748	55.7	27	S-adenosyl-L-homocysteine hydrolase	one-carbon metabolic process
w/ both the tags	Glyma.03G260600.1.p	92	66.4	4	D-arabinono-1,4-lactone oxidase family protein	oxidation-reduction process
w/ both the tags	Glyma.07G007000.1.p	611	96.9	16	lipoxygenase 1	oxidation-reduction process
w/ both the tags	Glyma.08G080200.1.p	28	60.2	1	FAD-binding Berberine family protein	oxidation-reduction process
w/ both the tags	Glyma.10G067500.1.p	104	46.9	5	monodehydroascorbate reductase 1	oxidation-reduction process
w/ both the tags	Glyma.11G062800.1.p	268	43.9	41	cytochrome P450, family 71, subfamily B, polypeptide 37	oxidation-reduction process
w/ both the tags	Glyma.08G332900.1.p	423	80.4	15	HEAT SHOCK PROTEIN 81.4	protein folding
w/ both the tags	Glyma.05G243500.1.p	26	50.3	1	Glycosyltransferase family 29 (sialyltransferase) family protein	protein glycosylation
w/ both the tags	Glyma.03G164400.1.p	26	162.8	1	ribophorin II (RPN2) family protein	protein N-linked glycosylation
w/ both the tags	Glyma.04G035100.1.p	26	79.8	1	Protein kinase superfamily protein	protein phosphorylation
w/ both the tags	Glyma.05G222700.2.p	32	76.9	2	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain	protein phosphorylation
w/ both the tags	Glyma.08G061000.1.p	22	95.6	1	S-locus lectin protein kinase family protein	protein phosphorylation
w/ both the tags	Glyma.13G223400.1.p	35	51.9	5	Integrin-linked protein kinase family	protein phosphorylation
w/ both the tags	Glyma.05G167800.1.p	117	50	4	RAB GDP dissociation inhibitor 2	protein transport
w/ both the tags	Glyma.11G069300.1.p	34	86.5	3	ARM repeat superfamily protein	protein ubiquitination
w/ both the tags	Glyma.02G204000.1.p	88	60	4	Cytosol aminopeptidase family protein	proteolysis

w/ both the tags	Glyma.02G299000.1.p	22	38.5	1	20S proteasome beta subunit PBB2	proteolysis involved in cellular protein catabolic process
w/ both the tags	Glyma.02G297400.1.p	72	29.7	2	WRKY family transcription factor family protein	regulation of transcription, DNA-dependent
w/ both the tags	Glyma.08G124200.1.p	25	23.1	1	nuclear factor Y, subunit A7	regulation of transcription, DNA-dependent
w/ both the tags	Glyma.09G248200.1.p	52	77.7	3	Integrase-type DNA-binding superfamily protein	regulation of transcription, DNA-dependent
w/ both the tags	Glyma.10G081700.1.p	319	72.7	42	BEL1-like homeodomain 1	regulation of transcription, DNA-dependent
w/ both the tags	Glyma.18G189700.1.p	33	64.6	3	POX (plant homeobox) family protein	regulation of transcription, DNA-dependent
w/ both the tags	Glyma.U009200.1.p	30	83.9	1	BEL1-like homeodomain 1	regulation of transcription, DNA-dependent
w/ both the tags	Glyma.04G018100.1.p	42	90.1	3	RIBOSOME BIOGENESIS PROTEIN TSR1 (20S RRNA ACCUMULATION	ribosome biogenesis
w/ both the tags	Glyma.16G041800.1.p	31	65.6	3	ARM repeat superfamily protein	spliceosomalsnRNP assembly
w/ both the tags	Glyma.02G158900.1.p	42	133.6	2	PIF1 helicase	telomere maintenance
w/ both the tags	Glyma.10G121300.1.p	25	104.1	1	PIF1 helicase	telomere maintenance
w/ both the tags	Glyma.04G180100.1.p	28	173.5	1	BAH domain ;TFIIS helical bundle-like domain	transcription, DNA-dependent
w/ both the tags	Glyma.02G220000.1.p	126	25.5	3	Ribosomal protein L16p/L10e family protein	translation
w/ both the tags	Glyma.05G193600.1.p	30	81.4	1	glucose-inhibited division family A protein	tRNA wobble uridine modification
w/ both the tags	Glyma.02G013600.1.p	226	26.1	10	20S proteasome alpha subunit E2	ubiquitin-dependent protein catabolic process
w/ both the tags	Glyma.01G117200.1.p	35	124.4	2	receptor like protein 7	
w/ both the tags	Glyma.01G182100.1.p	103	51.7	11	Transducin/WD40 repeat-like superfamily protein	
w/ both the tags	Glyma.01G235000.1.p	25	28.1	1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin	

w/ both the tags	Glyma.02G000700.1.p	31	80.5	1	No functional domain available
w/ both the tags	Glyma.02G002500.1.p	32	110.5	1	Stabilizer of iron transporter SufD / Polynucleotidyl transferase
w/ both the tags	Glyma.02G034800.1.p	113	15.3	12	Nucleic acid-binding, OB-fold-like protein
w/ both the tags	Glyma.02G046200.1.p	26	55.8	1	Tetratricopeptide repeat (TPR)-like superfamily protein
w/ both the tags	Glyma.02G145200.1.p	41	223.7	2	kinase interacting (KIP1-like) family protein
w/ both the tags	Glyma.02G251800.1.p	68	64.6	4	DA1-related protein 2
w/ both the tags	Glyma.02G295600.1.p	31	64	3	protein phosphatase X 2
w/ both the tags	Glyma.03G001200.1.p	23	37	1	hydrolases, acting on acid anhydrides, in phosphorus-containing anhydrides;ATP-
w/ both the tags	Glyma.03G030200.1.p	32	186.2	1	No functional domain available
w/ both the tags	Glyma.03G246700.1.p	130	37.8	4	NAD(P)-linked oxidoreductase superfamily protein
w/ both the tags	Glyma.04G158000.1.p	24	70.5	1	No functional domain available
w/ both the tags	Glyma.04G183400.1.p	116	79.9	7	Tetratricopeptide repeat (TPR)-like superfamily protein
w/ both the tags	Glyma.05G004500.1.p	49	49.1	3	Galactose oxidase/kelch repeat superfamily protein
w/ both the tags	Glyma.05G044500.1.p	164	44.8	4	26S proteasome, regulatory subunit Rpn7;Proteasome component (PCI)
w/ both the tags	Glyma.05G078100.1.p	29	19.5	4	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
w/ both the tags	Glyma.05G146100.1.p	68	78.9	5	RNA helicase family protein
w/ both the tags	Glyma.06G009900.1.p	24	48.6	1	RING/U-box superfamily protein
w/ both the tags	Glyma.06G129200.1.p	30	45.5	2	defective in meristem silencing 3

w/ both the tags	Glyma.06G214700.1.p	52	44.5	3	SOUL heme-binding family protein
w/ both the tags	Glyma.06G240000.1.p	38	56.3	5	Domain of unknown function (DUF3598)
w/ both the tags	Glyma.07G229200.1.p	57	59.9	5	SET domain protein 35
w/ both the tags	Glyma.08G156100.1.p	52	50.9	6	TRICHOME BIREFRINGENCE-LIKE 7
w/ both the tags	Glyma.08G306800.1.p	59	24.9	2	glutathione S-transferase PHI 9
w/ both the tags	Glyma.09G025500.1.p	29	68.3	1	Tetratricopeptide repeat (TPR)-like superfamily protein
w/ both the tags	Glyma.09G092600.1.p	97	25.3	8	No functional domain available
w/ both the tags	Glyma.09G122700.1.p	25	56.6	1	hAT transposon superfamily
w/ both the tags	Glyma.09G224900.1.p	59	48.4	4	scarecrow-like 3
w/ both the tags	Glyma.09G229900.1.p	24	122.6	3	P-loop containing nucleoside triphosphate hydrolases superfamily
w/ both the tags	Glyma.09G233000.1.p	70	12	11	No functional domain available
w/ both the tags	Glyma.10G236100.1.p	28	104.1	2	Myosin heavy chain-related protein
w/ both the tags	Glyma.10G247700.1.p	43	77.5	2	Pentatricopeptide repeat (PPR) superfamily protein
w/ both the tags	Glyma.10G265600.1.p	28	61.6	3	No functional domain available
w/ both the tags	Glyma.10G290400.1.p	25	43.4	1	phraGmoplastin interacting protein 1
w/ both the tags	Glyma.10G291100.1.p	67	23.5	2	dehydroascorbate reductase 2
w/ both the tags	Glyma.11G254500.1.p	32	55.8	2	Exostosin family protein
w/ both the tags	Glyma.12G123400.1.p	174	118.6	32	disease resistance family protein / LRR family protein

w/ both the tags	Glyma.12G149600.1.p	66	24.6	4	Regulator of Vps4 activity in the MVB pathway protein	
w/ both the tags	Glyma.12G227500.1.p	35	30.4	2	Polynucleotidyl transferase, ribonuclease H-like superfamily protein	
w/ both the tags	Glyma.13G060800.1.p	23	62.1	1	Pentatricopeptide repeat (PPR) superfamily protein	
w/ both the tags	Glyma.13G098200.1.p	41	30.5	3	No functional domain available	
w/ both the tags	Glyma.13G111600.1.p	31	22	3	Cupin superfamily (DUF985)	
w/ both the tags	Glyma.14G079800.1.p	33	28.4	3	Transducin/WD40 repeat-like superfamily protein	
w/ both the tags	Glyma.16G177500.1.p	25	53.4	1	CBS domain-containing protein with a domain of unknown function (DUF21)	
w/ both the tags	Glyma.17G051900.1.p	23	174.2	1	Myosin family protein with Dil domain	
w/ both the tags	Glyma.18G295800.1.p	46	106.4	4	PHOSPHATIDYLINOSITOL N-ACETYLGLUCOSAMINYLTRANSF	
w/ both the tags	Glyma.19G249200.1.p	25	38.2	1	AT hook motif DNA-binding family protein	
YFP-GmMYB176 only	Glyma.05G092300.1.p	34	30.6	2	ATP synthase subunit 1	ATP hydrolysis coupled proton transport
YFP-GmMYB176 only	Glyma.11G214800.1.p	23	82.1	1	NADH-ubiquinonedehydrogenase, mitochondrial, putative	ATP synthesis coupled electron transport
YFP-GmMYB176 only	Glyma.01G213700.1.p	42	56.1	5	ACC synthase 10	biosynthetic process
YFP-GmMYB176 only	Glyma.04G235200.1.p	38	73.1	2	Glycogen/starch synthases, ADP-glucose type	biosynthetic process
YFP-GmMYB176 only	Glyma.07G185700.1.p	51	50.3	2	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein	biosynthetic process
YFP-GmMYB176 only	Glyma.08G138800.1.p	25	50.6	1	aspartate aminotransferase	biosynthetic process
YFP-GmMYB176 only	Glyma.10G172200.1.p	136	67.5	5	UDP-Glycosyltransferase superfamily protein	biosynthetic process
YFP-GmMYB176 only	Glyma.15G157500.1.p	138	33.6	22	phosphorylcholinecytidylyltransferase	biosynthetic process

YFP-GmMYB176 only	Glyma.11G094200.1.p	25	41.9	1	Radical SAM superfamily protein	biotin biosynthetic process
YFP-GmMYB176 only	Glyma.03G192300.1.p	31	99.2	1	starch branching enzyme 2.2	carbohydrate metabolic process
YFP-GmMYB176 only	Glyma.04G017700.1.p	49	102	2	starch branching enzyme 2.2	carbohydrate metabolic process
YFP-GmMYB176 only	Glyma.07G103000.2.p	26	149	1	ketose-bisphosphate aldolase class-II family protein	carbohydrate metabolic process
YFP-GmMYB176 only	Glyma.09G038600.1.p	31	84.9	1	beta-xylosidase 1	carbohydrate metabolic process
YFP-GmMYB176 only	Glyma.09G178100.1.p	27	58	1	B-S glucosidase 44	carbohydrate metabolic process
YFP-GmMYB176 only	Glyma.13G048800.1.p	24	80.9	1	glycosyl hydrolase family 81 protein	cell wall macromolecule catabolic process
YFP-GmMYB176 only	Glyma.05G090100.1.p	34	84.9	1	Cobalamin-independent synthase family protein	cellular amino acid biosynthetic process
YFP-GmMYB176 only	Glyma.04G080700.1.p	115	45.5	3	aspartate aminotransferase 3	cellular amino acid metabolic process
YFP-GmMYB176 only	Glyma.17G052000.1.p	25	56.7	1	Pyridoxal-5\'-phosphate-dependent enzyme family protein	cellular amino acid metabolic process
YFP-GmMYB176 only	Glyma.17G216000.1.p	187	50.7	9	aspartate aminotransferase 5	cellular amino acid metabolic process
YFP-GmMYB176 only	Glyma.05G160000.1.p	28	152.3	1	cellulose synthase family protein	cellulose biosynthetic process
YFP-GmMYB176 only	Glyma.19G242300.1.p	74	34.4	2	O-acetylserine (thiol) lyase (OAS-TL) isoform A1	cysteine biosynthetic process from serine
YFP-GmMYB176 only	Glyma.08G305400.1.p	25	108.7	1	NB-ARC domain-containing disease resistance protein	defense response
YFP-GmMYB176 only	Glyma.12G027100.1.p	30	135.2	2	disease resistance protein (TIR-NBS-LRR class), putative	defense response
YFP-GmMYB176 only	Glyma.14G008700.1.p	63	215	10	RPS5-like 1	defense response
YFP-GmMYB176 only	Glyma.15G127100.1.p	25	117.7	1	NB-ARC domain-containing disease resistance protein	defense response
YFP-GmMYB176 only	Glyma.16G159100.1.p	23	133.1	1	Disease resistance protein (TIR-NBS-LRR class) family	defense response

YFP-GmMYB176 only	Glyma.18G093900.1.p	22	106.4	1	NB-ARC domain-containing disease resistance protein	defense response
YFP-GmMYB176 only	Glyma.09G047300.1.p	25	96.3	1	minichromosome maintenance (MCM2/3/5) family protein	DNA replication
YFP-GmMYB176 only	Glyma.12G080900.1.p	95	86.9	8	DNA ligase 1	DNA replication
YFP-GmMYB176 only	Glyma.01G234600.1.p	25	114.7	1	zinc knuckle (CCHC-type) family protein	DNA-dependent transcription, initiation
YFP-GmMYB176 only	Glyma.05G070900.1.p	31	38.6	2	Cyclin-like family protein	DNA-dependent transcription, initiation
YFP-GmMYB176 only	Glyma.08G024700.1.p	33	52.6	1	3-ketoacyl-acyl carrier protein synthase I	fatty acid biosynthetic process
YFP-GmMYB176 only	Glyma.20G241600.1.p	23	25.1	1	Chalcone-flavanone isomerase family protein	flavonoid biosynthetic process
YFP-GmMYB176 only	Glyma.04G177000.1.p	22	29.4	1	5-formyltetrahydrofolate cycloligase	folic acid-containing compound biosynthetic process
YFP-GmMYB176 only	Glyma.04G032600.1.p	40	62.8	2	Sugar isomerase (SIS) family protein	gluconeogenesis
YFP-GmMYB176 only	Glyma.14G162300.1.p	25	234.9	1	NADH-dependent glutamate synthase 1	glutamate biosynthetic process
YFP-GmMYB176 only	Glyma.02G014000.1.p	62	41.2	2	thioredoxin family protein	glycerol ether metabolic process
YFP-GmMYB176 only	Glyma.14G152000.1.p	49	56	1	PDI-like 1-1	glycerol ether metabolic process
YFP-GmMYB176 only	Glyma.13G222300.1.p	85	58.8	5	serine hydroxymethyltransferase 3	glycine metabolic process
YFP-GmMYB176 only	Glyma.04G239900.2.p	25	25.4	1	glycolipid transfer protein 2	glycolipid transport
YFP-GmMYB176 only	Glyma.01G109300.1.p	24	51.3	1	tubulin beta chain 3	GTP catabolic process
YFP-GmMYB176 only	Glyma.02G258200.1.p	27	151.8	1	Clathrin, heavy chain	intracellular protein transport
YFP-GmMYB176 only	Glyma.06G077600.1.p	68	84.5	16	cyclic nucleotide-gated channel 14	ion transport
YFP-GmMYB176 only	Glyma.12G152300.4.p	31	56	5	cyclic nucleotide gated channel 5	ion transport

YFP-GmMYB176 only	Glyma.01G012100.1.p	26	64.2	1	Mono-/di-acylglycerol lipase, N-terminal;Lipase, class 3	lipid metabolic process
YFP-GmMYB176 only	Glyma.07G003400.1.p	30	66.7	1	D-3-phosphoglycerate dehydrogenase	L-serine biosynthetic process
YFP-GmMYB176 only	Glyma.02G147900.1.p	22	50.1	1	Aluminium activated malate transporter family protein	malate transport
YFP-GmMYB176 only	Glyma.01G239500.1.p	65	35.3	2	alpha/beta-Hydrolases superfamily protein	metabolic process
YFP-GmMYB176 only	Glyma.01G239600.1.p	56	35.5	2	alpha/beta-Hydrolases superfamily protein	metabolic process
YFP-GmMYB176 only	Glyma.04G181100.1.p	25	49.2	1	aldehyde dehydrogenase 2B7	metabolic process
YFP-GmMYB176 only	Glyma.06G031600.1.p	25	70.7	1	AMP-dependent synthetase and ligase family protein	metabolic process
YFP-GmMYB176 only	Glyma.07G183600.1.p	30	58	1	aldehyde dehydrogenase 6B2	metabolic process
YFP-GmMYB176 only	Glyma.08G292000.2.p	33	30.8	1	IAA-leucine resistant (ILR)-like 2	metabolic process
YFP-GmMYB176 only	Glyma.09G222600.1.p	22	56.2	1	GLN phosphoribosyl pyrophosphate amidotransferase 1	metabolic process
YFP-GmMYB176 only	Glyma.13G125400.1.p	70	45.7	5	ATP citrate lyase (ACL) family protein	metabolic process
YFP-GmMYB176 only	Glyma.16G175400.1.p	24	52.7	1	UDP-glucosyl transferase 88A1	metabolic process
YFP-GmMYB176 only	Glyma.17G031800.1.p	24	63	1	acyl-activating enzyme 7	metabolic process
YFP-GmMYB176 only	Glyma.19G107300.2.p	30	85	1	acetyl-CoA synthetase	metabolic process
YFP-GmMYB176 only	Glyma.09G136400.1.p	25	76.8	1	ATP binding microtubule motor family protein	microtubule-based movement
YFP-GmMYB176 only	Glyma.10G045500.1.p	28	118.6	1	ATP binding microtubule motor family protein	microtubule-based movement
YFP-GmMYB176 only	Glyma.13G310700.1.p	27	155.4	1	P-loop containing nucleoside triphosphate hydrolases superfamily	microtubule-based movement
YFP-GmMYB176 only	Glyma.14G011100.1.p	30	127.5	4	kinesin-like protein 1	microtubule-based movement

YFP-GmMYB176 only	Glyma.15G252800.1.p	29	118	1	Di-glucose binding protein with Kinesin motor domain	microtubule-based movement
YFP-GmMYB176 only	Glyma.16G062200.1.p	36	82.4	3	global transcription factor group E8	microtubule-based movement
YFP-GmMYB176 only	Glyma.10G019200.1.p	27	106	1	DNA mismatch repair protein, putative	mismatch repair
YFP-GmMYB176 only	Glyma.20G073400.1.p	27	75.3	1	MUTL protein homolog 1	mismatch repair
YFP-GmMYB176 only	Glyma.02G034300.1.p	28	104.5	1	splicing factor PWI domain-containing protein / RNA recognition motif	mRNA processing
YFP-GmMYB176 only	Glyma.01G008600.1.p	29	69.1	2	chromatin protein family	mRNA splicing, via spliceosome
YFP-GmMYB176 only	Glyma.10G031200.1.p	27	14.6	1	histone H2A 11	nucleosome assembly
YFP-GmMYB176 only	Glyma.10G133500.1.p	175	11.4	7	Histone superfamily protein	nucleosome assembly
YFP-GmMYB176 only	Glyma.01G235600.1.p	25	45.4	1	12-oxophytodienoate reductase 2	oxidation-reduction process
YFP-GmMYB176 only	Glyma.02G272200.1.p	27	61.8	1	abscisic aldehyde oxidase 3	oxidation-reduction process
YFP-GmMYB176 only	Glyma.03G237300.1.p	189	97.9	9	lipoxygenase 1	oxidation-reduction process
YFP-GmMYB176 only	Glyma.04G019500.1.p	25	60.9	1	SKU5 similar 5	oxidation-reduction process
YFP-GmMYB176 only	Glyma.07G225400.1.p	30	62.1	2	SKU5 similar 12	oxidation-reduction process
YFP-GmMYB176 only	Glyma.08G186500.1.p	35	48.8	3	oxidoreductase, 2OG-Fe(II) oxygenase family protein	oxidation-reduction process
YFP-GmMYB176 only	Glyma.10G153900.1.p	24	98.4	1	lipoxygenase 1	oxidation-reduction process
YFP-GmMYB176 only	Glyma.12G082100.1.p	22	64.3	1	Pheophorbide a oxygenase family protein with Rieske [2Fe-2S] domain	oxidation-reduction process
YFP-GmMYB176 only	Glyma.15G026400.1.p	2444	96.4	92	lipoxygenase 1	oxidation-reduction process
YFP-GmMYB176 only	Glyma.19G170800.1.p	35	48.2	1	Flavin-binding monooxygenase family protein	oxidation-reduction process

YFP-GmMYB176 only	Glyma.05G169400.1.p	26	88.7	1	Phosphatidylinositol-4-phosphate 5-kinase family protein	phosphatidylinositol metabolic process
YFP-GmMYB176 only	Glyma.08G283700.1.p	71	164.2	6	Pyruvate phosphate dikinase, PEP/pyruvate binding domain	phosphorylation
YFP-GmMYB176 only	Glyma.04G167900.1.p	24	28.2	1	light-harvesting chlorophyll-protein complex I subunit A4	photosynthesis, light harvesting
YFP-GmMYB176 only	Glyma.06G294600.1.p	24	28.9	1	expansin A20	plant-type cell wall organization
YFP-GmMYB176 only	Glyma.12G229700.2.p	28	45.4	2	Uroporphyrinogen decarboxylase	porphyrin-containing compound biosynthetic process
YFP-GmMYB176 only	Glyma.14G003200.1.p	22	43.5	1	Coproporphyrinogen III oxidase	porphyrin-containing compound biosynthetic process
YFP-GmMYB176 only	Glyma.05G117500.1.p	28	95	1	K+ uptake permease 7	potassium ion transmembrane transport
YFP-GmMYB176 only	Glyma.20G131500.1.p	22	41.6	1	Protein phosphatase 2C family protein	protein dephosphorylation
YFP-GmMYB176 only	Glyma.07G103600.1.p	39	58.5	7	Molecular chaperone Hsp40/DnaJ family protein	protein folding
YFP-GmMYB176 only	Glyma.10G193200.1.p	565	61.7	20	heat shock protein 60	protein folding
YFP-GmMYB176 only	Glyma.06G202200.1.p	25	83.5	1	heat shock protein 101	protein metabolic process
YFP-GmMYB176 only	Glyma.13G101300.1.p	26	113.3	1	Double Clp-N motif-containing P-loop nucleoside triphosphate hydrolases	protein metabolic process
YFP-GmMYB176 only	Glyma.02G177000.1.p	31	83.7	1	STRUBBELIG-receptor family 2	protein phosphorylation
YFP-GmMYB176 only	Glyma.04G063600.1.p	26	43.6	1	Protein kinase superfamily protein	protein phosphorylation
YFP-GmMYB176 only	Glyma.07G173100.1.p	24	101.9	1	transmembrane kinase 1	protein phosphorylation
YFP-GmMYB176 only	Glyma.09G270400.1.p	42	19.5	3	cysteine-rich RLK (RECEPTOR-like protein kinase) 25	protein phosphorylation
YFP-GmMYB176 only	Glyma.10G120300.1.p	36	65.4	1	calcium dependent protein kinase 1	protein phosphorylation
YFP-GmMYB176 only	Glyma.13G283200.1.p	25	92.9	1	S-locus lectin protein kinase family protein	protein phosphorylation

