9-30-2016 12:00 AM

Identification and Characterization of the Isoflavonoid-Specific Prenyltransferase Gene Family to Prevent Stem and Root Rot in Soybean

Arjun Sukumaran, *The University of Western Ontario*

Supervisor: Dr. Sangeeta Dhaubhadel, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biology

Abstract

Soybean is one of the most predominantly grown legumes worldwide, however, one deterrent to maximizing its yield is the pathogen, *Phytophthora sojae*, which causes stem and root rot disease. Many strategies have been implemented to combat this pathogen such as use of pesticides and certain agricultural practices. However, these have been largely ineffective in completely preventing *P. sojae* infection. An alternative strategy would be to improve the innate resistance of soybean by promoting increased glyceollin production. Glyceollins are soybean-specific antimicrobial agents which are derived from the isoflavonoid branch of the general phenylpropanoid pathway. Soybeans produce 3 forms of glyceollin: glyceollin I, glyceollin II, and glyceollin III that result from the prenylation on either the C-2 or C-4 carbon of glycinol. Currently, only the C4-prenyltransferase (PT) referred to as glycinol 4-dimethylallyltransferase (G4DT), responsible for glyceollin I biosynthesis, has been identified. The objective of the present research is to identify and characterize the isoflavonoid-specific prenyltransferase gene family in soybean. This study has identified a family of 11(iso)flavonoid-specific soybean prenyltransferases (*GmPTs*) from a total of 63 *GmPTs*. Nine of these genes were characterized, from which, five were found to be induced upon pathogen infection. These induced GmPTs were shown to localize to plastids and display differential tissue-specific expression. They all contain promoter motifs associated with induction by a broad range of biotic and abiotic stress. In addition, one QTL was found to contain the loci for *GmPT01*. These findings demonstrate that there is isoflavonoid-specific *GmPT* family that play a role in glyceollin production.

Keywords

Soybean; Isoflavonoids; Specialized metabolite; Biosynthesis; Prenyltransferase; Metabolon; Biotic and abiotic stress; Subcellular localization; Tissue-specific expression; Enzyme Assay; Promoter Analysis; QTL

Acknowledgments

Long before and much after my Master's project, my family have always been my pillar of support. They have encouraged me throughout my academic career through both the highs and lows. And were it not for them, I would not be in the position I am right now. I am forever grateful for their love and support through these past two years.

I would like to express my deepest gratitude to my supervisor, Dr. Sangeeta Dhaubhadel who probably, more so than myself, believed in me and gave me this incredible opportunity. Her vision and support is what help to sculpt this thesis, and were it not for her guidance, I truly believe I would not have reached this day. Furthermore, I would like to thank my co-supervisor, Dr. Denis Maxwell who I knew would always be willing to help whenever I needed. I would also like to thank my advisory committee members, Dr. Ali Hannoufa and Dr. Mark Bernards, for their helpful advice and guidance throughout my project.

Through my time Master's project I saw many faces in the Dhaubhadel Lab, all of whom contributed in their own way to this project. I would have been lost was it not for Ling Chen who was always there with a helping hand. I am grateful that I was able to spend the past two years with my wonderful lab members (both past and present): Dr. Jaeju Yu, Dr. Mehran Dastmalchi, Arun Kumaran Anguraj Vadivel, Kishor Duwadi, Pravesh Lama, Hong Mo (Timothy) Kim and Caroline Sepiol, who I knew always had my back.

I owe a special thanks to Chelsea Ishmael from Dr. Mark Gijzen's lab for inoculating soybeans for this project. Thanks to both Tim Mcdowell and Dr. Kevin Burgess for their guidance and mentoring through the isolation of glycinol. I would like to thank Dr. Ben Rubin for his advice regarding statistics. Lastly, thanks to Alex Molnar for gracing this thesis with his incredible artistic abilities.

To the rest of the members of Agriculture Canada, I owe you a big thank you for making work an enjoyable place to be.