YFP-GmMYB176 only	Glyma.14G011500.1.p	24	119.9	1	Leucine-rich repeat receptor-like protein kinase family protein	protein phosphorylation
YFP-GmMYB176 only	Glyma.14G054500.1.p	26	93.8	1	S-locus lectin protein kinase family protein	protein phosphorylation
YFP-GmMYB176 only	Glyma.14G157600.1.p	32	84.9	1	wall-associated kinase 2	protein phosphorylation
YFP-GmMYB176 only	Glyma.19G191600.1.p	27	90.9	1	Protein kinase superfamily protein	protein phosphorylation
YFP-GmMYB176 only	Glyma.20G220400.1.p	27	67.9	1	Protein kinase superfamily protein	protein phosphorylation
YFP-GmMYB176 only	Glyma.U008400.1.p	42	36.9	5	carboxyl terminus of HSC70-interacting protein	protein ubiquitination
YFP-GmMYB176 only	Glyma.03G125200.1.p	724	38.1	24	serine carboxypeptidase-like 45	proteolysis
YFP-GmMYB176 only	Glyma.08G184200.1.p	31	248.9	1	homolog of separase	proteolysis
YFP-GmMYB176 only	Glyma.11G193800.1.p	25	55.1	1	serine carboxypeptidase-like 48	proteolysis
YFP-GmMYB176 only	Glyma.13G044300.1.p	28	43.3	1	metacaspase 3	proteolysis
YFP-GmMYB176 only	Glyma.17G240600.1.p	25	24.5	1	Subtilisin-like serine endopeptidase family protein	proteolysis
YFP-GmMYB176 only	Glyma.18G065600.1.p	26	86.5	1	FTSH protease 9	proteolysis
YFP-GmMYB176 only	Glyma.20G015600.1.p	25	49.3	1	serine carboxypeptidase-like 50	proteolysis
YFP-GmMYB176 only	Glyma.17G020600.1.p	110	104.9	7	pyruvate orthophosphate dikinase	pyruvate metabolic process
YFP-GmMYB176 only	Glyma.03G022900.1.p	115	79.2	7	SEC7-like guanine nucleotide exchange family protein	regulation of ARF protein signal transduction
YFP-GmMYB176 only	Glyma.13G031500.1.p	25	57	1	Cyclin A1;1	regulation of cyclin-dependent protein serine/threonine kinase
YFP-GmMYB176 only	Glyma.04G222200.1.p	46	33.1	3	Basic-leucine zipper (bZIP) transcription factor family protein	regulation of transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.07G104700.1.p	24	137.1	1	NF-X-like 1	regulation of transcription, DNA-dependent

YFP-GmMYB176 only	Glyma.07G229100.1.p	24	26.9	1	NAC-like, activated by AP3/PI	regulation of transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.07G239600.1.p	26	86.4	1	PAS domain-containing protein tyrosine kinase family protein	regulation of transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.09G080000.1.p	25	51.2	1	WRKY family transcription factor	regulation of transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.10G204700.1.p	25	48	1	NAC (No Apical Meristem) domain transcriptional regulator superfamily	regulation of transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.13G180200.1.p	28	25.1	1	heat shock transcription factor B2A	regulation of transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.19G118000.1.p	30	54.1	1	cycling DOF factor 3	regulation of transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.09G277800.1.p	62	34.8	2	Peroxidase superfamily protein	response to oxidative stress
YFP-GmMYB176 only	Glyma.10G222500.1.p	33	38.5	1	Peroxidase superfamily protein	response to oxidative stress
YFP-GmMYB176 only	Glyma.11G125100.1.p	23	139.1	1	P-loop containing nucleoside triphosphate hydrolases superfamily	ribosome biogenesis
YFP-GmMYB176 only	Glyma.07G012000.1.p	28	58.1	1	Protein phosphatase 2A, regulatory subunit PR55	signal transduction
YFP-GmMYB176 only	Glyma.03G093100.1.p	94	43.8	16	PIF1 helicase	telomere maintenance
YFP-GmMYB176 only	Glyma.19G090300.1.p	27	137.8	1	PIF1 helicase	telomere maintenance
YFP-GmMYB176 only	Glyma.19G012300.1.p	25	169.9	1	BAH domain ;TFIIS helical bundle-like domain	transcription, DNA-dependent
YFP-GmMYB176 only	Glyma.02G002400.1.p	24	8.3	1	Ribosomal L28 family	translation
YFP-GmMYB176 only	Glyma.07G198500.1.p	22	32	1	Ribosomal protein S5 family protein	translation
YFP-GmMYB176 only	Glyma.10G106200.1.p	24	15.2	1	glycine-rich RNA-binding protein 2	translation
YFP-GmMYB176 only	Glyma.11G231000.1.p	31	30.2	1	Ribosomal protein L1p/L10e family	translation
YFP-GmMYB176 only	Glyma.15G095300.1.p	28	16.4	1	Ribosomal protein S11 family protein	translation

YFP-GmMYB176 only	Glyma.07G005300.1.p	27	33.2	1	Mitochondrial substrate carrier family protein	transmembrane transport
YFP-GmMYB176 only	Glyma.09G067800.1.p	30	83.7	1	mechanosensitive channel of small conductance-like 10	transmembrane transport
YFP-GmMYB176 only	Glyma.02G053200.1.p	31	63.9	1	SEC14-like 12	transport
YFP-GmMYB176 only	Glyma.06G290600.1.p	43	39.8	2	ADP/ATP carrier 3	transport
YFP-GmMYB176 only	Glyma.10G137600.1.p	27	138.9	1	ATP binding cassette subfamily B4	transport
YFP-GmMYB176 only	Glyma.16G061600.1.p	29	103.3	2	glutamate receptor 2	transport
YFP-GmMYB176 only	Glyma.18G080900.1.p	29	146.1	1	multidrug resistance-associated protein 3	transport
YFP-GmMYB176 only	Glyma.06G126300.1.p	27	42.2	1	trehalose-6-phosphate phosphatase	trehalose biosynthetic process
YFP-GmMYB176 only	Glyma.17G223600.1.p	25	60.8	1	Class II aminoacyl-tRNA and biotin synthetases superfamily protein	tRNAaminoacylation for protein translation
YFP-GmMYB176 only	Glyma.02G106500.1.p	48	85.4	5	cullin 3	ubiquitin-dependent protein catabolic process
YFP-GmMYB176 only	Glyma.17G111000.1.p	38	25.8	2	vesicle-associated membrane protein 713	vesicle-mediated transport
YFP-GmMYB176 only	Glyma.01G001200.1.p	22	91	1	ARM repeat superfamily protein	
YFP-GmMYB176 only	Glyma.01G053100.1.p	25	118.7	1	Argonaute family protein	
YFP-GmMYB176 only	Glyma.01G055600.1.p	22	63.5	1	uncharacterized	
YFP-GmMYB176 only	Glyma.01G086400.1.p	23	141.4	1	PIF1 helicase	
YFP-GmMYB176 only	Glyma.01G097900.1.p	23	73.3	1	Zinc finger (C3HC4-type RING finger) family protein	
YFP-GmMYB176 only	Glyma.01G107000.1.p	60	106.6	4	Protein of unknown function (DUF3527)	
YFP-GmMYB176 only	Glyma.01G147500.3.p	36	68.6	1	Exostosin family protein	

YFP-GmMYB176 only	Glyma.01G175100.1.p	25	57.4	1	U1 small nuclear ribonucleoprotein-70K
YFP-GmMYB176 only	Glyma.01G221100.1.p	27	32.3	1	CCT motif family protein
YFP-GmMYB176 only	Glyma.01G233300.1.p	43	140.7	3	Insulinase (Peptidase family M16) family protein
YFP-GmMYB176 only	Glyma.02G013900.1.p	28	42.2	1	myb domain protein 12
YFP-GmMYB176 only	Glyma.02G021400.1.p	26	31.6	1	SPFH/Band 7/PHB domain-containing membrane-associated protein family
YFP-GmMYB176 only	Glyma.02G118600.1.p	23	88.1	1	pentatricopeptide (PPR) repeat-containing protein
YFP-GmMYB176 only	Glyma.02G133400.1.p	24	219.8	1	autophagy 2
YFP-GmMYB176 only	Glyma.02G144200.1.p	26	404	1	Beige/BEACH domain ;WD domain, G-beta repeat protein
YFP-GmMYB176 only	Glyma.02G144600.2.p	24	101.8	1	ETO1-like 1
YFP-GmMYB176 only	Glyma.02G213400.1.p	24	115.6	1	ubiquitin-specific protease 17
YFP-GmMYB176 only	Glyma.02G219100.1.p	26	19.1	1	No functional domain available
YFP-GmMYB176 only	Glyma.02G222200.1.p	29	52.2	1	Calcium-binding endonuclease/exonuclease/phosphatase
YFP-GmMYB176 only	Glyma.02G301600.1.p	278	80.6	13	Sec14p-like phosphatidylinositol transfer family protein
YFP-GmMYB176 only	Glyma.03G055300.1.p	24	23.1	1	receptor like protein 47
YFP-GmMYB176 only	Glyma.03G064800.1.p	27	103.5	2	Sterile alpha motif (SAM) domain-containing protein
YFP-GmMYB176 only	Glyma.03G087400.1.p	25	24.8	1	receptor like protein 23
YFP-GmMYB176 only	Glyma.03G116000.1.p	24	54.8	1	F-box family protein
YFP-GmMYB176 only	Glyma.03G155100.1.p	228	55.3	10	RNI-like superfamily protein

YFP-GmMYB176 only	Glyma.03G163500.4.p	804	56.3	30	RmlC-like cupins superfamily protein
YFP-GmMYB176 only	Glyma.03G163500.5.p	105	74.9	15	RmlC-like cupins superfamily protein
YFP-GmMYB176 only	Glyma.03G168000.1.p	634	71.8	26	pleiotropic drug resistance 12
YFP-GmMYB176 only	Glyma.03G185100.1.p	27	68.2	1	Major facilitator superfamily protein
YFP-GmMYB176 only	Glyma.03G198400.1.p	23	39.8	1	decapping 1
YFP-GmMYB176 only	Glyma.04G002700.2.p	24	18.7	1	No functional domain available
YFP-GmMYB176 only	Glyma.04G012300.1.p	221	33.9	10	NmrA-like negative transcriptional regulator family protein
YFP-GmMYB176 only	Glyma.04G057900.1.p	22	44.7	1	hydroxyproline-rich glycoprotein family protein
YFP-GmMYB176 only	Glyma.04G075000.1.p	27	41.6	1	COP9-signalosome 5B
YFP-GmMYB176 only	Glyma.04G088600.1.p	304	16.3	42	No functional domain available
YFP-GmMYB176 only	Glyma.04G096900.1.p	27	411.3	1	ubiquitin-protein ligase 1
YFP-GmMYB176 only	Glyma.04G150900.1.p	26	104.4	1	Stabilizer of iron transporter SufD / Polynucleotidyl transferase
YFP-GmMYB176 only	Glyma.04G206400.1.p	24	71.2	2	Protein of unknown function (DUF668)
YFP-GmMYB176 only	Glyma.04G233100.1.p	25	31.4	1	nucleotide binding;nucleic acid binding
YFP-GmMYB176 only	Glyma.05G037400.1.p	30	54.6	2	peroxin 14
YFP-GmMYB176 only	Glyma.05G045300.1.p	27	58.5	1	RGA-like 1
YFP-GmMYB176 only	Glyma.05G077300.1.p	22	20.6	1	No functional domain available
YFP-GmMYB176 only	Glyma.05G098900.1.p	31	92.1	1	RNA recognition motif.

YFP-GmMYB176 only	Glyma.05G134400.1.p	26	57.3	1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
YFP-GmMYB176 only	Glyma.05G183900.1.p	23	90.5	1	Tetratricopeptide repeat (TPR)-like superfamily protein
YFP-GmMYB176 only	Glyma.05G196900.1.p	27	64.9	1	MAC/Perforin domain-containing protein
YFP-GmMYB176 only	Glyma.06G021300.1.p	38	9.7	2	No functional domain available
YFP-GmMYB176 only	Glyma.06G032600.1.p	51	198.3	6	GYF domain-containing protein
YFP-GmMYB176 only	Glyma.06G046300.1.p	23	49.6	1	interferon-related developmental regulator family protein / IFRD protein
YFP-GmMYB176 only	Glyma.06G050300.1.p	24	29.1	1	CCCH-type zinc finger family protein
YFP-GmMYB176 only	Glyma.06G075900.1.p	34	64.9	1	No functional domain available
YFP-GmMYB176 only	Glyma.06G139800.1.p	23	121.5	1	No functional domain available
YFP-GmMYB176 only	Glyma.06G151300.1.p	45	75.2	2	Transducin/WD40 repeat-like superfamily protein
YFP-GmMYB176 only	Glyma.06G154500.1.p	29	70.3	1	ABC-2 type transporter family protein
YFP-GmMYB176 only	Glyma.06G159000.1.p	25	47	1	No functional domain available
YFP-GmMYB176 only	Glyma.06G209800.1.p	23	69.8	1	Hydroxyproline-rich glycoprotein family protein
YFP-GmMYB176 only	Glyma.06G243200.1.p	25	9.7	1	phosphofructokinase 7
YFP-GmMYB176 only	Glyma.06G311400.1.p	31	32.6	1	nuclear factor Y, subunit C11
YFP-GmMYB176 only	Glyma.06G324500.1.p	89	58.4	5	Smg-4/UPF3 family protein
YFP-GmMYB176 only	Glyma.07G031600.1.p	27	15	1	Alba DNA/RNA-binding protein
YFP-GmMYB176 only	Glyma.07G103100.1.p	26	20.4	1	No functional domain available

YFP-GmMYB176 only	Glyma.07G128000.1.p	29	50.5	1	1-amino-cyclopropane-1-carboxylate synthase 7
YFP-GmMYB176 only	Glyma.07G133100.1.p	25	83.6	1	GRAS family transcription factor
YFP-GmMYB176 only	Glyma.07G139600.1.p	25	26.1	1	glutathione S-transferase tau 7
YFP-GmMYB176 only	Glyma.07G178800.1.p	22	88.6	1	endonucleases
YFP-GmMYB176 only	Glyma.07G205800.1.p	29	30.3	1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
YFP-GmMYB176 only	Glyma.07G228400.1.p	30	93.9	1	EXS (ERD1/XPR1/SYG1) family protein
YFP-GmMYB176 only	Glyma.07G232500.1.p	84	45.9	23	CRT (chloroquine-resistance transporter)-like transporter 3
YFP-GmMYB176 only	Glyma.07G245500.1.p	34	79.5	1	Tetratricopeptide repeat (TPR)-like superfamily protein
YFP-GmMYB176 only	Glyma.07G247200.1.p	24	59.4	1	No functional domain available
YFP-GmMYB176 only	Glyma.08G021400.1.p	26	163	1	COP1-interactive protein 1
YFP-GmMYB176 only	Glyma.08G099800.1.p	28	28.8	1	No functional domain available
YFP-GmMYB176 only	Glyma.08G134700.1.p	28	36.1	1	C2H2 and C2HC zinc fingers superfamily protein
YFP-GmMYB176 only	Glyma.08G148300.1.p	36	209.3	1	guanyl-nucleotide exchange factors;GTPasebinding;GTP binding
YFP-GmMYB176 only	Glyma.08G175200.1.p	39	25.7	2	glutathione S-transferase TAU 19
YFP-GmMYB176 only	Glyma.08G201000.1.p	29	75.4	1	hydroxyproline-rich glycoprotein family protein
YFP-GmMYB176 only	Glyma.08G238800.1.p	25	103.9	1	Phosphoinositide phosphatase family protein
YFP-GmMYB176 only	Glyma.08G279900.1.p	25	30.9	1	Chaperone DnaJ-domain superfamily protein
YFP-GmMYB176 only	Glyma.08G341900.1.p	36	20.6	1	drought-repressed 4

YFP-GmMYB176 only	Glyma.08G348200.1.p	25	6.8	1	wound-responsive protein-related
YFP-GmMYB176 only	Glyma.09G014600.1.p	24	187.7	1	binding
YFP-GmMYB176 only	Glyma.09G018300.1.p	25	91.8	1	Membrane trafficking VPS53 family protein
YFP-GmMYB176 only	Glyma.09G030100.1.p	25	52.6	1	hydroxyproline-rich glycoprotein family protein
YFP-GmMYB176 only	Glyma.09G046300.1.p	36	186	2	CW-type Zinc Finger
YFP-GmMYB176 only	Glyma.09G052200.1.p	25	137.6	1	SET domain protein 25
YFP-GmMYB176 only	Glyma.09G079600.1.p	31	64.7	1	hydroxyproline-rich glycoprotein family protein
YFP-GmMYB176 only	Glyma.09G079600.3.p	36	64.7	2	hydroxyproline-rich glycoprotein family protein
YFP-GmMYB176 only	Glyma.09G083900.1.p	25	38.9	1	hAT transposon superfamily
YFP-GmMYB176 only	Glyma.09G165600.1.p	31	25.8	1	Translation initiation factor eIF3 subunit
YFP-GmMYB176 only	Glyma.09G215100.1.p	46	133.4	1	actin binding
YFP-GmMYB176 only	Glyma.09G226400.1.p	32	126	1	transducin family protein / WD-40 repeat family protein
YFP-GmMYB176 only	Glyma.10G004100.1.p	27	137.7	1	COP1-interacting protein-related
YFP-GmMYB176 only	Glyma.10G028300.1.p	36	58.4	1	cupin family protein
YFP-GmMYB176 only	Glyma.10G037100.1.p	105	64.2	5	RmlC-like cupins superfamily protein
YFP-GmMYB176 only	Glyma.10G050400.1.p	23	44.8	1	No functional domain available
YFP-GmMYB176 only	Glyma.10G124800.1.p	23	36.8	1	potassium channel beta subunit 1
YFP-GmMYB176 only	Glyma.10G168600.1.p	23	101.3	1	plastid transcriptionally active 3

YFP-GmMYB176 only	Glyma.10G204800.1.p	23	40.6	1	NAD(P)-binding Rossmann-fold superfamily protein
YFP-GmMYB176 only	Glyma.10G246300.1.p	113	72.5	6	cupin family protein
YFP-GmMYB176 only	Glyma.10G257100.1.p	26	33.1	1	cyclin-related
YFP-GmMYB176 only	Glyma.10G258400.1.p	25	66.7	1	Family of unknown function (DUF566)
YFP-GmMYB176 only	Glyma.11G003100.1.p	24	34.7	1	No functional domain available
YFP-GmMYB176 only	Glyma.11G044500.1.p	31	55.2	1	No functional domain available
YFP-GmMYB176 only	Glyma.11G077500.1.p	29	72.4	1	armadillo repeat only 2
YFP-GmMYB176 only	Glyma.11G096200.1.p	109	58.5	5	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein
YFP-GmMYB176 only	Glyma.11G109500.1.p	23	111.4	1	Frigida-like protein
YFP-GmMYB176 only	Glyma.11G121100.1.p	28	59.6	1	NAD(P)-binding Rossmann-fold superfamily protein
YFP-GmMYB176 only	Glyma.11G136700.1.p	27	114.2	1	pumilio 2
YFP-GmMYB176 only	Glyma.11G137300.1.p	26	40.5	1	zinc finger (C3HC4-type RING finger) family protein
YFP-GmMYB176 only	Glyma.11G138500.1.p	25	76.6	1	SCARECROW-like 14
YFP-GmMYB176 only	Glyma.11G143800.1.p	23	38.6	1	RmlC-like cupins superfamily protein
YFP-GmMYB176 only	Glyma.11G152400.1.p	23	21.5	1	NADPH:quinone oxidoreductase
YFP-GmMYB176 only	Glyma.11G152500.1.p	26	92	1	ARM repeat superfamily protein
YFP-GmMYB176 only	Glyma.11G183900.1.p	28	28.5	1	alpha/beta-Hydrolases superfamily protein
YFP-GmMYB176 only	Glyma.11G242400.1.p	24	22.2	1	No functional domain available

YFP-GmMYB176 only	Glyma.12G020000.1.p	26	39.1	2	Protein of unknown function (DUF3411)
YFP-GmMYB176 only	Glyma.12G083000.1.p	31	26	1	PLATZ transcription factor family protein
YFP-GmMYB176 only	Glyma.12G100900.1.p	26	12	1	Ubiquitin-like superfamily protein
YFP-GmMYB176 only	Glyma.12G166900.1.p	225	60.5	41	DNA-binding bromodomain-containing protein
YFP-GmMYB176 only	Glyma.12G186700.1.p	24	52.8	1	Leucine-rich repeat (LRR) family protein
YFP-GmMYB176 only	Glyma.12G188700.1.p	26	3.8	1	conserved peptide upstream open reading frame 6
YFP-GmMYB176 only	Glyma.12G204600.1.p	35	25	2	No functional domain available
YFP-GmMYB176 only	Glyma.12G209500.1.p	34	57.3	1	late embryogenesis abundant domain-containing protein / LEA domain-
YFP-GmMYB176 only	Glyma.12G214100.1.p	28	24.5	1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
YFP-GmMYB176 only	Glyma.12G224900.1.p	26	35.5	1	RNA-binding S4 domain-containing protein
YFP-GmMYB176 only	Glyma.13G047700.2.p	34	58.1	1	Proline-rich spliceosome-associated (PSP) family protein / zinc knuckle
YFP-GmMYB176 only	Glyma.13G056100.1.p	34	52	1	HXXXD-type acyl-transferase family protein
YFP-GmMYB176 only	Glyma.13G088200.1.p	26	44.1	1	Endoplasmic reticulum vesicle transporter protein
YFP-GmMYB176 only	Glyma.13G112200.1.p	24	31.8	1	Protein of unknown function (DUF581)
YFP-GmMYB176 only	Glyma.13G120800.1.p	27	122.7	1	P-loop containing nucleoside triphosphate hydrolases superfamily
YFP-GmMYB176 only	Glyma.13G133000.1.p	26	20.1	1	C2H2 and C2HC zinc fingers superfamily protein
YFP-GmMYB176 only	Glyma.13G165000.1.p	23	15.5	1	No functional domain available
YFP-GmMYB176 only	Glyma.13G287000.1.p	28	99.7	1	homolog of yeast autophagy 18 (ATG18) F

YFP-GmMYB176 only	Glyma.13G292200.1.p	30	70.3	1	global transcription factor group E4
YFP-GmMYB176 only	Glyma.13G299600.1.p	29	36.6	1	Peroxisomal membrane 22 kDa (Mpv17/PMP22) family protein
YFP-GmMYB176 only	Glyma.14G074900.1.p	28	81.1	1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
YFP-GmMYB176 only	Glyma.14G087300.1.p	26	52.4	1	Protein of unknown function (DUF620)
YFP-GmMYB176 only	Glyma.14G123800.1.p	34	22	2	No functional domain available
YFP-GmMYB176 only	Glyma.14G133100.1.p	25	20.4	1	selenium binding
YFP-GmMYB176 only	Glyma.14G150900.1.p	25	102.5	1	Plant protein of unknown function (DUF869)
YFP-GmMYB176 only	Glyma.15G025700.1.p	24	26.6	1	serine/threonine phosphatase 7
YFP-GmMYB176 only	Glyma.15G033000.1.p	23	88	1	Putative glycosyl hydrolase of unknown function (DUF1680)
YFP-GmMYB176 only	Glyma.15G061000.1.p	47	43.6	3	LisH/CRA/RING-U-box domains-containing protein
YFP-GmMYB176 only	Glyma.15G078400.1.p	31	39.8	1	Guanylate-binding family protein
YFP-GmMYB176 only	Glyma.15G092000.1.p	22	162.8	1	breast cancer susceptibility1
YFP-GmMYB176 only	Glyma.15G131600.1.p	24	68.7	1	Tetratricopeptide repeat (TPR)-like superfamily protein
YFP-GmMYB176 only	Glyma.15G205600.1.p	26	147	1	homolog of yeast FIP1 [V]
YFP-GmMYB176 only	Glyma.15G229400.1.p	24	44.7	1	No functional domain available
YFP-GmMYB176 only	Glyma.15G237600.1.p	27	73.6	1	No functional domain available
YFP-GmMYB176 only	Glyma.15G252200.1.p	141	26.1	4	glutathione S-transferase TAU 19
YFP-GmMYB176 only	Glyma.16G053700.1.p	35	72.7	2	Tetratricopeptide repeat (TPR)-like superfamily protein

YFP-GmMYB176 only	Glyma.16G058700.1.p	24	95.7	1	Tetratricopeptide repeat (TPR)-like superfamily protein
YFP-GmMYB176 only	Glyma.16G070300.1.p	23	36.7	1	ARM repeat superfamily protein
YFP-GmMYB176 only	Glyma.16G071800.1.p	58	23.4	2	Oleosin family protein
YFP-GmMYB176 only	Glyma.16G111400.1.p	59	38.5	9	No functional domain available
YFP-GmMYB176 only	Glyma.16G205800.1.p	30	7.8	1	No functional domain available
YFP-GmMYB176 only	Glyma.17G013900.1.p	33	23.9	2	Protein of unknown function, DUF584
YFP-GmMYB176 only	Glyma.17G024700.1.p	23	54.1	1	homolog of yeast sucrose nonfermenting 4
YFP-GmMYB176 only	Glyma.17G025500.1.p	33	108.2	1	No functional domain available
YFP-GmMYB176 only	Glyma.17G076400.1.p	26	190.6	1	Microfibrillar-associated protein MFAP1
YFP-GmMYB176 only	Glyma.17G115100.1.p	27	48.7	1	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family
YFP-GmMYB176 only	Glyma.17G164300.1.p	23	41.3	1	No functional domain available
YFP-GmMYB176 only	Glyma.17G171700.1.p	29	22.4	1	zinc finger protein 4
YFP-GmMYB176 only	Glyma.17G178100.1.p	24	22.3	1	No functional domain available
YFP-GmMYB176 only	Glyma.17G193100.1.p	26	21.6	1	zinc ion binding;nucleic acid binding
YFP-GmMYB176 only	Glyma.17G197200.1.p	28	317.6	1	Uncharacterized conserved coiled-coil protein
YFP-GmMYB176 only	Glyma.17G203800.1.p	28	50.4	1	RING/U-box superfamily protein
YFP-GmMYB176 only	Glyma.17G228900.1.p	44	35.1	4	No functional domain available
YFP-GmMYB176 only	Glyma.18G017700.1.p	22	51	1	No functional domain available

YFP-GmMYB176 only	Glyma.18G037000.1.p	24	40.1	1	RING-H2 finger C2A
YFP-GmMYB176 only	Glyma.18G040400.1.p	32	31.5	2	ALWAYS EARLY 4
YFP-GmMYB176 only	Glyma.18G060500.1.p	25	30.1	1	RING/U-box superfamily protein
YFP-GmMYB176 only	Glyma.18G063400.1.p	25	167.7	1	pleiotropic drug resistance 3
YFP-GmMYB176 only	Glyma.18G074800.1.p	23	77.9	1	ABC-2 type transporter family protein
YFP-GmMYB176 only	Glyma.18G084900.1.p	24	69.6	1	Ankyrin repeat family protein
YFP-GmMYB176 only	Glyma.18G108900.1.p	26	51.5	1	Leucine-rich repeat (LRR) family protein
YFP-GmMYB176 only	Glyma.18G132300.1.p	36	60.9	2	No functional domain available
YFP-GmMYB176 only	Glyma.18G178200.1.p	25	43.2	1	FRS (FAR1 Related Sequences) transcription factor family
YFP-GmMYB176 only	Glyma.18G294000.1.p	26	20.1	1	Protein of unknown function (DUF674)
YFP-GmMYB176 only	Glyma.19G011400.1.p	28	22.4	1	HSP20-like chaperones superfamily protein
YFP-GmMYB176 only	Glyma.19G014000.1.p	23	114.9	1	CRM family member 2
YFP-GmMYB176 only	Glyma.19G032700.1.p	25	128.6	1	ARM repeat superfamily protein
YFP-GmMYB176 only	Glyma.19G037300.1.p	22	34	1	ortholog of human splicing factor SC35
YFP-GmMYB176 only	Glyma.19G063400.2.p	61	23.6	2	Oleosin family protein
YFP-GmMYB176 only	Glyma.19G069500.1.p	73	54.9	10	HXXXD-type acyl-transferase family protein
YFP-GmMYB176 only	Glyma.19G075100.1.p	27	54.1	1	Mitochondrial transcription termination factor family protein
YFP-GmMYB176 only	Glyma.19G082000.1.p	30	44	1	zinc finger (C3HC4-type RING finger) family protein

YFP-GmMYB176 only	Glyma.19G102600.1.p	28	87.9	1	hAT transposon superfamily
YFP-GmMYB176 only	Glyma.19G130400.1.p	23	19	1	VQ motif-containing protein
YFP-GmMYB176 only	Glyma.19G164900.1.p	371	54.8	16	RmlC-like cupins superfamily protein
YFP-GmMYB176 only	Glyma.19G168800.1.p	23	65.5	1	Exostosin family protein
YFP-GmMYB176 only	Glyma.19G181400.1.p	25	9.8	1	No functional domain available
YFP-GmMYB176 only	Glyma.19G207500.1.p	31	36.9	1	Cysteine proteinases superfamily protein
YFP-GmMYB176 only	Glyma.19G215400.1.p	23	64.5	1	Heat shock protein DnaJ with tetratricopeptide repeat
YFP-GmMYB176 only	Glyma.19G227600.2.p	41	19.5	1	UNCHARACTERIZED
YFP-GmMYB176 only	Glyma.19G254400.1.p	25	10.4	1	No functional domain available
YFP-GmMYB176 only	Glyma.20G016400.1.p	28	47	1	RWP-RK domain
YFP-GmMYB176 only	Glyma.20G045000.1.p	53	0	9	No functional domain available
YFP-GmMYB176 only	Glyma.20G072100.1.p	25	54.9	1	Phototropic-responsive NPH3 family protein
YFP-GmMYB176 only	Glyma.20G119100.2.p	27	47.1	1	Protein phosphatase 2C family protein
YFP-GmMYB176 only	Glyma.20G129900.1.p	24	25.3	1	Plant basic secretory protein (BSP) family protein
YFP-GmMYB176 only	Glyma.20G140300.1.p	24	37.8	1	cysteine-rich RLK (RECEPTOR-like protein kinase) 25
YFP-GmMYB176 only	Glyma.20G148200.1.p	117	50.5	3	cupin family protein
YFP-GmMYB176 only	Glyma.20G148400.1.p	393	70.5	12	cupin family protein
YFP-GmMYB176 only	Glyma.20G163700.1.p	25	38.8	1	VIRB2-interacting protein 1

YFP-GmMYB176 only	Glyma.20G235500.1.p	22	72.2	1	acyl-CoA binding protein 4	
YFP-GmMYB176 only	Glyma.U002800.1.p	28	35.9	1	F-box family protein	
YFP-GmMYB176 only	Glyma.U009400.1.p	26	66.6	1	BED zinc finger ;hAT family dimerisation domain	
YFP-GmMYB176 only	Glyma.U027200.2.p	25	151	1	chromatin remodeling 42	
YFP-GmMYB176 only	Glyma.U034800.1.p	27	8.4	1	No functional domain available	
YFP-GmMYB176 only	Glyma.U041000.1.p	50	7.5	4	No functional domain available	
YFP-GmMYB176^2	Glyma.13G270900.1.p	47	76.5	3	tetratricopetide-repeat thioredoxin-like 1	cell redox homeostasis
YFP-GmMYB176^2	Glyma.04G144300.1.p	483	52.5	9	enolase 1	glycolysis
YFP-GmMYB176^2	Glyma.10G039000.1.p	133	45.7	5	ATP citrate lyase (ACL) family protein	metabolic process
YFP-GmMYB176^2	Glyma.15G007100.1.p	25	276.8	1	Pre-mRNA-processing-splicing factor	mRNA splicing, via spliceosome
YFP-GmMYB176^2	Glyma.02G163800.1.p	49	88.7	4	phosphatidyl inositol monophosphate 5 kinase 4	phosphatidylinositol metabolic process
YFP-GmMYB176^2	Glyma.04G013300.1.p	36	69.1	2	Protein kinase superfamily protein	protein phosphorylation
YFP-GmMYB176^2	Glyma.10G278100.2.p	30	27.5	1	vesicle-associated membrane protein 727	vesicle-mediated transport
YFP-GmMYB176^2	Glyma.01G121100.1.p	56	14	2	Plant protein of unknown function (DUF639)	
YFP-GmMYB176^2	Glyma.02G029400.1.p	66	28.3	2	C2H2 and C2HC zinc fingers superfamily protein	
YFP-GmMYB176^2	Glyma.03G207600.1.p	27	128.9	1	sister-chromatid cohesion protein 3	
YFP-GmMYB176^2	Glyma.05G002100.1.p	31	132.3	1	HEAT repeat-containing protein	
YFP-GmMYB176^2	Glyma.06G042700.1.p	50	36.5	1	Serine/threonine-protein kinase WNK (With No Lysine)-related	

YFP-GmMYB176[2]	Glyma.13G351100.1.p	24	30.8	1	phospholipase D alpha 4
YFP-GmMYB176[2]	Glyma.16G211200.1.p	28	87	1	kokopelli

Appendix 2 Pathway enrichment of the protein candidates that co-immunoprecipitated with GmMYB176.

'Pathway Enrichment' was done using PhytoMine tool, which uses data from KEGG and PlantCyc, from version 11 of Phytozome and with the *Glycine max* database as the reference. The number of gene products not analyzed in this test, due to a lack of pathway classification, was 560.The statistical test used was Benjamini-Hochberg, generating p values for all the listed pathway categories. The columns indicate: pathway name, number of gene products in the candidate list matching the pathway, and the *p* value (determined by a binomial test of the number of matches against the expected values).

Pathways	**p-Value**	**Matches**
Proteasome	0.011575	10
Ribosome	0.239357	17
Propanoate metabolism	0.243249	4
Carbon fixation in photosynthetic organisms	0.261969	8
Glycine, serine and threonine metabolism	0.276346	7
Cysteine and methionine metabolism	0.283468	9
Citrate cycle (TCA cycle)	0.335657	6
Linoleic acid metabolism	0.369247	4
Selenocompound metabolism	0.42539	3
alpha-Linolenic acid metabolism	0.465572	6
Mismatch repair	0.595804	4
Inositol phosphate metabolism	0.609849	5
Glyoxylate and dicarboxylate metabolism	0.697882	5
Glycolysis / Gluconeogenesis	0.724517	9
Pyruvate metabolism	0.757709	7

Alanine, aspartate and glutamate metabolism	0.975124	4
Porphyrin and chlorophyll metabolism	0.982913	3
Homologous recombination	0.998488	3
Pentose phosphate pathway	1	3
Pentose and glucuronate interconversions	1	2
Fructose and mannose metabolism	1	4
Galactose metabolism	1	1
Ascorbate and aldarate metabolism	1	3
Fatty acid degradation	1	1
Oxidative phosphorylation	1	3
Photosynthesis - antenna proteins	1	1
Purine metabolism	1	3
Pyrimidine metabolism	1	2
Valine, leucine and isoleucine degradation	1	2
Valine, leucine and isoleucine biosynthesis	1	1
Lysine biosynthesis	1	1
Lysine degradation	1	1
Arginine and proline metabolism	1	3
Histidine metabolism	1	1
Tyrosine metabolism	1	1
Phenylalanine metabolism	1	2
Tryptophan metabolism	1	1
beta-Alanine metabolism	1	3
Cyanoamino acid metabolism	1	2
Glutathione metabolism	1	3
Starch and sucrose metabolism	1	6
Amino sugar and nucleotide sugar metabolism	1	1
Glycerolipid metabolism	1	1
Glycerophospholipid metabolism	1	1
Ether lipid metabolism	1	1
C5-Branched dibasic acid metabolism	1	1
One carbon pool by folate	1	1

Vitamin B6 metabolism	1	1
Biotin metabolism	1	1
Terpenoid backbone biosynthesis	1	1
Limonene and pinene degradation	1	2
Carotenoid biosynthesis	1	1
Nitrogen metabolism	1	2
Sulfur metabolism	1	2
Phenylpropanoid biosynthesis	1	4
Flavonoid biosynthesis	1	3
Stilbenoid, diarylheptanoid and gingerol biosynthesis	1	1
Aminoacyl-tRNA biosynthesis	1	2
Ribosome biogenesis in eukaryotes	1	5
RNA transport	1	5
mRNA surveillance pathway	1	4
RNA degradation	1	5
Basal transcription factors	1	1
DNA replication	1	4
Spliceosome	1	6
Protein export	1	1
Base excision repair	1	2
Nucleotide excision repair	1	3
Phosphatidylinositol signaling system	1	2
Plant hormone signal transduction	1	4
Ubiquitin mediated proteolysis	1	3
SNARE interactions in vesicular transport	1	1
Regulation of autophagy	1	1
Protein processing in endoplasmic reticulum	1	9
Endocytosis	1	4
Phagosome	1	2
Peroxisome	1	2
Plant-pathogen interaction	1	7
Circadian rhythm - plant	1	1

Appendix 3 Regulatory element (RE) sites on the *CHS8*Δ*P5* fragment of *GmCHS8* gene promoter.

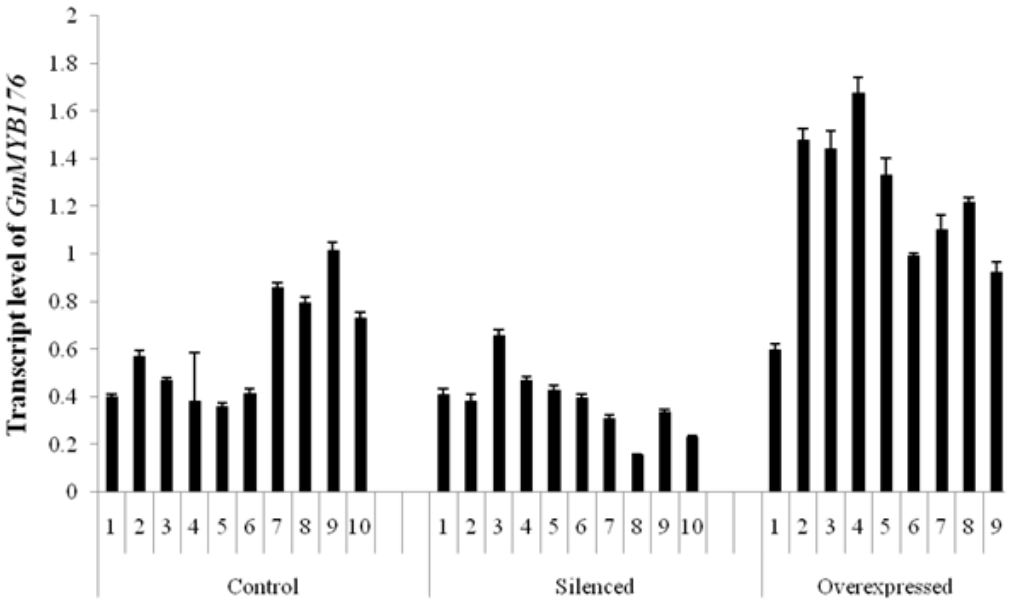
*CHS8*Δ*P5* promoter fragment was analysed for regulatory element sites in NSITE-PL Version 5.2013. Data from REGSITE DB: Plant Transcription Regulatory Sites [2779 elements; Last update: May 03, 2014]; Softberry Inc. with search parameters as follows, expected mean number of 0.1000000; Statistical Significance Level of 0.9900000; Level of homology between known RE and motif was 100%; Variation of Distance between RE Blocks was 20%;

Accession No of RE	Organism / species	Gene	RE name	Binding factor	Sequence
RSP00010	*Catharanthusroseus*	Str	G-box	TAF-1	CACGTG
RSP00016	*Glycine max*	beta-	RY	unknown nuclear	CATGCAC
RSP00024	*Pisumsativum*	petE	A/T-rich BOX3	PCF1	TATTATTT
RSP00112	*Glycine max*	GH3	TGA1	unknown nuclear	TGACGTAA
RSP00134	*Oryzasativa*	rifa-7-P-	G-box	Different bZIP	CACGTG
RSP00172	*Phaseolus vulgaris*	beta-phaseolin,	CACTTC motif	unknown nuclear	CACTTC
RSP00241	*Zea mays*	Bronze2	C1-motif (2)	unknown nuclear	CGGTCA
RSP00302	*Brassica napus*	napA	G-box	ABI3	CACGTG
RSP00329	*Arabidopsis thaliana*	erd1	ACGT motif	unknown nuclear	TACGTC
RSP00390	*Phaseolus vulgaris*	chs15	G-BOX	unknown nuclear	CACGTG
RSP00543	*Lycopersiconesculentum*	rbcS3B	GT-1-like K	unknown nuclear	AGGCATT
RSP00615	*Lycopersiconesculentum*	rbcS3A	C-rich GG1	unknown nuclear	CTCACC
RSP00657	*Phaseolus vulgaris*	beta-phaseolin	G-box	unknown	CACGTG
RSP00726	*Nicotiana tabacum*	Ubi.U4	G-box	unknown	CACGTG
RSP00735	*Arabidopsis thaliana*	CAB2	CCA1 BS	DET1; CCA1	AAAAATCA
RSP00786	*Oryzasativa*	osRACD	G-box	unknown	CACGTG
RSP00815	*Arabidopsis thaliana*	RAV1	G box	Pti4	CACGTG
RSP00864	*Arabidopsis thaliana*	STK	GA-5	BPC1	AGAGAGAGA
RSP00957	*Lycopersiconesculentum*	pin2	G box	JAMUC1;	CACGTG
RSP00990	*Arabidopsis thaliana*	CYCB1;1	p33TCP20 BS1	p33TCP20	GCCCG

RSP01060	*Helianthus annuus*	OLIGOs	Hahb-4 BS	Hahb-4	CAATTATTG
RSP01077	*CraterostiGmaplantagin*	CpC2	HDZIP BS	CpbZIP1; CpbZIP2	CAATAATTG
RSP01453	*Pinustaeda*	Synthetic	AC-I	PtMYB4	ACCTACC
RSP01471	*Pisumsativum*	PsCHS1	G-box	PsGBF	CACGTG
RSP01532	*Pisumsativum*	PsCHS3	G box	unknown	CACGTG
RSP01533	*Pisumsativum*	PsCHS4	G box	unknown	CACGTG
RSP01534	*Pisumsativum*	PsCHS5	G-box	unknown	CACGTG
RSP01690	*Arabidopsis thaliana*	AOX1a	CARE H	unknown nuclear	GTCATC
RSP01692	*Arabidopsis thaliana*	AOX1a	CARE J	unknown nuclear	TTCGATCA
RSP01697	*Arabidopsis thaliana*	NDB2	CARE H	unknown nuclear	GTCATC
RSP01699	*Arabidopsis thaliana*	At2g21640/UP	CARE B1	unknown nuclear	ACGTGAT
RSP01700	*Arabidopsis thaliana*	At3g50930/BC	CARE H	unknown nuclear	GTCATC
RSP01749	*Arabidopsis thaliana*	COX5b-1	G-box	AREB2/ABF4	CACGTG
RSP01770	*Arabidopsis thaliana*	Synthetic	G box	CIB1	CACGTG
RSP01781	*Triticumaestivum*	VRN1	A/C motif	TaFDL2	TACGTC
RSP01782	*Triticumaestivum*	VRN1	G-box	TaFDL2	CACGTG
RSP01819	*Nicotianatabacum*	Pr-1a	WK box	NtWRKY12	TTTTCCAC
RSP01821	*Nicotianatabacum*	Pr-1a	MBSII box	Myb1	GTTTGGT
RSP01830	*Oryzasativa*	KAR	GAMYB BS	GAMYB	TAACTGAAT
RSP01832	*Arabidopsis thaliana*	Cytc-2	G-box	AREB2/ABF4;	CACGTG
RSP01857	*Arabidopsis thaliana*	GAI	G-box	PIL5	CACGTG
RSP01859	*Arabidopsis thaliana*	RGA/RGA1	G-box	PIL5	CACGTG
RSP01881	*Arabidopsis thaliana*	Synthetic	Athb-1 BS	Athb-1	CAATTATTG
RSP01955	*Arabidopsis thaliana*	MT2B	G-box	unknown nuclear	CACGTG
RSP01977	*Arabidopsis thaliana*	FLS2	LS2 EIN3/EIL1	EIN3/EIL1	ATGCATCT
RSP01978	*Arabidopsis thaliana*	FLS2	FLS2	EIN3/EIL1	ATGTATGT
RSP02018	*Eucalyptus gunnii*	EgCAD2	MYBa	MYB	CACCTACC
RSP02020	*Eucalyptus gunnii*	EgCCR	MYB	MYB	GGTAGGTG
RSP02037	*Medicagotruncatula*	LBD1	HB1	HB1	CAATAATTG
RSP02067	*Glycine max*	CHS8	GmMYB176	GmMYB176	GCGTGAAAATAT
RSP02068	*Triticumaestivum*	synthetic	MYB/W	MYB/W	GAAAATATAGTT
RSP02070	*Petunia hybrid*	Synthetic	MYB/Ph&W	MYB/Ph&W	AGCGTGAAAATA