Table of Contents

Abstract i

Acknowledgments ii

Table of Contents iii

List of Tables vi

List of Figures vii

List of Abbreviation viii

Chapter 1: Introduction 1

1.1 Soybeans 1

1.2 Soybean agriculture 1

1.2.1 *Phytophthora sojae* 2

1.3 Defense strategies against stem and root rot disease in soybean 2

1.3.1 Agricultural practices 3

1.3.2 Cultivar development 3

1.4 Isoflavonoids 7

1.4.1 Role of isoflavonoids in plants 7

1.4.2 The benefits of isoflavonoids to human health 8

1.5 Isoflavonoid biosynthesis 9

1.5.1 Isoflavonoid metabolon 15

1.6 Prenyltransferases 16

1.6.1 Structure and function 16

1.6.2 Glycinol-4-dimethylallyltransferase 17

1.7 Soybean genome 18

1.8 Hypothesis and objectives of the study 21

Chapter 2: Materials and Methods 22

2.1 Plant materials and growth conditions 22

2.2 Bacterial and yeast strains 22

2.3 *In silico* and phylogenetic analysis 22

2.3.1 Promoter analysis 23

2.4 Pathogen infection and $AgNO_3$ treatment 24

2.5 RNA extraction, RT-PCR and qPCR 26

2.6 Statistical analysis 26

2.7 Cloning and plasmid construction 27

2.8 Subcellular localization 31

2.8.1 Transient expression of GmPTs in *N. benthamiana* 31

2.8.2 Confocal microscopy 31

2.9 Yeast microsome preparation 32

2.10 Glycinol production 32

2.11 Enzyme assay 34

2.12 QTL analysis 34

Chapter 3: Results 36

3.1 The soybean (iso)flavonoid prenyltransferase gene family contains 11 members . 36

3.1.1 Conservation of catalytic sites in candidate GmPTs 37

3.2 *GmPT* expression analysis 44

3.2.1 Candidate *GmPTs* are induced during *P. sojae* infection 44

3.2.2 $AgNO_3$-treatment mimics expression pattern of *P. sojae* infection 46

3.3 Candidate *GmPTs* show tissue-specific expression 49

3.4 Promoter analysis 51

3.5 Candidate GmPTs localize to plastids 53

3.6 Functional analysis of GmPTs 56

3.6.1 Production and isolation of glycinol 56

3.6.2 Enzyme assay of GmPT10a 68

3.7 QTL analysis 68

Chapter 4: Discussion 71

4.1 Soybean (iso)flavonoid prenyltransferase family contains 11 members 71

4.2 Five *GmPTs* are induced upon infection 73

4.3 *GmPTs* are spatially distributed under normal conditions 74

4.4 All GmPTs localize to plastids 75

4.5 Evolution of isoflavonoid-specific GmPTs 76

4.6 QTL and promoter analysis 77

4.7 Update on the isoflavonoid metabolon 78

4.8 Isoflavonoid biosynthesis in response to pathogen infection 81

4.9 Future directions 84

Appendix 86

References 97

Curriculum Vitae 109

List of Tables

Table 2.1 List of primers used for subcellular localization, qPCR and protein expression 29

Table 3.1 Characteristics of candidate *GmPTs* 40

Table 3.2.2 Pairwise comparison of nucleotides and amino acid sequences of isoflavonoid-specific candidate GmPTs 41

Table 3.3 Promoter elements found upstream of induced *GmPTs* 52

Table 3.4 QTL marker associated with *P. sojae* resistance contains *GmPT01* 70

Table 3.5 QTL markers that flanks induced *GmPTs* 70

List of Figures

Figure 1.1 The isoflavonoid biosynthetic pathway in soybeans. 13

Figure 1.2 Predicted folding of G4DT in the plastid membrane. 19

Figure 1.3Preylation of glycinol by glycinol-4-dimethylallyltransferase (G4DT). 20

Figure 2.1 $AgNO_3$ treatment of soybean hypocotyls. 25

Figure 3.1 Phylogenetic analysis of soybean prenyltransferases. 38

Figure 3.2 Identification of catalytic site residues in isoflavonoid-specific candidate GmPTs 45

Figure 3.4 Soybean hypocotyl physiology during $AgNO_3$ treatment. 47

Figure 3.5 Gene expression of *GmPTs* in response to $AgNO_3$ treatment. 48

Figure 3.6 Tissue-specific expression of candidate *GmPTs* during normal growth and development. 50

Figure 3.7 Subcellular localization of the candidate GmPTs 54

Figure 3.8 Metabolite profile of control and $AgNO_3$-treated soybean seeds. 58

Figure 3.9 Effect of inhibitors on metabolite profile. 60

Figure 3.10 Altering the solvent gradient alleviates co-elution of metabolites. 62

Figure 3.11LC-MS of $AgNO_3$-treated soybean seeds 64

Figure 3.12 Purification and Confirmation of Glycinol 66

Figure 3.13 Enzyme assay of GmPT10a 69

Figure 4.1 Generalized model of the isoflavonoid metabolon. 80

Figure 4.2 Isoflavonoid-specific prenyltransferase distribution in soybean. 82

List of Abbreviation

2HDR	2’-hydroxydaidzein reductase
2HID	2-hydroxyisoflavone dehydratase
3,9DPO	3,9’-dihydroxypterocarpan 6a-monooxyenase
4CL	4-coumarate-Co:A-ligase
ACN	acetonitrile
ADT	arogenate dehydratase
$AgNO_3$	silver nitrate
ATP	adenosine triphosphate
ATPS2	ATP sulfurylase 2
att	attachment
AVG	aminoethoxyvinylglycine
Avr	avirulence
BLAST	basic local alignment search tool
BLASTP	protein BLAST
bp	base pair
C4H	cinnamate-4-hydroxylate
cDNA	complementary DNA
CFP	cyan fluorescent protein
CHI	chalcone isomerase
CHR	chalcone reductase
CHS	chalcone synthase
CPA1	cation/proton antiporter 1

cv.	cultivar
DAP	days after pollination
DFR	dihydroflavonol 4-reductase
DMAPP	dimethylallyl pyrophosphate
DNA	deoxyribonucleic acid
ER	endoplasmic reticulum
EtOH	ethanol
ETR1	ethylene receptor
FA	formic acid
FARM	first aspartate-rich motif
G2DT	glycinol 2-dimethylallyltransferase
G4DT	glycinol 4-dimethylallyltransferase
GmPT	soybean prenyltransferase
GS	glyceollin synthase
H63	harosoy 63
HGA	homogentisate acid
HPLC	high-performance liquid chromatography
IFS	isoflavone synthase
kDa	kilo-Dalton
KPB	potassium phosphate buffer
LB	lysogeny broth
LC-MS	liquid chromatography-mass spectrometry
LC-MS/MS	liquid chromatography-tandem mass spectrometry

LDL	low-density lipoprotein
LOD	logarithm of odds
mAU	milli absorbance units
MEP	methylerythritol phosphate pathway
MES	2-(*N*-morpholino)ethanesulfonic acid
$MgCl_2$	magnesium chloride
MS^2	tandem mass spectrometry
MT	malonyltransferase
MVA	melvonate
N_2	nitrogen gas
N8DT	naringenin 8-dimethylallyltransferase
NADH	nicotinamide adenine dinucleotide
NMR	nuclear magnetic resonance
NTC	no template control
OD	optical density
PAL	phenylalanine ammonialyase
PCR	polymerase chain reaction
PGT	p-hydroxybenzoate:geranyltransferase
PHB	p–hydroxybenzoic acid
PLAM	plastid-associated membranes
PT	prenyltransferase
PTS	pterocarpan synthase
qPCR	real-time polymerase chain reaction

QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
Rps	resistance to Phytophthora sojae
RT-PCR	reverse transcription-polymerase chain reaction
SARM	second aspartate-rich motif
SD	synthetic dextrose
SEM	standard error of the mean
SNP	single nucleotide polymorphism
SSR	simple sequence repeats
TAE	tris-acetate-EDTA buffer
TBE	tris-borate-EDTA buffer
TBP	TATA-binding protein
UGT	glycosyltransferase
YFP	yellow fluorescent protein

Chapter 1: Introduction

1.1 Soybeans

An excellent source of proteins, essential fatty acids, carbohydrates, vitamins and minerals, soybean (*Glycine max* [L.] Merr.) is one of the most widely grown grain legumes worldwide. In 2015/2016 there was 313.26 million tons of soybean produced globally, with an additional 4.35% increase expected for 2016/2017 (GlobalSoybeanProduction.com, 2016). Canada ranks as the seventh largest producer in the world (GlobalSoybeanProduction.com, 2016). Primarily known for being a human food source, soybeans are also used to produce industrial products, animal feed, and biofuels. Soybean has been a staple in Asian diets for many centuries. Towards the end of the 20th century, soybean's health benefits led to a rise in popularity within Western countries. It was initially seen as a high-quality protein source that is low in saturated fat, thus, it was seen as a better alternative to animal sources of protein. Already valued for its nutritional content, soybean's popularity would only continue to grow as plant research would identify numerous biologically active chemical compounds naturally produced in the plant. These compounds, referred to as phytochemicals, are abundantly found in soybean, and are believed to confer many health benefits to humans (Messina, 2010b, a; Messina and Messina, 2010).