RSP02071	*Hordeum vulgare*	Synthetic	MYB-like/WT	MYB-like/WT	CCACCGTCTACT
RSP02072	*Hordeum vulgare*	Synthetic	MYB-like1	MYB-like1	TAGTTTGTGACC
RSP02120	*Arabidopsis thaliana*	ech-42	LFY	LFY	CCAATGT
RSP02140	*Zea mays*	F5H	MBS-1	ZmMYB31	TCCACC
RSP02162	*Capsicum annuum*	CaPR-1	G-box	CabZIP1	CACGTG
RSP02169	*Glycine max*	CHS8	Fp2	unknown nuclear	ACTTTAATT
RSP02171	*Glycine max*	CHS8	Fp5/II	unknown nuclear	ATAGGTCACAAA
RSP02172	*Glycine max*	CHS8	Fp6/II	unknown nuclear	ACTCATA
RSP02173	*Glycine max*	CHS8	Fp7/II	unknown nuclear	GCGGGCATATGC
RSP02174	*Glycine max*	CHS8	Fp9/III	unknown nuclear	GGGGTAGAGGC
RSP02175	*Glycine max*	CHS8	Fp11/III	unknown nuclear	GAGAATAC
RSP02176	*Glycine max*	CHS8	Fp12/III	unknown nuclear	ATTTTTTTAT
RSP02177	*Glycine max*	CHS8	Fp13/III	unknown nuclear	TTCCATCATCGA
RSP02178	*Glycine max*	CHS8	Fp16/IV	unknown nuclear	TGAATAGCGTGA
RSP02210	*Solanum melongena*	SmPT5	P1BS (SmPT5)	PHR1 (P1)	GCATATGC
RSP02213	*Nicotianatabacum*	NtPT5	P1BS (NtPT5)	PHR1 (P1)	GCATATGC
RSP02246	*Arabidopsis thaliana*	RPP8	W-RPP8-A/B/C	WRKY transcription	GTTGACTT
RSP02254	*Arabidopsis thaliana*	Ep5C	H box	unknown nuclear	CCTACC
RSP02315	*Nicotianatabacum*	MATE1	G-box	NtMYC2	GCACGTGA
RSP02324	*Arabidopsis thaliana*	JAZ2	G-box	MYC2; MYC3;	CACGTG
RSP02329	*Populustrichocarpa*	Synthetic	SMRE 3	PtrMYB3;	ACCAAAC
RSP02334	*Populustrichocarpa*	Synthetic	SMRE 8	PtrMYB3;	ACCTACC
RSP02401	*Arabidopsis thaliana*	COX5b	G-box	AREB2/ABF4	CACGTG
RSP02426	*Arabidopsis thaliana*	IAA19	G-box	PIF4	CACGT
RSP02427	*Arabidopsis thaliana*	IAA29	G-box	PIF4	CACGTG
RSP02443	*Arabidopsis thaliana*	ANAC019	G-box	CBF1; CBF2	CACGTG
RSP02453	*Arabidopsis thaliana*	ANAC055	GBF1/GBF2	GBF1; GBF2	CACGTG
RSP02469	*Arabidopsis thaliana*	PIL1	G-box	PIF3	CACGTG
RSP02470	*Arabidopsis thaliana*	PHYB	G-box	PIF3	CACGTG
RSP02471	*Arabidopsis thaliana*	RGA1	G-box	PIF3	CACGTG
RSP02475	*Arabidopsis thaliana*	ATHB-2	G-box	PIF3	CACGTG
RSP02478	*Arabidopsis thaliana*	HAT3	G-box	PIF3	CACGTG

RSP02480	*Arabidopsis thaliana*	WAG2	G-box	PIF3	CACGTG
RSP02482	*Arabidopsis thaliana*	RCI2B	G-box	PIF3	CACGTG
RSP02484	*Arabidopsis thaliana*	AT1G19310	G-box	PIF3	CACGTG
RSP02486	*Arabidopsis thaliana*	AT4G03010	G-box	PIF3	CACGTG
RSP02492	*Arabidopsis thaliana*	s	AC-I (5)	AtMYB61	ACCAAC
RSP02564	*Arabidopsis thaliana*	TAT1	G-box	JAM1; MYC2	CACGTG
RSP02587	*Piceaglauca [Moench]*	Pg4CL	AC-I	PGmYB8	ACCTACC
	Piceaglauca [Moench]	PgDHS2	AC-I	PGmYB8	ACCTACC
RSP02612	*Arabidopsis thaliana*	At5g64940	G-box	bHLH122	CACGTG
RSP02614	*Arabidopsis thaliana*	ERF6	G-box	bHLH122	CACGTG
RSP02647	*Arabidopsis thaliana*	ORA59	G-box	MYC2	CACGTG
RSP02680	*Musa acuminata*	MaNAC1	MYC RE1	MaICE1	CACGTG



Appendix 4 The expression level of GmMYB*176* in GmMYB176-Si, GmMYB176-OE and control hairy roots.

The transcript abundance of *GmMYB176* was analysed by qPCR in 10 biological replicates of GmMYB176-Si and controls, and in 9 biological replicates of GmMYB176-OE samples. On average, the *GmMYB176* expression levels were significantly altered in GmMYB176-Si and GmMYB176-OE replicates. Four replicates from each GmMYB176-Si, GmMYB176-OE and control were chosen for RNAseq experiment.

Appendix 5 List of differentially produced metabolites identified by LC-MS/MS in GmMYB176-Si and GmMYB176-OE hairy roots compared to the controls.

	mz	rt	P value	log2 FC	Sample (vs Control)	ESI
1	905.6764465	2.6257495	0.009826846	-3.168083831	GmMYB176-Si	positive
2	575.5017243	4.1766934	0.009826846	-2.576334769	GmMYB176-OE	positive
3	440.3567616	4.1770342	0.026173129	-2.12638696	GmMYB176-Si	positive
4	163.0383851	0.544963182	0.026173129	-2.212202066	GmMYB176-Si	positive
5	435.0927455	2.6381056	0.026173129	1.472828653	GmMYB176-OE	negative
6	176.0912844	1.553851467	0.026173129	-2.867654231	GmMYB176-Si	positive
7	627.4305331	2.873569833	0.026173129	-2.138438047	GmMYB176-Si	positive
8	650.4158986	2.872079167	0.026173129	-2.568568052	GmMYB176-Si	positive
9	1122.804613	5.618056933	0.026173129	-2.703965212	GmMYB176-Si	positive
10	725.5129543	5.53643815	0.026173129	-1.986972247	GmMYB176-Si	positive
11	133.0493266	0.791747867	0.026173129	1.398347935	GmMYB176-OE	negative
12	542.4242755	4.3634812	0.026173129	-2.747180265	GmMYB176-Si	positive
13	112.9842595	0.628470033	0.026173129	-1.667004932	GmMYB176-OE	negative
14	575.5017243	4.1766934	0.026173129	-1.769922806	GmMYB176-Si	positive
15	570.9207489	5.619310533	0.026173129	-2.031905193	GmMYB176-Si	positive
16	594.4915969	5.771200033	0.026173129	-2.648043468	GmMYB176-Si	positive
17	556.4038888	4.63845155	0.026173129	-2.052734177	GmMYB176-Si	positive
18	574.391949	4.456709633	0.026173129	-2.107911447	GmMYB176-Si	positive
19	640.430378	4.413162433	0.026173129	-2.152154936	GmMYB176-Si	positive
20	102.0550663	0.494602943	0.026173129	-2.795887214	GmMYB176-Si	positive
21	1184.840963	5.541501433	0.026173129	-2.143429311	GmMYB176-Si	positive
22	196.059992	2.720047367	0.026173129	-2.429281981	GmMYB176-Si	positive
23	491.8738142	5.605198567	0.026173129	-2.447158916	GmMYB176-Si	positive
24	640.9323287	4.409722417	0.026173129	-2.127638072	GmMYB176-Si	positive
25	688.5177803	4.537834483	0.026173129	-1.947595546	GmMYB176-Si	positive
26	1079.782416	5.6336028	0.026173129	-2.185029416	GmMYB176-Si	positive
27	118.0651108	2.191369767	0.026173129	-1.190570716	GmMYB176-Si	positive
28	565.3211784	2.4742222	0.026173129	-2.42621506	GmMYB176-OE	negative

29	726.5695875	5.7382551	0.026173129	-2.549739818	GmMYB176-Si	positive
30	506.4398061	5.777374817	0.026173129	-2.410996594	GmMYB176-Si	positive
31	703.4997895	5.546063183	0.026173129	-2.309350238	GmMYB176-Si	positive
32	120.0806844	2.2412213	0.026173129	-2.661056437	GmMYB176-Si	positive
33	702.9984042	5.546063183	0.026173129	-2.483196997	GmMYB176-Si	positive
34	302.1953861	3.3292887	0.026173129	-2.914827948	GmMYB176-Si	positive
35	355.1738406	3.757660067	0.026173129	-1.916928613	GmMYB176-Si	positive
36	322.2090701	2.325851483	0.026173129	-2.551272418	GmMYB176-Si	positive
37	132.0804974	3.7402256	0.026173129	-2.085897705	GmMYB176-Si	positive
38	724.5096132	5.5383673	0.026173129	-2.495555777	GmMYB176-Si	positive
39	906.6797801	2.62560535	0.026173129	-2.133568444	GmMYB176-Si	positive
40	713.4734135	2.57449135	0.026173129	-1.466303563	GmMYB176-Si	negative
41	682.543651	5.7516621	0.026173129	-2.456974862	GmMYB176-Si	positive
42	181.1217304	3.740732733	0.026173129	-1.932301328	GmMYB176-Si	positive
43	628.4342823	2.8705885	0.026173129	-2.08288909	GmMYB176-Si	positive
44	662.4434505	4.3999343	0.026173129	-2.038943912	GmMYB176-Si	positive
45	314.2190378	2.870276967	0.026173129	-1.613728183	GmMYB176-Si	positive
46	991.7290257	5.665802767	0.026173129	-2.677251181	GmMYB176-Si	positive
47	215.1060485	3.721645167	0.026173129	-1.464622112	GmMYB176-Si	positive
48	638.5175763	5.757938233	0.026173129	-2.42172896	GmMYB176-Si	positive
49	526.8947048	5.650398933	0.026173129	-3.198906435	GmMYB176-Si	positive
50	636.4580454	5.5808563	0.026173129	-1.919661637	GmMYB176-Si	positive
51	1097.792168	5.56127215	0.026173129	-2.150032909	GmMYB176-Si	positive
52	180.10135	2.242383483	0.026173129	-2.665037486	GmMYB176-Si	positive
53	770.5955773	5.725231233	0.026173129	-2.655559901	GmMYB176-Si	positive
54	755.5279296	5.455351233	0.026173129	-2.092307341	GmMYB176-Si	positive
55	135.0801198	3.7399755	0.026173129	-1.768843782	GmMYB176-Si	positive
56	569.4795513	5.712663217	0.026173129	-2.039210303	GmMYB176-Si	positive
57	163.0748437	3.74088465	0.026173129	-1.750132529	GmMYB176-Si	positive
58	107.0856061	3.7402256	0.026173129	-1.85350385	GmMYB176-Si	positive
59	711.5020172	5.4744835	0.026173129	-2.189691044	GmMYB176-Si	positive
60	325.1842061	4.817634633	0.026173129	-1.034572334	GmMYB176-Si	negative

61	114.0913335	2.2880056	0.026173129	-1.878060609	GmMYB176-Si	positive
62	712.0037103	5.4745121	0.026173129	-2.171257691	GmMYB176-Si	positive
63	243.100861	3.72094755	0.026173129	-1.949324479	GmMYB176-Si	positive
64	592.43209	5.608091033	0.026173129	-2.22929971	GmMYB176-Si	positive
65	725.0114092	5.5383673	0.026173129	-2.431187335	GmMYB176-Si	positive
66	593.435307	5.608505033	0.026173129	-2.166033498	GmMYB176-Si	positive
67	662.9457099	4.3999665	0.026173129	-1.979213349	GmMYB176-Si	positive
68	680.4836829	5.553948533	0.026173129	-2.306489823	GmMYB176-Si	positive
69	680.985519	5.5587396	0.026173129	-2.471435664	GmMYB176-Si	positive
70	990.7257099	5.665690033	0.026173129	-2.698312934	GmMYB176-Si	positive
71	858.647419	5.7026032	0.026173129	-2.721106545	GmMYB176-Si	positive
72	270.1407823	3.741781233	0.026173129	-1.791637141	GmMYB176-Si	positive
73	550.4657531	5.7726573	0.026173129	-2.475571567	GmMYB176-Si	positive
74	727.5729369	5.7382551	0.026173129	-2.413095939	GmMYB176-Si	positive
75	301.1635559	3.745058267	0.026173129	-1.871182455	GmMYB176-Si	positive
76	903.6770146	5.690161267	0.026173129	-2.640530216	GmMYB176-Si	positive
77	255.1218659	3.772731017	0.026173129	-1.9076381	GmMYB176-Si	positive
78	644.4557094	4.579722767	0.026173129	-2.072479317	GmMYB176-Si	positive
79	209.1166041	3.741036567	0.026173129	-1.759056518	GmMYB176-Si	positive
80	188.0706788	2.830873367	0.026173129	2.531285685	GmMYB176-OE	negative
81	548.4061431	5.6369555	0.026173129	-2.257762146	GmMYB176-Si	positive
82	280.1621986	2.093940733	0.026173129	-1.284221208	GmMYB176-Si	positive
83	335.278079	4.195070167	0.026173129	-1.778713755	GmMYB176-Si	positive
84	683.5468544	5.752605933	0.026173129	-2.370459072	GmMYB176-Si	positive
85	177.1268484	3.7785737	0.026173129	-1.728195184	GmMYB176-Si	positive
86	269.1374725	3.744061467	0.026173129	-1.778863403	GmMYB176-Si	positive
87	379.304071	4.1858561	0.026173129	-1.672842206	GmMYB176-Si	positive
88	702.4969219	5.542589633	0.026173129	-2.546188777	GmMYB176-Si	positive
89	244.1041957	3.7208423	0.026173129	-1.924849428	GmMYB176-Si	positive
90	946.6996854	5.675356483	0.026173129	-2.711276149	GmMYB176-Si	positive
91	902.6736396	5.688365117	0.026173129	-2.590537966	GmMYB176-Si	positive
92	571.4221759	5.619310533	0.026173129	-2.26025762	GmMYB176-Si	positive

93	825.4640292	3.6021271	0.026173129	1.636811445	GmMYB176-OE	negative
94	548.9078759	5.636985933	0.026173129	-2.22153816	GmMYB176-Si	positive
95	769.0378941	5.522415533	0.026173129	-2.487404211	GmMYB176-Si	positive
96	859.6508788	5.7036941	0.026173129	-2.647123054	GmMYB176-Si	positive
97	1140.814928	5.551190667	0.026173129	-2.13660155	GmMYB176-Si	positive
98	684.4562629	4.380882583	0.026173129	-1.934335612	GmMYB176-Si	positive
99	323.2035223	4.078754367	0.026173129	-1.898830164	GmMYB176-Si	positive
100	1096.788794	5.560900467	0.026173129	-2.140788936	GmMYB176-Si	positive
101	387.2477734	2.654671967	0.026173129	-1.955662416	GmMYB176-Si	positive
102	194.1169328	0.513964645	0.026173129	-2.549356272	GmMYB176-Si	positive
103	678.4759237	4.3634812	0.026173129	-3.13523039	GmMYB176-Si	positive
104	733.5149295	5.464573383	0.026173129	-2.086348352	GmMYB176-Si	positive
105	814.6212736	5.714029	0.026173129	-2.306982438	GmMYB176-Si	positive
106	734.0166334	5.4648436	0.02621709	-2.163198354	GmMYB176-Si	positive
107	349.2934914	4.4082335	0.02621709	-2.044012945	GmMYB176-Si	positive
108	116.0341551	0.493363677	0.02621709	-2.755504102	GmMYB176-Si	positive
109	283.2812614	4.825304533	0.02621709	-1.671533603	GmMYB176-Si	positive
110	826.4673925	3.60252185	0.02621709	1.642721281	GmMYB176-OE	negative
111	1002.53601	3.398146933	0.026415335	2.16757948	GmMYB176-OE	negative
112	407.2454806	3.774411267	0.026415335	-2.007027385	GmMYB176-Si	positive
113	505.3829152	5.6622508	0.026415335	-2.058584152	GmMYB176-Si	positive
114	693.4215014	3.810947767	0.026415335	1.688134624	GmMYB176-OE	negative
115	318.1900595	3.7402256	0.026415335	-1.867679068	GmMYB176-Si	positive
116	624.9543159	5.522512333	0.026415335	-1.97066273	GmMYB176-Si	positive
117	357.1894765	3.7841287	0.026415335	-2.021668253	GmMYB176-Si	positive
118	371.2182187	2.4663322	0.026415335	-1.115871839	GmMYB176-OE	negative
119	132.8664328	2.470145933	0.026415335	-1.861546233	GmMYB176-Si	negative
120	276.1302687	3.7209714	0.026415335	-1.973884292	GmMYB176-Si	positive
121	331.2534418	2.5697501	0.026415335	-1.560356618	GmMYB176-Si	positive
122	746.5227552	5.52688715	0.026415335	-2.441189136	GmMYB176-Si	positive
123	694.4247953	3.810049733	0.026415335	1.70676869	GmMYB176-OE	negative
124	965.713704	5.599857517	0.026415335	-2.024023997	GmMYB176-Si	positive

125	242.145965	3.7402256	0.026415335	-1.791416291	GmMYB176-Si	positive
126	712.5052513	5.4744549	0.026415335	-1.997965692	GmMYB176-Si	positive
127	788.6058395	5.657698	0.026415335	-1.961700066	GmMYB176-Si	positive
128	144.1379879	3.0289841	0.026415335	-2.341003413	GmMYB176-Si	positive
129	778.0430148	5.45302085	0.026428324	-1.972807003	GmMYB176-Si	positive
130	920.6839298	5.618120367	0.026428324	-1.981821561	GmMYB176-Si	positive
131	408.3387628	4.6111461	0.026428324	-2.12904784	GmMYB176-Si	positive
132	327.1790818	3.759237117	0.026428324	-1.73952743	GmMYB176-Si	positive
133	658.972415	5.569791633	0.026428324	-2.341770244	GmMYB176-Si	positive
134	158.1535866	3.45694095	0.026428324	-1.061673069	GmMYB176-Si	positive
135	130.1223752	2.671697683	0.026428324	-2.295887841	GmMYB176-Si	positive
136	275.1268375	3.721308283	0.02651415	-1.935622696	GmMYB176-Si	positive
137	524.4502017	5.714971867	0.02651415	-1.988826177	GmMYB176-Si	positive
138	406.2421895	3.774411267	0.02651415	-1.986106092	GmMYB176-Si	positive
139	921.6874266	5.6191109	0.02651415	-1.745876584	GmMYB176-Si	positive
140	525.4537186	5.715610333	0.02651415	-1.961794122	GmMYB176-Si	positive
141	130.049606	0.473503318	0.02651415	-2.47078817	GmMYB176-Si	positive
142	404.2265187	3.755217633	0.02651415	-1.89079927	GmMYB176-Si	positive
143	568.4761452	5.713778667	0.02651415	-1.928549669	GmMYB176-Si	positive
144	612.5020789	5.704362833	0.02651415	-1.919275075	GmMYB176-Si	positive
145	1053.765817	5.57482465	0.02651415	-2.052850497	GmMYB176-Si	positive
146	681.5153289	2.5700409	0.02651415	-1.749430243	GmMYB176-Si	positive
147	947.7031247	5.675356483	0.02651415	-2.580442968	GmMYB176-Si	positive
148	320.2865168	4.6205924	0.02651415	-2.245742155	GmMYB176-Si	positive
149	689.990463	5.4841915	0.02651415	-2.011480025	GmMYB176-Si	positive
150	656.5280268	5.6947978	0.02651415	-1.891055157	GmMYB176-Si	positive
151	659.4741008	5.570234317	0.02651415	-2.191917943	GmMYB176-Si	positive
152	1052.762704	5.5731736	0.02651415	-2.097728213	GmMYB176-Si	positive
153	700.5540198	5.6851124	0.02651415	-1.914667607	GmMYB176-Si	positive
154	298.3460176	3.571549733	0.02651415	-1.723694569	GmMYB176-Si	positive
155	789.6093119	5.6580089	0.02651415	-1.949478775	GmMYB176-Si	positive
156	689.4888272	5.4841915	0.02651415	-2.002326535	GmMYB176-Si	positive

157	768.5356215	5.5226018	0.02651415	-2.400750987	GmMYB176-Si	positive
158	148.0599901	0.494602943	0.02651415	-2.881214195	GmMYB176-Si	positive
159	451.3611257	4.6021219	0.02651415	-2.381982813	GmMYB176-Si	positive
160	1008.736491	5.589413233	0.02651415	-2.059721083	GmMYB176-Si	positive
161	109.0744272	2.031932133	0.02651415	-1.998956153	GmMYB176-Si	positive
162	701.5575796	5.684673267	0.02651415	-1.908213761	GmMYB176-Si	positive
163	173.9843502	0.551122963	0.02651415	-1.886959582	GmMYB176-Si	positive
164	679.5092785	2.569965967	0.02651415	-1.705562713	GmMYB176-Si	positive
165	680.5122884	2.5702311	0.02651415	-1.710452022	GmMYB176-Si	positive
166	389.2537372	2.653886483	0.02651415	-1.705792006	GmMYB176-Si	positive
167	744.5799768	5.675094267	0.026518351	-1.929369759	GmMYB176-Si	positive
168	373.1288986	3.4194301	0.026518351	1.847517095	GmMYB176-OE	negative
169	618.417722	4.4273292	0.026518351	-1.905179364	GmMYB176-Si	positive
170	964.7102504	5.60110345	0.026518351	-2.024235041	GmMYB176-Si	positive
171	1001.532602	3.3980771	0.026649227	2.105110115	GmMYB176-OE	negative
172	1009.739987	5.589438117	0.026649227	-2.049235023	GmMYB176-Si	positive
173	526.3930959	5.6481396	0.026790828	-2.290202977	GmMYB176-Si	positive
174	247.1320776	3.726244433	0.026790828	-1.894648928	GmMYB176-Si	positive
175	498.3368519	2.4845829	0.026790828	-1.271761051	GmMYB176-OE	negative
176	690.4920224	5.48424625	0.026790828	-1.970550446	GmMYB176-Si	positive
177	226.179592	3.08775855	0.026790828	-3.191333053	GmMYB176-OE	positive
178	491.2972059	4.533723733	0.026790828	-2.565908937	GmMYB176-Si	positive
179	504.3799246	5.6657456	0.026816877	-2.224170761	GmMYB176-Si	positive
180	198.1483894	2.84065065	0.026816877	-1.834984884	GmMYB176-Si	positive
181	1078.778663	5.636985933	0.026816877	-2.603673621	GmMYB176-Si	positive
182	413.1084374	2.609868367	0.026816877	1.406628381	GmMYB176-OE	negative
183	512.4139886	4.5930309	0.026816877	-2.343411636	GmMYB176-Si	positive
184	312.2271371	2.40612905	0.026816877	-2.826414696	GmMYB176-Si	positive
185	409.1735409	3.640553283	0.026816877	-1.785598013	GmMYB176-Si	positive
186	173.1091903	2.03262695	0.026816877	-1.869602898	GmMYB176-Si	positive
187	971.5213655	3.503293267	0.026816877	1.689052766	GmMYB176-OE	negative
188	682.5183721	2.5698397	0.026816877	-1.711145816	GmMYB176-Si	positive