1.2 Soybean agriculture

For many decades, most of the soybean production in Canada occurred in Southern Ontario, but since the 1970s, production has begun to spread to other provinces including Manitoba, Quebec, Saskatchewan, Alberta and the Maritimes. Canada has been seen as a provider of `top-quality` premium soybeans, but has now begun to expand towards production of commodity soybeans that are processed into soy proteins, oil and other industrial products. Two thirds of the country`s soybean production are destined for export. Soybean is Canada's fourth largest field crop grown, and has become a major part of the country's economy (SoyCanada, 2016)

1.2.1 *Phytophthora sojae*

One of the major deterrents to maximizing the yield of soybean is the oomycete *P. sojae*. This pathogen infects soybean and certain species of lupins (Tyler, 2007). First identified in the United States in the 1950s (Kaufmann and Gerdemann, 1958), it has since spread around the globe. *P. sojae* causes stem and root rot disease in soybean, and is responsible for approximately $1-2 billion in annual soybean losses worldwide (Tyler, 2007).

P. sojae produces three kinds of spores: sporangia, zoospores and chlamydospores. Sporangia either germinate to produce hyphae or differentiate to produce zoospores. The main method of dispersal for *P. sojae* is through zoospores, which are short-lived spores that have two flagella used for swimming to their host. Guided by chemotactic attraction of isoflavonoids found in soybean root exudates, they swim towards the soybean roots (Morris and Ward, 1992). Once they reach the roots, zoospores differentiate to adhesive cysts, thus plugging water-conducting tissues. From then, the cyst produces a hyphae that will penetrate the root. Lastly, chlamydospores are long-lived thick-walled spores that exist for preservation through unfavourable conditions (Tyler, 2007).

P. sojae favours cool, damp environments in low and/or poorly drained areas. The disease can occur at any stage of soybean development, but primarily affects seeds and seedlings. The main symptom of the disease is pre- and post-emergence “dampening-off” of the stem and roots. Seedlings that have not died before emergence tend to germinate slowly and rapidly die (OMAFRA, 2014). Infection of already emerged plants results in water-soaked regions of the stem that soften and brown, resulting in wilting of the stem and eventual death of the plant (Schmitthenner, 1985).

1.3 Defense strategies against stem and root rot disease in soybean

Successfully defending against *P. sojae* infection is essential for farmers to ensure minimal to no crop loss. Defense strategies to combat *P. sojae* include using resistant cultivars, improved cultural practices, fungicides and biological control.

1.3.1 Agricultural practices

Certain agricultural practices can prevent or mitigate potential crop loss. *P. sojae* prefers cool and poorly drained areas, therefore, proper soil management is critical to preventing ideal conditions for growth. Practices that are implemented include preventing soil compaction, which can be done so by improving soil structure and drainage. Tillage is important for slow-draining fields and will help increase the soil temperature. Seed treatment using fungicides can protect germinating seeds and seedlings temporarily until disease resistance is built. Rotation with other crops is relatively ineffective as *P. sojae* oospores can remain in the soil for long periods of time. The most important strategy that has proven to have been the most effective in preventing crop loss is cultivar selection. Proper selection aims to select a cultivar that is tolerant and/or resistant to the race(s) of *P. sojae* present in the field (OMAFRA, 2014).

1.3.2 Cultivar development

Soybean cultivars that are tolerant and/or resistant to *P. sojae* infection have been identified and have been widely grown. Furthermore, cultivars resistant to stem and root rot would have an added benefit of reduced production expenses to farmers. In the past the development of such cultivars relied on traditional breeding, in which extensive cross-breeding is performed to introduce a desirable trait (e.g. stem and root rot resistance) into a genetic background with another desirable trait (e.g. high yield). Continuous backcrossing would be performed to ensure that levels of both traits were maximized. Conventional breeding methods were prominent in the 1960s to 80s using phenotypic assays to determine whether seedlings possessed resistance to *P. sojae* (Rps) genes. However, this method was costly and time-consuming for breeders (Young, 1999).