189	410.3460884	4.399821	0.026816877	-2.04331179	GmMYB176-Si	positive
190	613.505499	5.705680167	0.026816877	-1.920443285	GmMYB176-Si	positive
191	341.2650921	4.620633967	0.027004448	-2.0373153	GmMYB176-Si	positive
192	369.2374265	2.65382	0.027004448	-1.673536936	GmMYB176-Si	positive
193	552.3787157	4.468695017	0.027050935	-1.856124535	GmMYB176-Si	positive
194	134.8640452	4.163994717	0.027163484	-1.449384925	GmMYB176-OE	negative
195	724.5082313	4.350463133	0.027221486	-2.728992522	GmMYB176-Si	positive
196	174.9549785	2.5803663	0.027258686	-1.297894319	GmMYB176-Si	negative
197	197.1166735	3.58201285	0.027258686	-1.677451818	GmMYB176-Si	positive
198	717.4648335	2.5699188	0.027258686	-1.577785655	GmMYB176-Si	positive
199	475.3237738	4.258713467	0.027258686	-1.977283036	GmMYB176-Si	positive
200	734.5179463	5.4648436	0.027258686	-1.788455012	GmMYB176-Si	positive
201	512.3780886	4.669289467	0.027258686	-2.159185087	GmMYB176-Si	positive
202	261.1113318	3.742803267	0.027358573	-1.989400619	GmMYB176-Si	positive
203	605.1819015	3.549884283	0.027379515	-2.435285863	GmMYB176-Si	positive
204	667.9775802	5.49422555	0.027384162	-1.957985321	GmMYB176-Si	positive
205	291.2520033	4.1954773	0.027384162	-1.856295294	GmMYB176-Si	positive
206	297.1686614	3.784088367	0.027481995	-1.810773487	GmMYB176-Si	positive
207	192.1588587	0.668192213	0.027481995	3.231790004	GmMYB176-OE	positive
208	548.4042395	4.3999343	0.027481995	-2.457715074	GmMYB176-Si	positive
209	146.1172866	2.1465509	0.027481995	-2.174090619	GmMYB176-Si	positive
210	352.3046442	4.19461535	0.027481995	-1.659594278	GmMYB176-Si	positive
211	498.3981566	4.379441833	0.027481995	-1.993978837	GmMYB176-Si	positive
212	396.3308692	4.18591025	0.027481995	-1.616722447	GmMYB176-Si	positive
213	624.4530576	5.522282167	0.027481995	-1.929558979	GmMYB176-Si	positive
214	644.4915818	4.5541427	0.027481995	-1.908500946	GmMYB176-Si	positive
215	341.7628378	2.569720467	0.027481995	-1.464341964	GmMYB176-Si	positive
216	454.3722233	4.39038295	0.027481995	-2.010834782	GmMYB176-Si	positive
217	157.0491452	1.918936467	0.027481995	-2.477365215	GmMYB176-Si	positive
218	319.2831749	4.621035933	0.027481995	-2.116655397	GmMYB176-Si	positive
219	183.1488278	2.799785683	0.027481995	-1.851705034	GmMYB176-Si	positive
220	246.1480609	2.475467733	0.027481995	-1.469418244	GmMYB176-Si	positive

221	695.4274642	3.810049733	0.027859154	1.690045737	GmMYB176-OE	negative
222	276.7269164	2.569735283	0.027949713	-1.598990718	GmMYB176-Si	positive
223	592.9334681	5.6085232	0.027949713	-2.109149011	GmMYB176-Si	positive
224	614.9465233	5.598155133	0.027949713	-2.122611373	GmMYB176-Si	positive
225	268.082544	3.173008067	0.027949713	1.448828266	GmMYB176-OE	negative
226	697.519715	2.373767467	0.027949713	-1.955644965	GmMYB176-Si	positive
227	508.1854324	3.856897367	0.027949713	-1.842409949	GmMYB176-Si	positive
228	625.9274191	5.522091483	0.027949713	-1.60358852	GmMYB176-Si	positive
229	645.4629504	5.512209017	0.027949713	-1.902152703	GmMYB176-Si	positive
230	596.9065619	4.43758215	0.027949713	-1.752450151	GmMYB176-Si	positive
231	747.0244272	5.52688715	0.027949713	-2.246528535	GmMYB176-Si	positive
232	369.2374903	2.8844934	0.027965065	-1.732351655	GmMYB176-Si	positive
233	323.1454366	3.745557033	0.02797843	-1.314078699	GmMYB176-Si	positive
234	265.1426285	3.77906675	0.02797843	-1.916370209	GmMYB176-Si	positive
235	667.4759496	5.4939009	0.02797843	-1.909263517	GmMYB176-Si	positive
236	298.2564824	3.367288217	0.02797843	-2.325387703	GmMYB176-Si	positive
237	238.0914848	1.91184905	0.02797843	-2.614496095	GmMYB176-Si	positive
238	254.0665991	2.833200733	0.02797843	1.236756018	GmMYB176-OE	negative
239	349.2635723	2.369948633	0.02797843	-1.822327082	GmMYB176-Si	positive
240	623.4500817	5.522415533	0.02797843	-1.858022007	GmMYB176-Si	positive
241	130.158748	2.2692669	0.02797843	-2.009402224	GmMYB176-Si	positive
242	227.1747135	2.261798567	0.02797843	-1.767880426	GmMYB176-Si	positive
243	340.2586386	2.569965967	0.028311	-1.469363075	GmMYB176-Si	positive
244	602.4401316	5.5323092	0.028311	-1.910706622	GmMYB176-Si	positive
245	233.2217832	1.8861698	0.028311	-1.892302958	GmMYB176-Si	positive
246	343.3056486	1.9966802	0.028311	-2.02456753	GmMYB176-Si	positive
247	174.9549785	2.5803663	0.028311	-1.630495401	GmMYB176-OE	negative
248	637.4611305	5.580173933	0.028311	-2.199450683	GmMYB176-Si	positive
249	623.9516642	5.522314867	0.028311	-1.853393702	GmMYB176-Si	positive
250	636.9595773	5.580173933	0.028311	-2.160805392	GmMYB176-Si	positive
251	329.1947219	3.785045633	0.028311	-1.855923668	GmMYB176-Si	positive
252	658.4706926	5.569761867	0.028311	-2.206206797	GmMYB176-Si	positive

253	658.5071825	4.78678395	0.028311	-2.12800923	GmMYB176-Si	positive
254	646.4660547	5.510941667	0.028311	-1.889980458	GmMYB176-Si	positive
255	229.1428044	3.581266533	0.028661215	-1.548187619	GmMYB176-Si	positive
256	340.7598906	2.578758133	0.028704117	-1.450105979	GmMYB176-Si	positive
257	601.9387172	5.533702433	0.028704117	-1.838925046	GmMYB176-Si	positive
258	877.6614605	5.6299839	0.028704117	-1.911403947	GmMYB176-Si	positive
259	363.3091348	4.617158333	0.028704117	-2.303220839	GmMYB176-Si	positive
260	291.1220137	3.4242188	0.028704117	-1.979880378	GmMYB176-Si	positive
261	359.2316983	2.57048525	0.028884329	-1.464740403	GmMYB176-Si	positive
262	815.6248605	5.713980533	0.028884329	-2.381807191	GmMYB176-Si	positive
263	645.9645706	5.512132733	0.028884329	-1.856869095	GmMYB176-Si	positive
264	424.361636	4.611216467	0.028884329	-2.315239523	GmMYB176-Si	positive
265	453.8435793	2.621807567	0.028884329	-1.594145934	GmMYB176-Si	positive
266	349.765188	2.369948633	0.028884329	-1.799175014	GmMYB176-Si	positive
267	575.5014838	3.7935551	0.028884329	-3.256407411	GmMYB176-OE	positive
268	476.3050626	2.3743399	0.028884329	-1.641330747	GmMYB176-Si	positive
269	649.412388	2.869914783	0.028884329	-2.036529769	GmMYB176-Si	positive
270	243.0468793	1.908977933	0.029313632	-2.251856555	GmMYB176-Si	positive
271	681.4871428	5.5574278	0.029313632	-2.004993613	GmMYB176-Si	positive
272	173.0780175	1.957918667	0.029337295	-2.112489371	GmMYB176-Si	positive
273	364.312469	4.618588667	0.029488733	-2.289737298	GmMYB176-Si	positive
274	513.4171938	4.5925469	0.029488733	-2.252228305	GmMYB176-Si	positive
275	614.4448326	5.592429867	0.029488733	-2.105351244	GmMYB176-Si	positive
276	136.0864889	5.838759033	0.029488733	1.303764988	GmMYB176-OE	positive
277	556.4398505	4.582226133	0.029488733	-2.079595942	GmMYB176-Si	positive
278	613.4150984	2.701672233	0.029488733	-1.989288716	GmMYB176-Si	positive
279	407.3351844	4.611216467	0.029488733	-2.33613766	GmMYB176-Si	positive
280	180.064933	2.654854233	0.029488733	-1.993076411	GmMYB176-Si	positive
281	393.245799	2.2430242	0.029488733	-1.78061799	GmMYB176-Si	positive
282	702.4941708	2.5741843	0.029488733	-1.532919003	GmMYB176-Si	positive
283	507.1821017	3.856897367	0.029488733	-1.795461588	GmMYB176-Si	positive
284	227.1747297	2.4867378	0.029488733	-1.579569713	GmMYB176-Si	positive

285	229.0852654	3.757660067	0.029488733	-1.788867859	GmMYB176-Si	positive
286	271.1529666	4.427900917	0.029488733	-1.967441857	GmMYB176-Si	positive
287	701.4909113	2.570636567	0.029488733	-1.526192758	GmMYB176-Si	positive
288	771.5989014	5.723602967	0.029555958	-2.417306852	GmMYB176-Si	positive
289	679.4792414	4.3659161	0.029555958	-2.674231306	GmMYB176-Si	positive
290	1034.752122	5.6467206	0.029555958	-2.560878988	GmMYB176-Si	positive
291	304.2989619	3.687339933	0.029555958	-1.82936428	GmMYB176-Si	positive
292	336.3096406	4.6207022	0.029555958	-2.170918265	GmMYB176-Si	positive
293	703.4973593	2.574050817	0.029555958	-1.552467431	GmMYB176-Si	positive
294	213.177195	2.475994667	0.029555958	-1.739021305	GmMYB176-Si	positive
295	419.3714678	5.713903833	0.029555958	-1.871290248	GmMYB176-Si	positive
296	369.1898244	3.913588667	0.029555958	-1.808195092	GmMYB176-Si	positive
297	619.4770691	3.1929447	0.029555958	-1.671295383	GmMYB176-Si	positive
298	332.329956	3.644079767	0.029555958	-1.794239414	GmMYB176-Si	positive
299	210.1482836	2.673383433	0.029556682	-1.780688272	GmMYB176-Si	positive
300	235.1070033	2.200920233	0.029562413	-3.441474292	GmMYB176-Si	positive
301	425.3650209	4.6111461	0.029576003	-2.338340422	GmMYB176-Si	positive
302	221.2006507	1.947076133	0.029789495	-1.944242854	GmMYB176-Si	positive
303	403.2690262	2.883681567	0.030238827	-1.64876844	GmMYB176-Si	positive
304	537.2940185	2.476821233	0.030391226	-1.809988912	GmMYB176-Si	positive
305	527.3964581	5.651546667	0.030391226	-2.013748899	GmMYB176-Si	positive
306	179.0339997	2.759489767	0.03046217	3.284263976	GmMYB176-OE	negative
307	469.3912941	4.6018714	0.030627004	-2.356843123	GmMYB176-Si	positive
308	745.5834736	5.673996867	0.030634335	-1.859073526	GmMYB176-Si	positive
309	418.2423151	3.913017333	0.030634335	-1.888961705	GmMYB176-Si	positive
310	523.1558875	3.856953967	0.030634335	-1.411851337	GmMYB176-Si	positive
311	432.2791272	2.3407544	0.030819367	-1.872160077	GmMYB176-Si	positive
312	411.1975533	3.775429033	0.031203223	-1.484463252	GmMYB176-Si	positive
313	241.1426332	3.7414089	0.031213392	-1.552566284	GmMYB176-Si	positive
314	410.1768315	3.639572767	0.031213392	-1.522820727	GmMYB176-Si	positive
315	557.4435047	4.582354233	0.031213392	-2.046557943	GmMYB176-Si	positive
316	280.2626805	3.940147733	0.031213392	-1.513781788	GmMYB176-Si	positive

317	468.3879808	4.603677567	0.031213392	-2.023042504	GmMYB176-Si	positive
318	198.0754586	2.0047134	0.031213392	-2.430785803	GmMYB176-Si	positive
319	668.4791585	5.4945502	0.032264002	-1.809923213	GmMYB176-Si	positive
320	373.2321129	2.4667265	0.032278786	-2.066332691	GmMYB176-Si	positive
321	499.3399534	2.4841965	0.032278786	-1.044422071	GmMYB176-OE	negative
322	471.352508	2.27071085	0.032278786	-1.618162229	GmMYB176-Si	positive
323	383.3087037	2.951761967	0.032278786	-2.24776511	GmMYB176-Si	positive
324	405.2258547	4.224347867	0.032278786	-2.028016384	GmMYB176-Si	positive
325	381.3387337	4.620787367	0.032329148	-2.334286298	GmMYB176-Si	positive
326	145.1332116	0.836668005	0.032329148	-1.745605095	GmMYB176-Si	positive
327	359.7334104	2.570115833	0.032329148	-1.439879948	GmMYB176-Si	positive
328	166.0858371	2.0249264	0.032329148	-2.950390167	GmMYB176-Si	positive
329	565.3211784	2.4742222	0.032329148	-1.545967708	GmMYB176-Si	negative
330	196.1327852	2.701168167	0.032382456	-1.815826969	GmMYB176-Si	positive
331	273.2237254	4.591937733	0.032432656	-2.438709451	GmMYB176-Si	positive
332	436.3981001	5.714029	0.032494908	-1.856641208	GmMYB176-Si	positive
333	274.2269355	4.591883933	0.032494908	-2.361415755	GmMYB176-Si	positive
334	243.0467617	2.236890917	0.03266082	-2.786538076	GmMYB176-Si	positive
335	367.7197158	2.577513933	0.032961214	-1.478164449	GmMYB176-Si	positive
336	102.0551049	0.78513735	0.032961214	-2.912885158	GmMYB176-Si	positive
337	480.4241204	5.718757133	0.032961214	-1.817589385	GmMYB176-Si	positive
338	779.457991	3.6021463	0.032961214	1.383551208	GmMYB176-OE	negative
339	377.2593695	2.691595167	0.032961214	-1.88220118	GmMYB176-Si	positive
340	370.2938825	4.42726285	0.032961214	-1.819276003	GmMYB176-Si	positive
341	274.1642683	2.769917867	0.033517697	-1.914826788	GmMYB176-Si	positive
342	927.6583499	2.626248033	0.033599738	-1.761650144	GmMYB176-Si	positive
343	120.0807065	2.0296342	0.033711353	-3.030551411	GmMYB176-Si	positive
344	401.2635607	2.8869755	0.033711353	-1.540712106	GmMYB176-Si	positive
345	329.1347793	3.7247609	0.033711353	-1.313150531	GmMYB176-Si	positive
346	191.1060541	3.7402256	0.033711353	-1.319479254	GmMYB176-Si	positive
347	389.1971964	3.636447767	0.033711353	-1.928575468	GmMYB176-Si	positive
348	367.2178504	2.577513933	0.033711353	-1.469584301	GmMYB176-Si	positive

349	580.9288541	5.54921465	0.034548138	-1.692666063	GmMYB176-Si	positive
350	133.0604588	0.436518797	0.034548138	-2.387147978	GmMYB176-Si	positive
351	615.4845663	4.79774665	0.034548138	-2.050116668	GmMYB176-Si	positive
352	374.2403242	2.654201167	0.034548138	-1.957320557	GmMYB176-Si	positive
353	209.1642263	2.261005333	0.034548138	-1.552339302	GmMYB176-Si	positive
354	614.4814553	4.7992311	0.034548138	-2.04772998	GmMYB176-Si	positive
355	331.1293475	3.640069633	0.034562881	-1.806456651	GmMYB176-Si	positive
356	129.018056	1.388040383	0.034562881	-2.896308033	GmMYB176-Si	negative
357	336.8586114	4.201642033	0.035205552	1.429128855	GmMYB176-Si	negative
358	756.0293883	5.4562847	0.035386291	-1.702288582	GmMYB176-Si	positive
359	579.424158	5.551140033	0.035386291	-1.720439094	GmMYB176-Si	positive
360	373.1763973	3.6554122	0.03594912	-2.23763817	GmMYB176-Si	positive
361	255.0699216	2.838793633	0.036352488	1.284756237	GmMYB176-OE	negative
362	137.0957175	3.582708533	0.036352488	-1.522156346	GmMYB176-Si	positive
363	616.4389801	2.8357953	0.036352488	-2.154285339	GmMYB176-Si	positive
364	600.4659511	4.5703138	0.036352488	-2.065522864	GmMYB176-Si	positive
365	571.458595	4.815596167	0.036352488	-2.035664578	GmMYB176-Si	positive
366	332.1326142	3.639572767	0.036352488	-1.889639588	GmMYB176-Si	positive
367	227.1270581	3.745058267	0.036352488	-1.639920334	GmMYB176-Si	positive
368	398.3237791	2.246104783	0.036352488	-1.850071224	GmMYB176-Si	positive
369	373.1288986	3.4194301	0.036352488	-2.755549329	GmMYB176-Si	negative
370	464.3335405	2.624752733	0.036352488	-1.558004134	GmMYB176-Si	positive
371	579.9258563	5.551150633	0.036352488	-1.724209565	GmMYB176-Si	positive
372	485.3455115	4.41381615	0.036352488	-2.152422974	GmMYB176-Si	positive
373	601.4696257	4.5725391	0.036962143	-1.870011964	GmMYB176-Si	positive
374	646.9675748	5.5069295	0.037020771	-1.730146992	GmMYB176-Si	positive
375	118.0861703	0.46978136	0.037020771	-1.510376204	GmMYB176-Si	positive
376	128.1067532	2.671424583	0.037020771	-2.328564293	GmMYB176-Si	positive
377	272.2576326	4.19821525	0.037020771	-2.721155289	GmMYB176-Si	positive
378	242.2834489	3.2332658	0.037020771	-1.664972928	GmMYB176-Si	positive
379	249.0273036	0.693764955	0.037020771	-2.067711404	GmMYB176-Si	positive
380	341.1949124	3.91691595	0.037020771	-1.661155804	GmMYB176-Si	positive

381	353.3878273	3.8549621	0.037020771	-3.617146592	GmMYB176-Si	positive
382	570.4190988	5.6227414	0.037020771	-2.04388183	GmMYB176-Si	positive
383	307.2346991	3.326882533	0.037414174	-1.385872666	GmMYB176-Si	positive
384	704.5002396	2.570115833	0.037584433	-1.463230017	GmMYB176-Si	positive
385	482.3670988	5.6753079	0.037584433	-1.763401769	GmMYB176-Si	positive
386	387.1918028	3.640069633	0.037584433	-1.768800198	GmMYB176-Si	positive
387	596.4045791	4.438286067	0.037584433	-1.84784695	GmMYB176-Si	positive
388	131.0350674	4.817137833	0.037584433	-1.118305584	GmMYB176-Si	negative
389	388.1949353	3.639572767	0.037718151	-1.765257116	GmMYB176-Si	positive
390	245.1852789	2.068652967	0.037718151	-2.149247845	GmMYB176-Si	positive
391	387.1917862	3.894761667	0.037718151	-2.003271102	GmMYB176-Si	positive
392	399.1636563	3.567498433	0.037895082	-2.052690751	GmMYB176-Si	positive
393	257.1376128	3.581467567	0.037954328	-1.536749973	GmMYB176-Si	positive
394	171.1487663	1.9667822	0.038630993	-1.780491292	GmMYB176-Si	positive
395	147.0760344	0.474634317	0.038731878	-2.572083761	GmMYB176-Si	positive
396	724.5050343	2.580428133	0.03883995	-1.115421656	GmMYB176-Si	negative
397	397.2934364	4.42342965	0.039130448	-1.992115305	GmMYB176-Si	positive
398	922.6906714	5.618652017	0.039130448	-1.704559341	GmMYB176-Si	positive
399	284.2844768	4.825167567	0.039130448	-1.32294751	GmMYB176-Si	positive
400	636.45607	4.3806815	0.039417886	-2.159777371	GmMYB176-Si	positive
401	528.4087669	4.150580183	0.039417886	-1.518019448	GmMYB176-Si	positive
402	397.1136666	2.736589233	0.039417886	1.496126535	GmMYB176-OE	negative
403	463.3011948	4.4191502	0.039550778	-2.529511876	GmMYB176-Si	positive
404	575.5014838	3.7935551	0.039550778	-2.56842792	GmMYB176-Si	positive
405	353.3878273	3.8549621	0.039550778	-7.421824918	GmMYB176-OE	positive
406	456.3508896	2.476821233	0.039799917	-1.504612069	GmMYB176-Si	positive
407	440.3588801	3.6832877	0.039799917	-1.649532109	GmMYB176-Si	positive
408	453.3421206	2.5041791	0.039889668	-1.503121458	GmMYB176-Si	positive
409	832.6319681	5.646671133	0.039993328	-1.797957601	GmMYB176-Si	positive
410	833.6355811	5.646621667	0.039993328	-1.746395903	GmMYB176-Si	positive
411	109.0285512	2.096595767	0.040119395	-2.234000931	GmMYB176-Si	positive
412	591.42724	4.39031475	0.040871414	-2.160961553	GmMYB176-Si	positive

413	394.3134435	3.4597078	0.040871414	-1.501689792	GmMYB176-Si	positive
414	484.3829061	4.1666291	0.040871414	-1.487524379	GmMYB176-Si	positive
415	504.8720109	5.589413233	0.04095748	-1.641886738	GmMYB176-Si	positive
416	549.4093915	5.636985933	0.041028105	-2.068859223	GmMYB176-Si	positive
417	590.4240468	4.3897095	0.041078269	-2.13023849	GmMYB176-Si	positive
418	156.1378899	2.6257495	0.041088592	-1.78385047	GmMYB176-Si	positive
419	125.0975111	2.0315462	0.041088592	-1.801620307	GmMYB176-Si	positive
420	468.193008	3.85510745	0.041242018	-2.0565411	GmMYB176-Si	positive
421	217.1040547	2.096495133	0.041242018	-1.911924365	GmMYB176-Si	positive
422	281.265834	4.565123367	0.041525804	-1.280566272	GmMYB176-Si	positive
423	162.0755348	0.520491883	0.041580451	-2.80614507	GmMYB176-Si	positive
424	117.1385685	0.394189493	0.041963905	-1.783740143	GmMYB176-Si	positive
425	156.13797	3.039463183	0.041963905	-2.243984367	GmMYB176-Si	positive
426	158.1534448	3.174833333	0.041963905	-2.010392152	GmMYB176-Si	positive
427	280.2625482	4.5653786	0.041963905	-1.259086438	GmMYB176-Si	positive
428	453.1740274	3.85510745	0.042074227	-2.05177435	GmMYB176-Si	positive
429	399.3316114	2.262956833	0.042144453	-1.688467243	GmMYB176-Si	positive
430	602.4237035	2.739339033	0.042179921	-2.757438482	GmMYB176-Si	positive
431	140.1066154	2.672975583	0.042195707	-2.026997361	GmMYB176-Si	positive
432	391.2749408	2.8262078	0.042217429	-1.762707367	GmMYB176-Si	positive
433	198.14828	2.626519267	0.042665936	-1.494247501	GmMYB176-Si	positive
434	195.1010367	3.7401216	0.042665936	-1.674494925	GmMYB176-Si	positive
435	795.3580316	3.6392749	0.042721052	-1.582707772	GmMYB176-Si	positive
436	467.1896661	3.855752733	0.042739563	-2.007740604	GmMYB176-Si	positive
437	657.5315412	5.69443185	0.042739563	-1.735511777	GmMYB176-Si	positive
438	635.4533028	4.3806815	0.042739563	-2.118300222	GmMYB176-Si	positive
439	372.2481751	2.351832267	0.042882947	-1.659282497	GmMYB176-Si	positive
440	152.0699741	0.503550713	0.0434365	-2.908613363	GmMYB176-Si	positive
441	615.4480399	5.592429867	0.0434365	-2.181898784	GmMYB176-Si	positive
442	154.1222684	2.932271133	0.04347487	-2.404276113	GmMYB176-Si	positive
443	285.7418533	2.150805133	0.043589194	-1.727783862	GmMYB176-Si	positive
444	504.8816556	5.665690033	0.043589194	-2.036416655	GmMYB176-Si	positive

445	272.2576326	4.19821525	0.043589194	-2.961911813	GmMYB176-OE	positive
446	227.6762688	2.476821233	0.043708167	-1.455005554	GmMYB176-Si	positive
447	580.427203	5.550620733	0.043923092	-1.661798861	GmMYB176-Si	positive
448	171.0259608	1.4245162	0.043925844	-2.693147353	GmMYB176-Si	positive
449	526.8852529	5.57225565	0.043925844	-1.549097661	GmMYB176-Si	positive
450	210.148326	2.884257817	0.044094711	-1.67202602	GmMYB176-Si	positive
451	181.0490109	0.525002117	0.044094711	-1.538397066	GmMYB176-Si	positive
452	142.1223931	2.8288869	0.044120411	-2.230969511	GmMYB176-Si	positive
453	441.3194589	4.419155433	0.044120411	-2.05427588	GmMYB176-Si	positive
454	285.2400994	2.155911333	0.044120411	-1.705620379	GmMYB176-Si	positive
455	600.429677	4.609097667	0.044182492	-1.98917751	GmMYB176-Si	positive
456	602.941949	5.5333377	0.044182492	-1.713176176	GmMYB176-Si	positive
457	702.5333126	4.7769926	0.044182492	-1.848692406	GmMYB176-Si	positive
458	546.3980057	4.40423375	0.044314284	-2.12205393	GmMYB176-Si	positive
459	325.3567344	3.6854717	0.044418798	-3.682284843	GmMYB176-Si	positive
460	691.4060567	4.000705767	0.044481928	1.462911338	GmMYB176-OE	negative
461	371.2640297	2.2430242	0.04468513	-1.518364307	GmMYB176-Si	positive
462	386.2372804	2.658134367	0.04468513	-1.185768733	GmMYB176-Si	negative
463	145.0968053	0.777151942	0.044702321	-2.013617057	GmMYB176-Si	positive
464	547.4013784	4.400155767	0.045577282	-2.110744298	GmMYB176-Si	positive
465	195.1221307	4.767006567	0.045577282	-1.980073715	GmMYB176-Si	positive
466	209.1643353	3.0978887	0.045577282	-1.5999772	GmMYB176-Si	positive
467	230.1745034	3.121863033	0.045652219	-1.235092553	GmMYB176-Si	positive
468	128.0704231	0.782312518	0.045738983	-1.856596461	GmMYB176-Si	positive
469	482.4033702	4.834849733	0.04580896	-2.153001755	GmMYB176-Si	positive
470	529.3713459	3.02129	0.046246869	-2.063955124	GmMYB176-Si	positive
471	223.1413195	2.272911533	0.046393621	-1.564702283	GmMYB176-Si	positive
472	325.3567344	3.6854717	0.046408471	-7.039345443	GmMYB176-OE	positive
473	229.0311296	0.68058687	0.046474371	-1.919823553	GmMYB176-Si	positive
474	473.1039296	2.382753367	0.046767816	-3.228928617	GmMYB176-Si	positive
475	170.0594843	2.369696967	0.047138784	-2.103598984	GmMYB176-Si	positive
476	170.0418247	0.48930324	0.047138784	-1.818397028	GmMYB176-Si	positive

477	386.1513077	3.85632505	0.047471099	-2.008730506	GmMYB176-Si	positive
478	385.1479996	3.856897367	0.047471099	-1.963611954	GmMYB176-Si	positive
479	244.2012422	1.962718967	0.047471099	-1.896119528	GmMYB176-Si	positive
480	143.0851351	4.0919908	0.047509423	-1.540158063	GmMYB176-Si	positive
481	255.1945113	3.6535714	0.047509423	-1.536581201	GmMYB176-Si	positive
482	416.2872387	3.010599517	0.047546844	-1.8007972	GmMYB176-Si	positive
483	431.1326635	2.729703417	0.047546844	-2.082055605	GmMYB176-Si	positive
484	642.4622692	3.184376667	0.047546844	-1.922180369	GmMYB176-Si	positive
485	280.2627204	3.3100631	0.047546844	-2.281804294	GmMYB176-Si	positive
486	641.45891	3.184376667	0.047546844	-1.90256046	GmMYB176-Si	positive
487	237.1840692	3.649331133	0.047546844	-1.542508244	GmMYB176-Si	positive
488	201.1954949	2.261005333	0.047546844	-1.8665409	GmMYB176-Si	positive
489	423.1976027	3.9220003	0.047546844	-1.292060103	GmMYB176-Si	positive
490	393.3100044	3.451251567	0.047546844	-1.375745755	GmMYB176-Si	positive
491	257.1739458	3.894761667	0.047546844	-1.418204326	GmMYB176-Si	positive
492	172.1691686	3.393381667	0.047768535	-2.020534368	GmMYB176-Si	positive
493	502.3719378	4.414469867	0.047768535	-2.17599933	GmMYB176-Si	positive
494	282.2781366	4.82657545	0.047768535	-1.225143846	GmMYB176-Si	positive
495	286.2561707	3.326882533	0.04798997	-1.579311635	GmMYB176-Si	positive
496	390.2715701	2.82668015	0.048141453	-1.784968495	GmMYB176-Si	positive
497	526.4294571	4.825276133	0.048196936	-2.030897784	GmMYB176-Si	positive
498	415.2919966	3.460347933	0.048217093	-1.520590968	GmMYB176-Si	positive
499	409.2296335	2.656647217	0.048217093	-1.90760473	GmMYB176-Si	positive
500	414.3198983	4.427900917	0.048217093	-1.908881071	GmMYB176-Si	positive
501	258.1693372	3.107754533	0.048217093	-1.61164087	GmMYB176-Si	positive
502	507.327351	4.409743817	0.048262543	-2.304279815	GmMYB176-Si	positive
503	190.0704904	1.9476884	0.048365922	-2.481228543	GmMYB176-Si	positive
504	345.2446097	4.4575593	0.048365922	-2.218042414	GmMYB176-Si	positive
505	410.2329855	2.654251483	0.048365922	-1.994633918	GmMYB176-Si	positive
506	212.0524667	1.935471833	0.048365922	-1.199494823	GmMYB176-Si	positive
507	242.1759525	3.093676917	0.048365922	-2.152639538	GmMYB176-OE	negative
508	458.3461084	4.421196733	0.048365922	-2.015336882	GmMYB176-Si	positive

509	399.2493528	4.205297833	0.048446464	-1.838207476	GmMYB176-Si	positive
510	212.1640148	3.04609865	0.048446464	-1.798129254	GmMYB176-Si	positive
511	558.4145325	5.560797217	0.048743648	-1.611725357	GmMYB176-Si	positive
512	156.0650456	2.2825571	0.050243532	-1.892080089	GmMYB176-Si	positive
513	299.2684896	3.477076333	0.050243532	-1.584189352	GmMYB176-Si	positive
514	226.179592	3.08775855	0.050243532	-2.001249533	GmMYB176-Si	positive
515	406.3028688	3.029757767	0.050298615	-1.731156589	GmMYB176-Si	positive
516	283.1743257	2.2160971	0.050462272	-1.956253828	GmMYB176-Si	positive
517	224.0546775	2.88535935	0.05046623	-1.750687813	GmMYB176-Si	positive
518	515.3552408	2.8749371	0.050509468	-1.915044447	GmMYB176-Si	positive
519	157.1412759	3.039571517	0.050509468	-2.240714227	GmMYB176-Si	positive
520	441.3712179	3.3879491	0.050509468	1.598446688	GmMYB176-OE	positive
521	603.9146424	5.534509	0.050548288	-1.37448607	GmMYB176-Si	positive
522	777.541281	5.454064033	0.050784207	-1.828808432	GmMYB176-Si	positive
523	501.3397839	2.797011933	0.050908464	-2.098777808	GmMYB176-Si	positive
524	209.1529635	3.6107174	0.050908464	-1.596573841	GmMYB176-Si	positive
525	491.3554195	2.777891233	0.050908464	-1.973015201	GmMYB176-Si	positive
526	503.3752992	4.417771	0.050908464	-2.072107051	GmMYB176-Si	positive
527	202.1794895	0.485235263	0.050908464	-2.233790277	GmMYB176-Si	positive
528	169.1218208	3.585617867	0.050908464	-1.329243358	GmMYB176-Si	positive
529	463.3240755	2.634268867	0.050908464	-2.050084985	GmMYB176-Si	positive
530	857.4802696	3.519479833	0.050908464	1.350922225	GmMYB176-OE	negative
531	163.0788537	0.785200733	0.050908464	-2.695939971	GmMYB176-Si	positive
532	371.2182187	2.4663322	0.050908464	-1.153133764	GmMYB176-Si	negative
533	253.1789448	3.567517117	0.050908464	-1.550275114	GmMYB176-Si	positive
534	275.1270708	3.363675533	0.051294929	-1.62675565	GmMYB176-Si	positive
535	294.1073556	2.030391683	0.051294929	-1.633694336	GmMYB176-Si	positive
536	173.1643683	1.984616467	0.051294929	-2.00164542	GmMYB176-Si	positive
537	241.1427779	3.31889135	0.051294929	-1.600715401	GmMYB176-Si	positive
538	285.2527489	3.319549033	0.051294929	-1.501695168	GmMYB176-Si	positive
539	179.059901	0.379353905	0.051294929	-2.512463456	GmMYB176-Si	positive
540	163.0383851	0.544963182	0.051294929	-1.28546472	GmMYB176-OE	positive

541	267.1711122	4.069056367	0.051294929	-1.222784657	GmMYB176-Si	positive
542	557.4112788	5.56085065	0.051294929	-1.596258184	GmMYB176-Si	positive
543	256.2625989	4.747651967	0.051294929	-1.222411733	GmMYB176-Si	positive
544	601.4372502	5.5383673	0.051294929	-1.732799073	GmMYB176-Si	positive
545	195.1486926	2.888827567	0.051294929	-1.570467837	GmMYB176-Si	positive
546	162.0755602	0.78513735	0.051294929	-2.585014562	GmMYB176-Si	positive
547	241.1903519	2.691595167	0.051294929	-2.039124832	GmMYB176-Si	positive
548	620.4731306	2.970668033	0.051294929	-2.149562253	GmMYB176-Si	positive
549	231.1372927	4.107854867	0.05168374	-1.481748949	GmMYB176-Si	positive
550	422.2522897	4.214421367	0.05168374	-1.681350204	GmMYB176-Si	positive
551	132.1016603	0.984227217	0.05168374	-2.465454529	GmMYB176-Si	positive
552	378.2717903	2.7516856	0.05168374	-1.799533871	GmMYB176-Si	positive
553	423.2450563	2.884284967	0.05168374	-1.468067614	GmMYB176-Si	positive
554	249.1565232	2.26140195	0.05168374	-1.380811366	GmMYB176-Si	positive
555	314.2678442	3.2033837	0.05168374	-1.701004411	GmMYB176-Si	positive
556	401.3363662	3.3290454	0.05168374	-2.834100951	GmMYB176-Si	positive
557	507.3892661	3.01004315	0.05168374	-2.032723508	GmMYB176-Si	positive
558	459.3492462	4.42342965	0.05168374	-2.020067784	GmMYB176-Si	positive
559	586.4502873	4.350735433	0.05168374	-1.765954004	GmMYB176-Si	positive
560	415.3231802	4.4271965	0.05168374	-1.642015999	GmMYB176-Si	positive
561	126.1007687	2.031932133	0.051688989	-1.739300206	GmMYB176-Si	positive
562	470.3684843	2.197041	0.051880248	-1.637212959	GmMYB176-Si	positive
563	490.305986	2.475140333	0.051880248	-1.192035107	GmMYB176-Si	negative
564	856.4775345	3.519479833	0.051880248	1.282228587	GmMYB176-OE	negative
565	548.8982974	5.553948533	0.051880248	-1.582193804	GmMYB176-Si	positive
566	581.9017576	5.551140033	0.051880248	-1.400855371	GmMYB176-Si	positive
567	213.1591207	2.378699933	0.051927536	-1.967631256	GmMYB176-Si	positive
568	855.4742292	3.52096265	0.052117198	1.278907555	GmMYB176-OE	negative
569	385.2909348	4.612550333	0.052273712	-2.144825959	GmMYB176-Si	positive
570	570.4554724	4.8156123	0.052273712	-1.937241194	GmMYB176-Si	positive
571	156.1378751	3.241960117	0.052273712	-2.277670011	GmMYB176-Si	positive
572	134.0444357	0.447043443	0.052273712	-2.674368582	GmMYB176-Si	positive

573	263.2360967	4.563005333	0.052273712	-1.187697768	GmMYB176-Si	positive
574	558.9158352	5.560800833	0.052692714	-1.497233639	GmMYB176-Si	positive
575	185.1643756	2.070484267	0.052692714	-2.102717955	GmMYB176-Si	positive
576	412.352077	3.309722583	0.053222032	-1.567467061	GmMYB176-Si	positive
577	134.8641252	0.486283533	0.053222032	-1.501639283	GmMYB176-Si	negative
578	119.0854295	3.687917467	0.053318204	-2.163894303	GmMYB176-Si	positive
579	280.1173314	3.0881205	0.053760923	1.565776786	GmMYB176-OE	positive
580	433.2526483	3.6426914	0.053998974	-1.901948075	GmMYB176-Si	positive
581	424.2483985	2.8844934	0.054583831	-1.452724365	GmMYB176-Si	positive
582	149.063179	0.494602943	0.054852488	-2.550924188	GmMYB176-Si	positive
583	253.1612822	3.877560117	0.054872363	-1.425996877	GmMYB176-Si	positive
584	373.1997912	4.224763517	0.054872363	-1.625746136	GmMYB176-Si	positive
585	313.2838655	3.630033067	0.054872363	-1.408389875	GmMYB176-Si	positive
586	362.2400878	2.597885133	0.054872363	-1.643665719	GmMYB176-Si	positive
587	186.1847464	3.5921194	0.054872363	-2.100092969	GmMYB176-Si	positive
588	278.2469815	3.767401767	0.054872363	-1.51171819	GmMYB176-Si	positive
589	177.0540668	2.687570033	0.054872363	-1.989323252	GmMYB176-Si	positive
590	265.2515486	4.825254	0.054872363	-1.177610389	GmMYB176-Si	positive
591	435.1291579	2.465790883	0.054872363	-1.211219607	GmMYB176-OE	negative
592	275.1722506	2.768005333	0.054956881	-1.71466148	GmMYB176-Si	positive
593	244.1174223	2.845650883	0.055053089	-2.08065937	GmMYB176-Si	positive
594	219.0505362	2.233559417	0.055053089	-2.312336014	GmMYB176-Si	negative
595	133.0493266	0.791747867	0.055350229	-2.317848592	GmMYB176-Si	negative
596	298.2731543	3.3181775	0.055616873	-2.064747218	GmMYB176-Si	positive
597	163.0787888	0.5149165	0.055636252	-2.653448802	GmMYB176-Si	positive
598	170.1534958	3.2962643	0.056455598	-2.153194902	GmMYB176-Si	positive
599	380.3356678	4.618588667	0.056817915	-2.13556321	GmMYB176-Si	positive
600	672.4135565	3.805362117	0.056892127	1.543988159	GmMYB176-OE	positive
601	853.458889	3.6900113	0.056892127	1.506790734	GmMYB176-OE	negative
602	143.0339066	2.2334995	0.056892127	-2.204255394	GmMYB176-Si	negative
603	167.0785683	0.768580333	0.056892127	-1.67688841	GmMYB176-Si	positive
604	316.2836605	3.24290775	0.057021254	-1.798149907	GmMYB176-Si	positive

605	317.2869239	3.243475367	0.057021254	-1.795245447	GmMYB176-Si	positive
606	317.1224237	2.784872733	0.057225581	-2.066733767	GmMYB176-Si	positive
607	254.1938234	2.76076375	0.057438616	-1.872298596	GmMYB176-Si	positive
608	250.159806	2.261798567	0.057438616	-1.325512418	GmMYB176-Si	positive
609	808.4564084	3.6896099	0.057468547	1.425101026	GmMYB176-OE	negative
610	557.9129403	5.56085065	0.057539376	-1.565693259	GmMYB176-Si	positive
611	211.1321303	3.241657233	0.057685331	-1.474707778	GmMYB176-Si	positive
612	182.1535322	3.319380483	0.057685331	-2.443588195	GmMYB176-Si	positive
613	432.2494151	3.6413468	0.057685331	-1.894091651	GmMYB176-Si	positive
614	184.0574865	0.790297163	0.057685331	-1.488738992	GmMYB176-Si	positive
615	291.2306942	4.399821	0.057685331	-1.770655413	GmMYB176-Si	positive
616	432.3181798	3.232518067	0.057685331	-1.562229754	GmMYB176-Si	positive
617	171.1486521	0.833987447	0.057685331	-3.795913595	GmMYB176-OE	positive
618	394.3054774	3.00952065	0.057685331	-1.748763844	GmMYB176-Si	positive
619	206.0442472	2.884389183	0.057685331	-1.791870093	GmMYB176-Si	positive
620	537.8758173	5.5764757	0.057685331	-1.267063219	GmMYB176-Si	positive
621	407.2614364	2.278458967	0.057685331	-1.645057076	GmMYB176-Si	positive
622	193.1218427	3.9014998	0.057685331	-1.403030405	GmMYB176-Si	positive
623	220.0962931	3.0881618	0.057685331	1.453480282	GmMYB176-OE	positive
624	438.377315	4.84445785	0.057685331	-2.169399586	GmMYB176-Si	positive
625	368.2897943	2.8272262	0.057685331	-1.634371155	GmMYB176-Si	positive
626	265.1789299	3.5053849	0.057685331	-1.539707429	GmMYB176-Si	positive
627	204.0861055	2.242383483	0.057685331	1.256058218	GmMYB176-OE	positive
628	199.1436777	2.225120033	0.057685331	-1.881837125	GmMYB176-Si	positive
629	395.2756489	4.332670033	0.057685331	1.320306473	GmMYB176-OE	positive
630	311.2682695	3.498814133	0.057694767	-1.760573654	GmMYB176-Si	positive
631	527.4328937	4.825338317	0.057694767	-1.97741844	GmMYB176-Si	positive
632	215.1747891	2.493767383	0.057694767	-1.827636525	GmMYB176-Si	positive
633	476.3272268	2.5142451	0.057836078	-1.533610116	GmMYB176-Si	positive
634	184.1691502	3.489833233	0.058091231	-1.938298607	GmMYB176-Si	positive
635	112.086935	4.847990867	0.058091231	1.107448322	GmMYB176-OE	positive
636	395.3087334	3.0097624	0.058091231	-1.780077132	GmMYB176-Si	positive

637	238.0701969	3.1559518	0.058091231	-2.522065166	GmMYB176-Si	positive
638	483.406622	4.834695	0.058091231	-2.047252229	GmMYB176-Si	positive
639	427.2078471	4.224152867	0.058091231	-1.156606493	GmMYB176-Si	positive
640	482.8591061	5.6036663	0.0581404	-1.382485846	GmMYB176-Si	positive
641	171.1026296	2.031160267	0.058269428	-1.478598462	GmMYB176-Si	positive
642	503.3553364	2.845650883	0.058269428	-1.880451799	GmMYB176-Si	positive
643	270.2249964	3.0289841	0.058269428	-1.866814605	GmMYB176-Si	positive
644	144.0651941	2.0592308	0.05832133	-2.200320279	GmMYB176-Si	positive
645	253.1905192	2.76031585	0.058620183	-1.72914856	GmMYB176-Si	positive
646	275.1016844	2.2894818	0.05888164	-3.491517612	GmMYB176-Si	positive
647	435.3318988	3.0008725	0.05888164	-1.555314535	GmMYB176-Si	positive
648	216.0859795	2.2826583	0.05888164	-1.458738205	GmMYB176-Si	positive
649	419.2752099	4.42342965	0.05888164	-1.778314897	GmMYB176-Si	positive
650	622.4425555	4.594079267	0.05888164	-1.888237922	GmMYB176-Si	positive
651	346.2211826	3.738255867	0.05888164	-1.519643666	GmMYB176-Si	positive
652	257.221639	2.961554317	0.05888164	-1.623093743	GmMYB176-Si	positive
653	295.2371499	3.235870283	0.059012667	-1.726963378	GmMYB176-Si	positive
654	580.4418911	2.739943267	0.059251356	-2.49723096	GmMYB176-Si	positive
655	272.2324431	2.0913597	0.059420322	-2.124062559	GmMYB176-Si	positive
656	167.1061409	3.741003417	0.059614127	-1.329806274	GmMYB176-Si	positive
657	251.1456202	3.679031533	0.059614127	-1.510828416	GmMYB176-Si	positive
658	329.2427065	2.758128	0.059627817	-1.533613212	GmMYB176-Si	positive
659	809.4686485	3.522589017	0.059814417	1.214640435	GmMYB176-OE	negative
660	238.0703161	2.8838723	0.059814417	-1.685061371	GmMYB176-Si	positive
661	384.3210235	3.0298683	0.059814417	-1.686959905	GmMYB176-Si	positive
662	388.2528544	2.303643267	0.059814417	-1.581465184	GmMYB176-Si	positive
663	436.1324507	2.466069667	0.059814417	-1.108864231	GmMYB176-OE	negative
664	810.4723305	3.519479833	0.059814417	1.230445799	GmMYB176-OE	negative
665	229.1904467	2.6738668	0.059824544	-1.361553798	GmMYB176-Si	positive
666	184.0599537	2.282668083	0.059824544	-1.833019328	GmMYB176-Si	positive
667	320.1935476	2.350973117	0.059952215	-1.735861817	GmMYB176-Si	positive
668	600.9312378	4.60904805	0.059952215	-1.962808514	GmMYB176-Si	positive

669	434.2734262	4.060047183	0.059952215	-1.51578143	GmMYB176-Si	positive
670	535.9000265	5.57225565	0.059952215	-1.520964482	GmMYB176-Si	positive
671	480.3615402	2.787359533	0.060358098	-2.042893292	GmMYB176-Si	positive
672	513.8870223	5.589507867	0.060477321	-1.495541133	GmMYB176-Si	positive
673	257.2659372	4.747509083	0.060477321	-1.171717685	GmMYB176-Si	positive
674	284.2938563	5.237040767	0.060477321	-1.318468796	GmMYB176-Si	positive
675	399.1612344	3.251643567	0.060620285	-1.514548884	GmMYB176-Si	positive
676	571.2341737	3.562977767	0.060727963	-1.695345139	GmMYB176-Si	positive
677	404.2872423	2.94998715	0.060727963	-1.605836382	GmMYB176-Si	positive
678	291.2033959	3.0289841	0.060727963	-1.709411765	GmMYB176-Si	positive
679	172.059881	2.237673867	0.060750725	1.230611545	GmMYB176-OE	positive
680	492.3504516	2.200567967	0.060750725	-1.745093956	GmMYB176-Si	positive
681	202.1795042	0.808223603	0.060823772	-1.4155884	GmMYB176-Si	positive
682	255.0642669	2.3949993	0.060823772	-1.716279752	GmMYB176-Si	positive
683	341.2742307	2.316323767	0.060823772	-1.547163744	GmMYB176-Si	positive
684	357.2642906	3.817864117	0.060823772	1.390662954	GmMYB176-OE	negative
685	355.2773295	2.692320367	0.060856224	-1.868876586	GmMYB176-Si	positive
686	722.5019686	4.352008583	0.061634021	-1.900678591	GmMYB176-Si	positive
687	385.1480638	3.258902467	0.061634021	-1.799248397	GmMYB176-Si	positive
688	172.0599993	1.935471833	0.061648741	1.045347655	GmMYB176-OE	positive
689	723.5050899	4.3517726	0.061648741	-1.910416178	GmMYB176-Si	positive
690	551.3532247	4.399821	0.061648741	-2.024538557	GmMYB176-Si	positive
691	421.3506028	4.844415733	0.06170884	-2.108114937	GmMYB176-Si	positive
692	226.1908453	2.095217933	0.061759615	-1.951759694	GmMYB176-Si	positive
693	303.1065682	3.26106425	0.061882903	-1.875693685	GmMYB176-Si	positive
694	258.2167777	2.020452867	0.061882903	-1.836050661	GmMYB176-Si	positive
695	634.4499431	4.378744033	0.061943263	-2.084375184	GmMYB176-Si	positive
696	171.0051066	0.45797983	0.061949742	-3.795017628	GmMYB176-Si	negative
697	308.2324039	2.817877867	0.062300619	-1.52655441	GmMYB176-Si	positive
698	356.2898512	2.7496064	0.062300619	-1.661068456	GmMYB176-Si	positive
699	236.1047371	3.5442354	0.062319379	-1.111575455	GmMYB176-Si	negative
700	630.4762869	4.3328761	0.062319379	-1.727060156	GmMYB176-Si	positive

701	294.2168348	2.700782233	0.062319379	-1.769179	GmMYB176-Si	positive
702	185.1281151	2.144345917	0.062319379	-1.983990514	GmMYB176-Si	positive
703	304.1099463	3.261116333	0.062319379	-1.828311275	GmMYB176-Si	positive
704	145.1332407	0.506573517	0.062319379	-2.355156163	GmMYB176-Si	positive
705	283.1894407	3.504220667	0.062319379	-1.491608029	GmMYB176-Si	positive
706	225.1590633	2.5128417	0.062319379	-1.872046577	GmMYB176-Si	positive
707	382.305347	2.9563549	0.062773256	-1.625658626	GmMYB176-Si	positive
708	377.3246617	4.844432333	0.062819643	-2.189987176	GmMYB176-Si	positive
709	535.3984708	5.5731736	0.062927274	-1.50089842	GmMYB176-Si	positive
710	341.1737683	4.2189389	0.063045534	-1.509715407	GmMYB176-Si	positive
711	293.2184491	3.1559248	0.063071677	-1.566457146	GmMYB176-Si	positive
712	319.3026622	3.3181775	0.063074373	-2.002800875	GmMYB176-Si	positive
713	370.3055009	2.956679267	0.063705503	-1.517705431	GmMYB176-Si	positive
714	144.0884439	4.089985383	0.063751916	-1.198235463	GmMYB176-Si	positive
715	287.1955405	2.376275567	0.063751916	-1.551763611	GmMYB176-Si	positive
716	272.1960296	2.225338833	0.063751916	-1.445762776	GmMYB176-Si	positive
717	536.4016394	5.5764757	0.063751916	-1.534794265	GmMYB176-Si	positive
718	168.1378384	3.120816333	0.064072945	-2.09280717	GmMYB176-Si	positive
719	297.2528801	3.367031817	0.064147167	-1.50507619	GmMYB176-Si	positive
720	333.2508162	2.662061267	0.064307691	-1.879591405	GmMYB176-Si	positive
721	323.2205672	3.8168632	0.064337505	-1.517813922	GmMYB176-Si	positive
722	228.1587903	2.883722367	0.064337505	-1.500011073	GmMYB176-Si	positive
723	376.2560215	2.690974533	0.064399551	-1.73951742	GmMYB176-Si	positive
724	308.8638793	0.041060178	0.064399551	-2.362480571	GmMYB176-Si	negative
725	578.4166051	4.622179633	0.064399551	-1.822501437	GmMYB176-Si	positive
726	635.3969207	2.702821133	0.064399551	-2.377618722	GmMYB176-Si	positive
727	174.0754801	2.0138801	0.064399551	-2.315244313	GmMYB176-Si	positive
728	334.301869	4.844415733	0.064399551	-2.287677866	GmMYB176-Si	positive
729	187.0967144	2.730276533	0.064399551	-1.4178328	GmMYB176-Si	negative
730	357.2931353	2.752709433	0.064399551	-1.694771686	GmMYB176-Si	positive
731	283.7166649	2.6535718	0.064399551	-2.182822784	GmMYB176-Si	positive
732	177.0547003	3.105889033	0.064399551	-1.172576428	GmMYB176-Si	negative

733	491.297543	2.481403117	0.064399551	-1.383330276	GmMYB176-Si	positive
734	492.3756603	5.602134033	0.064399551	-1.442273241	GmMYB176-Si	positive
735	439.326637	2.6189929	0.064399551	-1.854675001	GmMYB176-Si	positive
736	402.2669162	2.887526433	0.064399551	-1.47969477	GmMYB176-Si	positive
737	298.2117267	2.352067567	0.064399551	-1.864774898	GmMYB176-Si	positive
738	364.2559379	2.65382	0.064399551	-1.78342726	GmMYB176-Si	positive
739	439.3806229	4.8376713	0.064399551	-2.1822651	GmMYB176-Si	positive
740	592.4302207	4.3903907	0.064404178	-1.869829426	GmMYB176-Si	positive
741	371.1972075	3.9780384	0.064713916	-1.701552481	GmMYB176-Si	positive
742	386.1513693	3.261012167	0.064741562	-1.827912079	GmMYB176-Si	positive
743	1035.755203	5.649658033	0.064741562	-2.32316653	GmMYB176-Si	positive
744	190.0704904	1.9476884	0.064793485	1.219257544	GmMYB176-OE	positive
745	278.246963	4.346001833	0.064793485	-2.054036482	GmMYB176-Si	positive
746	854.4618541	3.6904059	0.064793485	1.406444168	GmMYB176-OE	negative
747	296.2404766	3.237458633	0.064793485	-1.622145356	GmMYB176-Si	positive
748	588.4076466	2.654345433	0.064859201	-1.806674281	GmMYB176-Si	positive
749	344.2266572	2.263683	0.065206277	-1.755923569	GmMYB176-Si	positive
750	104.1070879	0.43853687	0.065286365	-2.237573676	GmMYB176-Si	positive
751	491.2972059	4.533723733	0.065286365	-1.468138201	GmMYB176-OE	positive
752	143.9377092	4.153422767	0.065286365	2.227088253	GmMYB176-Si	negative
753	876.6581338	5.63683415	0.065286365	-1.628962701	GmMYB176-Si	positive
754	502.2265511	3.856897367	0.065286365	-1.922827232	GmMYB176-Si	positive
755	434.1149563	2.541716633	0.065286365	-1.665528157	GmMYB176-Si	positive
756	195.1374236	3.204423767	0.065663196	-1.357156352	GmMYB176-Si	positive
757	130.0496227	0.787349767	0.065663196	-2.867445196	GmMYB176-Si	positive
758	226.0703615	3.116915217	0.065663196	1.337981081	GmMYB176-OE	positive
759	309.252721	3.369955767	0.06583162	-1.536483038	GmMYB176-Si	positive
760	148.0388331	2.505643	0.066197831	-1.552181916	GmMYB176-Si	positive
761	209.1538854	3.4678634	0.066685827	-1.367990253	GmMYB176-Si	negative
762	256.0807621	2.82822815	0.066686792	-2.624613749	GmMYB176-Si	positive
763	338.2430909	2.579156333	0.066686792	-1.608904846	GmMYB176-Si	positive
764	485.2000437	3.855502767	0.066938005	-1.862827865	GmMYB176-Si	positive

765	189.1339873	0.440575927	0.066977827	-1.736624056	GmMYB176-Si	positive
766	271.2366213	3.15950275	0.066977827	-1.5159958	GmMYB176-Si	positive
767	354.2740702	2.692320367	0.066977827	-1.648465496	GmMYB176-Si	positive
768	433.1116805	2.5414939	0.066977827	-1.643264667	GmMYB176-Si	positive
769	312.2634738	2.351596967	0.066977827	-1.791785286	GmMYB176-Si	positive
770	413.3114154	2.410579833	0.066977827	-1.804262497	GmMYB176-Si	positive
771	238.070377	3.460784133	0.066977827	-2.283976368	GmMYB176-Si	positive
772	227.0737341	3.1125166	0.066977827	1.219153578	GmMYB176-OE	positive
773	321.2235742	4.50459645	0.066977827	-2.065334694	GmMYB176-Si	positive
774	425.1403946	3.261276867	0.066977827	-1.617160457	GmMYB176-Si	positive
775	145.0684785	2.059087433	0.066977827	-2.088999182	GmMYB176-Si	positive
776	421.1094832	2.7307622	0.066993244	-2.387843655	GmMYB176-Si	positive
777	369.2931246	2.8262078	0.067211123	-1.602160504	GmMYB176-Si	positive
778	314.2427396	2.391772067	0.067468494	-1.700423118	GmMYB176-Si	positive
779	298.2479774	2.2334481	0.068194966	-1.7634991	GmMYB176-Si	positive
780	256.080757	3.1652764	0.068194966	-2.062168797	GmMYB176-Si	positive
781	381.1170529	2.730334667	0.068279663	-2.162701795	GmMYB176-Si	positive
782	300.2273634	2.334255067	0.068279663	-1.690470569	GmMYB176-Si	positive
783	477.3397422	2.693848	0.068279663	-1.92907215	GmMYB176-Si	positive
784	451.3268238	2.625970233	0.068513655	-1.907976039	GmMYB176-Si	positive
785	530.2577725	3.855774033	0.068666475	-2.028342716	GmMYB176-Si	positive
786	210.0756215	2.9728857	0.068666475	-2.290537937	GmMYB176-Si	positive
787	272.1004929	3.165196267	0.068696774	-2.106207981	GmMYB176-Si	positive
788	513.3855556	5.5899293	0.068893467	-1.454332598	GmMYB176-Si	positive
789	166.0493522	2.505643	0.068893467	-1.48083801	GmMYB176-Si	positive
790	409.3161972	2.816059367	0.068893467	-1.42247774	GmMYB176-Si	positive
791	327.226906	2.680612433	0.068893467	-1.512904535	GmMYB176-Si	positive
792	144.0804262	2.19011575	0.068893467	-1.626339403	GmMYB176-Si	positive
793	631.418051	3.515164167	0.068893467	1.636491889	GmMYB176-OE	positive
794	366.274321	2.806466567	0.068893467	-1.650608918	GmMYB176-Si	positive
795	491.3724043	5.6036663	0.068915293	-1.415501783	GmMYB176-Si	positive
796	313.2306144	2.4056618	0.068915293	-1.56802011	GmMYB176-Si	positive

797	439.3557184	3.99204205	0.068915293	1.653282761	GmMYB176-OE	positive
798	378.3280051	4.844329267	0.068989559	-2.155837732	GmMYB176-Si	positive
799	631.4180565	3.80402585	0.06927819	1.545175532	GmMYB176-OE	positive
800	348.7716454	2.315292967	0.069337051	-1.520569679	GmMYB176-Si	positive
801	334.2090607	2.408898583	0.069570108	-1.729342568	GmMYB176-Si	positive
802	397.315962	2.259324833	0.069570108	-1.781703522	GmMYB176-Si	positive
803	352.2584934	2.6651563	0.069728725	-1.690798786	GmMYB176-Si	positive
804	395.3545862	4.8442428	0.070009605	-2.09562901	GmMYB176-Si	positive
805	337.2724003	4.350735433	0.070009605	-1.943561003	GmMYB176-Si	positive
806	308.0518571	3.165196267	0.070088053	-1.865106993	GmMYB176-Si	positive
807	811.4749697	3.522445467	0.07011224	1.193388378	GmMYB176-OE	negative
808	170.0918661	0.416346147	0.07011224	-1.91314503	GmMYB176-Si	positive
809	421.2770879	2.343096333	0.07011224	-1.933441738	GmMYB176-Si	positive
810	191.0543765	3.80402585	0.070443656	1.494026061	GmMYB176-OE	positive
811	469.3737064	2.769049567	0.070779332	-1.684812071	GmMYB176-Si	positive
812	342.2740289	2.648309833	0.070779332	-1.574531916	GmMYB176-Si	positive
813	408.3207725	3.127251767	0.070779332	-1.535098161	GmMYB176-Si	positive
814	479.3580317	2.788212633	0.071090241	-1.848776578	GmMYB176-Si	positive
815	515.3270673	2.475930217	0.071090241	-1.127759785	GmMYB176-OE	negative
816	196.1691022	3.480022933	0.071090241	-2.05159061	GmMYB176-Si	positive
817	422.3483288	3.6823944	0.071090241	1.677131596	GmMYB176-OE	positive
818	239.0737144	3.461456467	0.071240967	-2.272588764	GmMYB176-Si	positive
819	343.277429	2.6535718	0.071240967	-1.541405344	GmMYB176-Si	positive
820	132.0553817	5.272182533	0.071369764	-1.183923774	GmMYB176-OE	negative
821	239.1746943	2.616706333	0.071369764	-1.589984169	GmMYB176-Si	positive
822	495.3890861	2.962511267	0.071369764	-1.658434873	GmMYB176-Si	positive
823	569.2638336	2.4841965	0.071605654	-1.081442785	GmMYB176-Si	negative
824	578.9182796	4.622179633	0.071605654	-1.759312495	GmMYB176-Si	positive
825	284.2362975	3.758502233	0.071605654	-1.181540784	GmMYB176-Si	positive
826	622.9442389	4.594079267	0.071654468	-1.729303577	GmMYB176-Si	positive
827	531.2611902	3.860955067	0.071799448	-2.017416261	GmMYB176-Si	positive
828	380.2897962	2.894093967	0.071799448	-1.550119341	GmMYB176-Si	positive

829	534.3908587	4.6510753	0.071810081	-1.753392366	GmMYB176-Si	positive
830	225.7954971	4.220605033	0.071916577	-1.578533869	GmMYB176-OE	negative
831	545.4018573	3.2133824	0.071986774	-1.744552305	GmMYB176-Si	positive
832	514.3886143	5.58969615	0.071986774	-1.449055871	GmMYB176-Si	positive
833	199.1800257	2.174211333	0.071986774	-1.77143694	GmMYB176-Si	positive
834	270.0963788	3.165322733	0.07224797	-1.958260582	GmMYB176-Si	positive
835	196.1325999	2.5033042	0.072316847	-1.658717434	GmMYB176-Si	positive
836	523.4201277	3.220120567	0.072316847	-1.689953789	GmMYB176-Si	positive
837	285.2393865	3.755272833	0.072475762	-1.213043273	GmMYB176-Si	positive
838	336.2245228	2.392499533	0.072846764	-1.746739778	GmMYB176-Si	positive
839	807.4531742	3.6901428	0.07306701	1.33451746	GmMYB176-OE	negative
840	134.071053	2.7202454	0.07306701	-1.833582313	GmMYB176-Si	positive
841	688.4820828	4.545331633	0.07306701	-1.411564117	GmMYB176-Si	positive
842	835.4684642	3.514217	0.07306701	1.240807413	GmMYB176-OE	positive
843	430.3028513	3.136657433	0.073076725	-1.477991237	GmMYB176-Si	positive
844	318.2994193	3.311251133	0.073302445	-1.982308908	GmMYB176-Si	positive
845	377.1793868	3.251643567	0.073304672	-2.043709849	GmMYB176-Si	positive
846	187.0571941	0.686314347	0.073304672	-1.595993903	GmMYB176-Si	positive
847	272.0625346	2.5333285	0.073672304	-1.655470722	GmMYB176-Si	positive
848	246.1330447	2.8576379	0.073672304	-1.752209994	GmMYB176-Si	positive
849	269.2216671	3.028090267	0.073672304	-1.514577574	GmMYB176-Si	positive
850	209.0447743	2.516263433	0.073672304	-1.512713754	GmMYB176-Si	negative
851	421.3160698	2.8653831	0.073672304	-1.360759172	GmMYB176-Si	positive
852	552.4102739	2.606858433	0.073672304	-1.81015675	GmMYB176-Si	positive
853	184.1327851	2.624883233	0.073672304	-1.46271419	GmMYB176-Si	positive
854	421.1094832	2.7307622	0.073672304	1.172273868	GmMYB176-OE	positive
855	328.2584183	2.510110067	0.073672304	-1.767981177	GmMYB176-Si	positive
856	453.1378922	2.981965617	0.073672304	-1.706666372	GmMYB176-Si	positive
857	566.4258227	2.6624474	0.073672304	-1.694917172	GmMYB176-Si	positive
858	221.091455	2.064383283	0.073672304	-4.091531401	GmMYB176-Si	positive
859	256.0676701	2.39181015	0.073672304	-2.044940583	GmMYB176-Si	positive
860	305.2188751	3.1906231	0.073672304	-1.597797568	GmMYB176-Si	positive

861	171.1486521	0.833987447	0.073672304	-2.21311736	GmMYB176-Si	positive
862	375.2516554	4.303454167	0.073672304	-2.442362431	GmMYB176-Si	positive
863	333.298769	4.8445629	0.073672304	-2.062383949	GmMYB176-Si	positive
864	235.1410438	2.378844367	0.073672304	-1.803709687	GmMYB176-Si	positive
865	242.1936934	2.719388733	0.073672304	-1.551544312	GmMYB176-Si	positive
866	440.3109039	2.7467226	0.073749982	-1.511188498	GmMYB176-Si	positive
867	442.3743566	3.80068045	0.073757797	1.497445335	GmMYB176-OE	positive
868	217.0674839	0.69823528	0.073757797	-1.242319968	GmMYB176-Si	positive
869	204.1225945	1.965487333	0.073757797	-2.900476643	GmMYB176-Si	positive
870	497.4222631	5.0913624	0.073757797	-2.310772144	GmMYB176-Si	positive
871	426.1437478	3.2615381	0.073776282	-1.605848675	GmMYB176-Si	positive
872	245.1527967	4.0850282	0.074155335	-1.229403346	GmMYB176-Si	positive
873	322.2002659	4.0805875	0.074155335	-1.488250591	GmMYB176-Si	positive
874	417.1167291	2.393337233	0.074155335	-1.672237164	GmMYB176-Si	positive
875	280.0917109	0.494339847	0.074155335	-2.863248439	GmMYB176-Si	positive
876	326.3600081	3.687339933	0.074155335	-4.004661928	GmMYB176-Si	positive
877	644.9573514	4.5799867	0.074155335	-1.665134097	GmMYB176-Si	positive
878	132.0288065	0.440850297	0.074155335	-1.943159032	GmMYB176-Si	negative
879	672.4135565	3.805362117	0.074183508	-1.748285606	GmMYB176-Si	positive
880	341.2612909	2.5790351	0.07430565	-1.221343222	GmMYB176-Si	positive
881	343.2486208	3.435862333	0.07430565	1.129381517	GmMYB176-OE	negative
882	415.2101312	3.6880477	0.074581664	-2.675098429	GmMYB176-Si	positive
883	833.4627967	3.514217	0.074875005	1.201607545	GmMYB176-OE	positive
884	284.2325773	2.1415262	0.075380466	-1.879917223	GmMYB176-Si	positive
885	381.2931591	2.8993151	0.075819242	-1.512345286	GmMYB176-Si	positive
886	411.3397031	3.232756833	0.076151514	-1.462020667	GmMYB176-Si	positive
887	441.3423077	2.61822155	0.076151514	-1.561809245	GmMYB176-Si	positive
888	371.3004225	2.187419767	0.076151514	-1.38592346	GmMYB176-Si	positive
889	138.054557	0.483268147	0.076151514	-3.41066079	GmMYB176-Si	positive
890	767.5312097	4.332796533	0.076151514	-1.792938418	GmMYB176-Si	positive
891	201.1591799	2.272327967	0.076151514	-1.65851189	GmMYB176-Si	positive
892	429.3170604	4.6110031	0.076151514	-1.892959953	GmMYB176-Si	positive

893	256.2092993	2.845739033	0.076255783	-1.571486824	GmMYB176-Si	positive
894	585.4745486	5.072168733	0.076255783	-2.050423089	GmMYB176-Si	positive
895	259.1321413	3.760088967	0.076433423	-2.762557617	GmMYB176-Si	positive
896	285.0749601	2.90334745	0.076564686	-1.784939104	GmMYB176-Si	positive
897	517.3710333	2.966399967	0.076564686	-1.652454506	GmMYB176-Si	positive
898	204.0861055	2.242383483	0.077335488	-1.761489434	GmMYB176-Si	positive
899	385.3159091	2.200567967	0.077429463	-1.617883848	GmMYB176-Si	positive
900	834.4661678	3.514217	0.077457707	1.207659481	GmMYB176-OE	positive
901	284.2362975	3.758502233	0.077457707	1.580027682	GmMYB176-OE	positive
902	415.2101312	3.6880477	0.077495699	-3.072112682	GmMYB176-OE	positive
903	283.1743257	2.2160971	0.077700809	-1.264716419	GmMYB176-OE	positive
904	465.3423159	2.693848	0.077878818	-1.726964388	GmMYB176-Si	positive
905	483.1484502	2.951857967	0.077878818	-1.814926482	GmMYB176-Si	positive
906	132.0553817	5.272182533	0.077878818	-1.107391439	GmMYB176-Si	negative
907	515.8629334	5.5899293	0.077948194	-1.144986887	GmMYB176-Si	positive
908	556.9054568	4.636581133	0.077951317	-1.68364824	GmMYB176-Si	positive
909	146.0596749	2.188309717	0.078344399	-1.658251014	GmMYB176-Si	positive
910	285.2393865	3.755272833	0.078344399	1.687996186	GmMYB176-OE	positive
911	341.2331843	3.327735267	0.078344399	1.098717402	GmMYB176-OE	negative
912	680.4820368	4.36824435	0.078344399	-1.854474592	GmMYB176-Si	positive
913	441.3711838	3.8043727	0.078344399	1.465043281	GmMYB176-OE	positive
914	187.1437098	2.137448617	0.078344399	-1.802674907	GmMYB176-Si	positive
915	310.2480521	2.913149467	0.078371802	-1.318110864	GmMYB176-Si	positive
916	191.0187332	0.483412947	0.078371802	-3.86648469	GmMYB176-Si	negative
917	251.1632183	3.7401216	0.078820244	-1.903426384	GmMYB176-Si	positive
918	136.86148	0.486283533	0.079110332	-1.168760761	GmMYB176-Si	negative
919	584.4709449	5.072127433	0.079110332	-1.985486025	GmMYB176-Si	positive
920	282.2167486	2.597786817	0.079110332	-1.378173749	GmMYB176-Si	positive
921	253.1440856	3.468294133	0.079110332	-1.381485738	GmMYB176-Si	negative
922	385.2795362	2.2865866	0.079110332	-1.679935065	GmMYB176-Si	positive
923	353.2619535	2.6723342	0.079110332	-1.773799378	GmMYB176-Si	positive
924	396.3209226	3.078788083	0.079336476	-1.548411719	GmMYB176-Si	positive

925	559.8889157	5.560900467	0.079617851	-1.16261902	GmMYB176-Si	positive
926	130.085949	0.649924683	0.079617851	-1.956274151	GmMYB176-Si	positive
927	261.0572165	1.3285878	0.079646335	-1.840867568	GmMYB176-Si	positive
928	813.4868975	3.514654833	0.079646335	-2.293820209	GmMYB176-Si	positive
929	381.1170529	2.730334667	0.07982028	1.425983611	GmMYB176-OE	positive
930	493.3733009	2.874316267	0.07982028	-1.59733924	GmMYB176-Si	positive
931	672.5229217	5.049275583	0.07982028	-1.780190254	GmMYB176-Si	positive
932	350.3250904	4.844432333	0.080104813	-2.033932293	GmMYB176-Si	positive
933	487.3240372	2.692072133	0.080146559	-1.742437312	GmMYB176-Si	positive
934	475.3238986	2.493773	0.080191136	-1.495328457	GmMYB176-Si	positive
935	300.263427	2.306078267	0.080191136	-1.5855437	GmMYB176-Si	positive
936	439.098486	2.3957221	0.08068303	-1.668794484	GmMYB176-Si	positive
937	172.059881	2.237673867	0.080977119	-1.746854828	GmMYB176-Si	positive
938	483.3891283	2.9224447	0.081054449	-1.961131793	GmMYB176-Si	positive
939	428.310782	2.6994083	0.081054449	-1.505570769	GmMYB176-Si	positive
940	394.3513279	4.844329267	0.081285485	-1.953847589	GmMYB176-Si	positive
941	435.3663056	5.0913624	0.081285485	-2.24942569	GmMYB176-Si	positive
942	425.3469459	2.3922932	0.081285485	-1.645178557	GmMYB176-Si	positive
943	324.2884611	4.4095121	0.081285485	-2.244461453	GmMYB176-Si	positive
944	203.1787047	3.799427133	0.081285485	1.659621097	GmMYB176-OE	positive
945	267.2060208	2.918454817	0.081285485	-1.521698372	GmMYB176-Si	positive
946	193.0861843	3.4194301	0.08165604	-1.516296535	GmMYB176-Si	negative
947	286.211461	2.25546915	0.08165604	-1.460613698	GmMYB176-Si	positive
948	191.0544048	3.510984167	0.08165604	1.401625085	GmMYB176-OE	positive
949	339.2881774	4.4286386	0.08165604	-2.232162617	GmMYB176-Si	positive
950	325.2917584	4.4082335	0.08165604	-2.241934708	GmMYB176-Si	positive
951	523.1408258	2.952011017	0.08165604	-1.986171873	GmMYB176-Si	positive
952	811.480549	3.513493267	0.08165604	-2.233544472	GmMYB176-Si	positive
953	217.0674853	0.48842267	0.08165604	-1.570918004	GmMYB176-Si	positive
954	395.2756955	4.521486633	0.08165604	1.374127854	GmMYB176-OE	positive
955	608.4726807	2.9246773	0.08165604	-1.676081798	GmMYB176-Si	positive
956	159.1124436	1.5359905	0.081749322	-1.721922577	GmMYB176-Si	positive

957	283.2371189	3.187284233	0.081769461	-1.635742502	GmMYB176-Si	positive
958	327.2461766	2.523099367	0.081844858	-1.51421047	GmMYB176-Si	positive
959	461.3082182	2.626519267	0.081925605	-1.848553968	GmMYB176-Si	positive
960	149.0227974	0.225257633	0.081951946	-1.059155228	GmMYB176-Si	positive
961	810.5539182	4.321931333	0.081959491	-1.743471486	GmMYB176-Si	positive
962	397.324317	3.07847765	0.082049453	-1.503717853	GmMYB176-Si	positive
963	394.2058725	3.2516755	0.08215218	-1.953841999	GmMYB176-Si	positive
964	367.2775694	2.80212235	0.08215218	-1.637863753	GmMYB176-Si	positive
965	639.3187131	4.495559333	0.082162145	-3.101942244	GmMYB176-Si	positive
966	640.3220174	4.488999983	0.082162145	-3.102602704	GmMYB176-Si	positive
967	641.3248425	4.485318833	0.082162145	-3.150881605	GmMYB176-Si	positive
968	197.807088	2.62835775	0.082162145	-1.062231272	GmMYB176-Si	negative
969	317.2191714	3.2332766	0.082162145	-1.453428194	GmMYB176-Si	positive
970	505.3709886	2.922746267	0.082162145	-1.883268721	GmMYB176-Si	positive
971	132.8665896	0.485458458	0.082162145	-1.250249413	GmMYB176-Si	negative
972	421.3449395	3.6832877	0.082162145	1.509866968	GmMYB176-OE	positive
973	811.5575309	4.318422067	0.082265788	-1.730786464	GmMYB176-Si	positive
974	671.4102486	3.804592117	0.082265788	-1.645926321	GmMYB176-Si	positive
975	392.3436309	5.0913624	0.082303424	-2.412877799	GmMYB176-Si	positive
976	195.8104118	2.26310585	0.082329043	-1.460460919	GmMYB176-Si	negative
977	208.1537391	0.84394082	0.082356444	5.535256758	GmMYB176-OE	positive
978	568.4418682	2.701168167	0.082356444	-1.769248232	GmMYB176-Si	positive
979	399.2954032	2.329879633	0.082427187	-1.79322347	GmMYB176-Si	positive
980	442.3744795	3.5128246	0.082534817	1.498948754	GmMYB176-OE	positive
981	315.2670685	3.5102651	0.082699825	1.422500702	GmMYB176-OE	positive
982	242.1759525	3.093676917	0.082831391	-1.729374891	GmMYB176-Si	negative
983	453.3961832	5.090474233	0.083006518	-2.270929949	GmMYB176-Si	positive
984	356.3511069	3.5770986	0.083053794	-2.233392356	GmMYB176-Si	positive
985	455.357838	2.701168167	0.083053794	-1.95228032	GmMYB176-Si	positive
986	410.3363075	3.23299575	0.083164305	-1.28560004	GmMYB176-Si	positive
987	184.0599537	2.282668083	0.083164305	1.01947766	GmMYB176-OE	positive
988	418.302534	3.0792336	0.083423833	-1.538214555	GmMYB176-Si	positive

989	221.1891941	3.516593917	0.083456026	-1.554309437	GmMYB176-Si	positive
990	427.3267127	2.505606133	0.083456026	-1.295901081	GmMYB176-Si	positive
991	348.3176442	5.091373767	0.083456026	-2.35725778	GmMYB176-Si	positive
992	666.4686611	4.564696267	0.083456026	-1.585087493	GmMYB176-Si	positive
993	424.3639527	3.80343575	0.083456026	1.425381661	GmMYB176-OE	positive
994	437.2881485	4.537814667	0.083456026	-2.0285158	GmMYB176-Si	positive
995	219.1121525	2.36584645	0.083456026	-1.614542361	GmMYB176-Si	positive
996	811.480549	3.513493267	0.083456026	1.485505687	GmMYB176-OE	positive
997	509.4049021	3.080871617	0.08393234	-1.72829028	GmMYB176-Si	positive
998	813.4868975	3.514654833	0.08393234	1.498797056	GmMYB176-OE	positive
999	485.1980557	3.4222481	0.08393234	-1.621953887	GmMYB176-Si	positive
1000	123.0075044	3.513493267	0.08398138	1.609189765	GmMYB176-OE	positive
1001	202.1794895	0.485235263	0.084436571	-1.507691599	GmMYB176-OE	positive
1002	391.3402564	5.091376233	0.084436571	-2.277868831	GmMYB176-Si	positive
1003	355.280631	4.8445629	0.084436571	-1.889639874	GmMYB176-Si	positive
1004	452.3931154	5.0914626	0.084464325	-2.172269733	GmMYB176-Si	positive
1005	257.0841617	2.826407767	0.084468138	-2.597865493	GmMYB176-Si	positive
1006	387.1295246	2.983720167	0.084743558	1.202096192	GmMYB176-OE	negative
1007	425.3671015	3.512155933	0.084743558	1.503901177	GmMYB176-OE	positive
1008	128.0190621	0.374393163	0.084743558	-1.241738361	GmMYB176-Si	positive
1009	189.073336	2.187588733	0.084743558	-1.616614149	GmMYB176-Si	positive
1010	193.1621043	0.499812913	0.084743558	2.95734014	GmMYB176-OE	positive
1011	419.1220962	2.393785433	0.084743558	-1.617173669	GmMYB176-Si	positive
1012	357.1147099	2.787570333	0.084746674	-1.92274032	GmMYB176-Si	positive
1013	408.2810879	4.7212942	0.084788305	-1.39857235	GmMYB176-Si	positive
1014	239.0737144	3.461456467	0.084802731	1.082567272	GmMYB176-OE	positive
1015	439.355451	3.6832877	0.084913477	1.442998022	GmMYB176-OE	positive
1016	409.3702258	5.0913713	0.085009975	-2.310872701	GmMYB176-Si	positive
1017	440.3590431	3.39588755	0.085070796	1.500584349	GmMYB176-OE	positive
1018	423.3606098	3.80414165	0.085137299	1.383183578	GmMYB176-OE	positive
1019	282.2249756	3.083208433	0.085137299	-1.617574244	GmMYB176-Si	positive
1020	264.9995908	0.683756963	0.085137299	-1.641405106	GmMYB176-Si	positive

1021	399.3065059	4.838349833	0.08522295	-1.919401451	GmMYB176-Si	positive
1022	125.0595088	2.626182	0.085305451	-1.505733508	GmMYB176-Si	positive
1023	540.4449708	5.081727033	0.085596188	-2.027168766	GmMYB176-Si	positive
1024	399.1916684	4.069304867	0.085596188	-1.63794966	GmMYB176-Si	positive
1025	133.1007579	3.516593917	0.085828446	-1.608947975	GmMYB176-Si	positive
1026	477.3303112	2.4856775	0.085850371	-1.327062059	GmMYB176-Si	positive
1027	217.194398	3.513493267	0.085990655	1.390395949	GmMYB176-OE	positive
1028	812.4839263	3.5128246	0.086136507	1.463310444	GmMYB176-OE	positive
1029	662.303821	4.491549467	0.086136507	-2.543944593	GmMYB176-Si	positive
1030	203.1788333	3.512155933	0.086168575	1.328671437	GmMYB176-OE	positive
1031	281.2216333	3.083208433	0.086168863	-1.634722952	GmMYB176-Si	positive
1032	613.4074696	3.513636217	0.086168863	1.508227578	GmMYB176-OE	positive
1033	191.0187332	0.483412947	0.086168863	2.146606868	GmMYB176-OE	negative
1034	594.4571844	2.8357953	0.086168863	-1.704211438	GmMYB176-Si	positive
1035	533.4016356	3.184376667	0.086168863	-1.483674802	GmMYB176-Si	positive
1036	202.0856758	2.365840967	0.086168863	-1.619505326	GmMYB176-Si	positive
1037	803.4516778	3.600904	0.086168863	1.313855312	GmMYB176-OE	positive
1038	521.4046217	3.117568533	0.086376697	-1.496479371	GmMYB176-Si	positive
1039	421.3448877	3.193892633	0.086376697	1.5466217	GmMYB176-OE	positive
1040	661.3005369	4.494966933	0.086376697	-2.493078821	GmMYB176-Si	positive
1041	831.4467254	3.684181	0.086376697	1.367624118	GmMYB176-OE	positive
1042	363.2491824	4.417508483	0.086376697	1.043625977	GmMYB176-OE	positive
1043	347.3142723	5.0918338	0.086514664	-2.350591566	GmMYB176-Si	positive
1044	195.1222548	2.090055933	0.086795911	-1.545873743	GmMYB176-Si	positive
1045	228.0860164	2.701672233	0.086795911	-1.518766421	GmMYB176-Si	positive
1046	405.3501996	3.5128246	0.086795911	1.418505754	GmMYB176-OE	positive
1047	656.3449754	4.485840117	0.086929835	-3.099864253	GmMYB176-Si	positive
1048	871.4697234	3.2071388	0.086957888	1.320282584	GmMYB176-OE	negative
1049	396.8004623	2.692072133	0.087050352	-2.500162004	GmMYB176-Si	positive
1050	421.3452502	3.40027435	0.087059615	-2.068583132	GmMYB176-Si	positive
1051	177.1631245	3.513493267	0.087178547	1.379809787	GmMYB176-OE	positive
1052	137.0230447	2.7343767	0.087178547	-1.098854122	GmMYB176-Si	negative

1053	266.0760513	0.507784087	0.087224933	-2.11234293	GmMYB176-Si	positive
1054	809.4649189	3.685872683	0.087277828	1.443129246	GmMYB176-OE	positive
1055	793.437816	3.400142833	0.087277828	2.779688549	GmMYB176-OE	negative
1056	172.1058201	2.022985883	0.087277828	-1.509624706	GmMYB176-Si	positive
1057	195.0291366	3.020476567	0.087598911	-1.382404248	GmMYB176-Si	negative
1058	335.1227615	2.692072133	0.087701181	-1.744485786	GmMYB176-Si	positive
1059	192.1588507	0.490893173	0.088365526	3.024448171	GmMYB176-OE	positive
1060	313.2111474	2.5790351	0.088522963	-1.338692658	GmMYB176-Si	positive
1061	487.3584031	4.828237767	0.089003286	-1.668644126	GmMYB176-Si	positive
1062	400.2986851	2.333497133	0.089003286	-1.737301552	GmMYB176-Si	positive
1063	335.1227615	2.692072133	0.089003286	1.116494643	GmMYB176-OE	positive
1064	496.4188534	5.0913624	0.089019382	-2.090035547	GmMYB176-Si	positive
1065	364.3406228	5.0913535	0.089214093	-2.40228736	GmMYB176-Si	positive
1066	325.2361719	4.166298883	0.089214093	-1.530301786	GmMYB176-Si	positive
1067	256.0806995	2.4339076	0.089214093	1.197893615	GmMYB176-OE	positive
1068	671.4102486	3.804592117	0.089214093	1.355019358	GmMYB176-OE	positive
1069	485.1296605	2.734505483	0.089214093	1.62228543	GmMYB176-OE	negative
1070	389.7923594	2.644849233	0.089214093	-2.063027084	GmMYB176-Si	positive
1071	810.4683599	3.687564367	0.089214093	1.441949673	GmMYB176-OE	positive
1072	443.3326016	4.834862483	0.089214093	-1.72693002	GmMYB176-Si	positive
1073	301.2110985	2.4812657	0.089280879	-1.396057491	GmMYB176-Si	positive
1074	228.178011	2.2665651	0.089280879	-1.35290443	GmMYB176-Si	positive
1075	278.0628165	2.798582533	0.089280879	1.620366487	GmMYB176-OE	positive
1076	423.3606953	3.388018267	0.089280879	1.441261291	GmMYB176-OE	positive
1077	301.2305464	2.317189633	0.089498983	-1.596563537	GmMYB176-Si	positive
1078	455.3479883	2.481403117	0.089722963	-1.25036457	GmMYB176-Si	positive
1079	189.1631588	3.513779167	0.089722963	1.304416519	GmMYB176-OE	positive
1080	206.0998828	2.187588733	0.089722963	-1.494176874	GmMYB176-Si	positive
1081	197.8073823	2.2552706	0.089868722	-1.458140529	GmMYB176-Si	negative
1082	406.353398	3.513493267	0.089868722	1.419252198	GmMYB176-OE	positive
1083	277.1879498	2.845562733	0.089868722	-1.455111146	GmMYB176-Si	positive
1084	461.3284512	3.160480817	0.090555459	-1.565977941	GmMYB176-Si	positive

1085	215.1786592	3.513493267	0.090661716	1.276644399	GmMYB176-OE	positive
1086	440.3590431	3.39588755	0.091067575	-1.655519761	GmMYB176-Si	positive
1087	825.464215	3.2075426	0.091433704	1.198236021	GmMYB176-OE	negative
1088	152.0562768	0.661732693	0.091433704	-1.946065025	GmMYB176-Si	positive
1089	260.0522491	2.8844934	0.091492536	-1.337461936	GmMYB176-Si	positive
1090	271.0591581	2.5408737	0.091492536	-1.585697103	GmMYB176-Si	positive
1091	229.1942628	3.513493267	0.091492536	1.342264377	GmMYB176-OE	positive
1092	487.3052291	2.4843897	0.091492536	-1.009740343	GmMYB176-Si	negative
1093	173.0440057	3.484257783	0.091653972	1.536531782	GmMYB176-OE	positive
1094	127.0388378	2.218918867	0.091826665	1.284900574	GmMYB176-OE	positive
1095	286.2478692	2.182424133	0.091826665	-1.358993727	GmMYB176-Si	positive
1096	421.3452502	3.40027435	0.09196696	1.882408181	GmMYB176-OE	positive
1097	298.2117267	2.352067567	0.092348935	-1.258591121	GmMYB176-OE	positive
1098	292.0780659	3.165234867	0.092348935	-1.706246032	GmMYB176-Si	positive
1099	307.2256606	3.752562667	0.09236894	1.39418049	GmMYB176-OE	positive
1100	405.3499566	3.803679	0.09236894	1.366148561	GmMYB176-OE	positive
1101	106.0396815	5.361417167	0.09236894	-1.023587037	GmMYB176-Si	negative
1102	408.3667396	5.0913535	0.09236894	-2.227178273	GmMYB176-Si	positive
1103	684.3762336	4.487075067	0.092423839	1.854492566	GmMYB176-OE	positive
1104	407.1323794	3.324577267	0.092811338	-1.556044014	GmMYB176-Si	positive
1105	481.3735025	2.845562733	0.092893907	-1.74121018	GmMYB176-Si	positive
1106	303.2034368	3.0881205	0.092990161	-1.397051238	GmMYB176-Si	positive
1107	193.1621043	0.499812913	0.093840593	-2.067144051	GmMYB176-Si	positive
1108	175.14753	3.513636217	0.093942393	-1.571443978	GmMYB176-Si	positive
1109	203.0889161	2.36450455	0.093970222	-1.559231343	GmMYB176-Si	positive
1110	475.1240763	2.7343767	0.094014977	-1.115143637	GmMYB176-Si	negative
1111	209.0292918	0.45459486	0.094739531	3.151536603	GmMYB176-OE	negative
1112	386.2638122	2.3957221	0.094850054	-1.421941311	GmMYB176-Si	positive
1113	201.1630078	3.5128246	0.094948681	1.411104069	GmMYB176-OE	positive
1114	387.1919635	4.247031167	0.094948681	-2.482106672	GmMYB176-Si	positive
1115	225.7954971	4.220605033	0.094948681	-1.028440663	GmMYB176-Si	negative
1116	112.086935	4.847990867	0.094948681	1.156211555	GmMYB176-Si	positive

1117	637.3032389	4.659503967	0.094948681	-3.2015801	GmMYB176-Si	positive
1118	424.3640554	3.510984167	0.094948681	1.348394615	GmMYB176-OE	positive
1119	220.1154564	2.3685575	0.094948681	-1.560811248	GmMYB176-Si	positive
1120	467.4119181	5.3841819	0.094948681	-2.539675929	GmMYB176-Si	positive
1121	160.096431	2.1965387	0.094948681	-2.714725584	GmMYB176-Si	positive
1122	128.0340142	0.821015213	0.094948681	1.128395071	GmMYB176-OE	negative
1123	284.7246424	2.701168167	0.094948681	-1.568693359	GmMYB176-Si	positive
1124	308.1933114	2.254219367	0.09503573	-1.269832524	GmMYB176-Si	positive
1125	766.5280853	4.332796533	0.095131011	-1.577548488	GmMYB176-Si	positive
1126	794.4411923	3.400142833	0.095270369	2.702984214	GmMYB176-OE	negative
1127	326.2427446	2.521919567	0.095314057	-1.500094471	GmMYB176-Si	positive
1128	440.1018243	2.39616165	0.095314057	-1.568467453	GmMYB176-Si	positive
1129	202.0468298	2.513566733	0.095844292	-1.562653135	GmMYB176-Si	positive
1130	439.3555267	3.194045	0.096170831	1.705668079	GmMYB176-OE	positive
1131	347.2665484	2.770266917	0.096274981	-1.746327569	GmMYB176-Si	positive
1132	303.1067021	2.9607561	0.096274981	-1.711502837	GmMYB176-Si	positive
1133	269.1738364	3.863835867	0.097293491	-1.196293198	GmMYB176-Si	positive
1134	290.7245387	2.7420543	0.097293491	-1.615657331	GmMYB176-Si	positive
1135	149.0227974	0.225257633	0.097293491	-1.063683659	GmMYB176-OE	positive
1136	155.0333751	3.5128246	0.097301804	1.29493307	GmMYB176-OE	positive
1137	194.1169328	0.513964645	0.097414801	-1.140282623	GmMYB176-OE	positive
1138	488.3087213	2.48400925	0.097470742	-1.223770219	GmMYB176-Si	negative
1139	493.8500871	5.6066448	0.098667061	-1.067729961	GmMYB176-Si	positive
1140	401.1067652	2.711875183	0.098667061	-1.63189124	GmMYB176-Si	positive
1141	457.3659821	3.203416167	0.098673691	1.666473023	GmMYB176-OE	positive
1142	581.1483572	2.711264733	0.099141086	-1.672606035	GmMYB176-Si	positive
1143	171.1486928	2.8937235	0.099267191	-1.018734385	GmMYB176-Si	positive
1144	460.3251233	3.1586073	0.099387014	-1.586486552	GmMYB176-Si	positive

Curriculum Vitae

Name: Arun Kumaran Anguraj Vadivel

Post-secondary Education and Degrees:

Karpagam University,
Coimbatore, India
2003-2006 B.Sc.

SKASC
Coimbatore, India
2006-2008 M.Sc

Western University
London, Ontario, Canada
2011-2017 Ph.D.
Passed Ph.D. comprehensive exam with distinction

Related Work Experience:

Senior Research Fellow (2009-2011) Sugarcane Breeding Institute, India
Teaching Assistant (2011-2017) Western University, Canada]

Publications:

Mainali, H., **Anguraj Vadivel, A. K.**, Li, X., Gijzen, M. and Dhaubhadel, S. (2017).Soybean cyclophilinGmCYP1 interacts with an isoflavonoid regulator GmMYB176. Scientific Reports 7, Article number: 39550.

Anguraj Vadivel, A. K. (2015). Gel-based proteomics in plants: time to move on from the tradition. Frontiers in Plant Science, 6, 369.

Anguraj Vadivel, A.K., Sukumaran, A., Li, X. and Dhaubhadel, S. (2015). Soybean isoflavonoids: role of GmMYB176 interactome and 14-3-3 proteins. Phytochemistry Reviews, pp. 1-13.