In the past few decades the development of new cultivars has relied on the identification of genes that govern desirable traits in soybean. The 1990s brought along technological advancements in the form of techniques such as restriction fragment length polymorphism (RFLP) and others. These techniques led to the development of genetic maps, which allowed for mapping genes and quantitative trait loci (QTL). A QTL is a

region of the genome that is correlated with the variation of a quantitative trait. QTLs are identified by its linkage to genes that control this phenotype (MacKay et al., 2009). The discovery of molecular markers, advances in genotyping and development of statistical methods have led to growth in the field of mapping QTLs. These molecular markers are commonly single nucleotide polymorphisms (SNPs), insertions, deletions or simple sequence repeats (SSRs). The development of SNPs markers has improved marker coverage and allowed for semi-automated screening of samples. The purpose of mapping QTLs is to understand the genetic basis of quantitative phenotypes. QTLs are mapped by their genetic linkage to segregating molecular markers for the phenotype of interest.

Commonly referred to as the "candidate gene approach", this approach relies on linkage mapping, positional cloning and insertional mutagenesis. As reviewed by Phlieger et al., this approach consists of three chronological steps, beginning with identification of candidate gene(s) through molecular or physiological studies, or through linkage data. Following identification, the candidate gene is screened for a polymorphism that is likely to alter the gene's expression. Finally, experiments must be performed to statistically validate the polymorphism's impact on the phenotype of interest (Pflieger et al., 2001). The cumulative information has allowed researchers to identify molecular markers associated with genes and QTLs of interest (Sugimoto et al., 2011). Several studies have identified QTLs associated with partial resistance to *P. sojae* infection, and further identified molecular markers that aid in cultivar development (Han et al., 2008; Tucker et al., 2010).

1.3.2.1 Complete vs. partial resistance

Originally coined by J.E. Van der Plank, there are two forms of resistance: vertical (complete) and horizontal (partial) resistance. Complete resistance is monogenic, with a single gene having a large and differential response (Fry, 1982). Thus far, 15 *Rps* alleles (*Rps1-9*) mapped to nine loci have been identified in soybean. Within these 15, there are five alleles of *Rps1 (Rps1a, Rps1b, Rps1c, Rps1d* and *Rps1k*) and three alleles of *Rps3 (Rps3a, Rps3b* and *Rps3c*). All the *Rps* genes provides complete resistance with the

exceptional of the root-specific *Rps2* that confers incomplete resistance (Mideros et al., 2007).

All *Rps* genes share the common function of recognizing corresponding *avirulence genes (Avr)* in *P. sojae*. Extensive analysis of *Avr1b-1* identified a small secreted protein that may be recognized by Rps1b in soybean, therefore, triggering a defense response (Shan et al., 2004). The structure of the secreted protein contains a signal peptide with an RxLR and/or dEER motif downstream, which plays a role in delivery to host cell's cytosol (Dou et al., 2008). *Rps* genes are predicted to trigger an effector-triggered immune response upon recognition of the Avr secreted protein (Dong et al., 2011). Soybeans containing *Rps1b* were resistant to infection from *P. sojae* strains expressing Avr1b-1 (Tyler, 2007). *P. sojae* virulent to *Rps 1b*-containing soybean cultivars either possessed *Avr1b-1* with a heavily mutated coding region (indicating strong divergent selection) or the gene is transcriptionally regulated to not express Avr1b-1 (Shan et al., 2004).

Complete resistance is conferred against some but not all races of the pathogen. Whereas, partial resistance provides incomplete protection against all races (Sugimoto et al., 2011). The benefit of complete resistance for breeders is that so long as the soybean cultivar being grown has all the *Rps* genes to the *P. sojae* present in the soil, the farmer will be safe from crop loss. These *Rps* genes provide effective resistance for 8-15 years (Schmitthenner, 1985), after which, identification and deployment of a new *Rps* gene-containing soybean cultivar is required. With continual usage of *Rps* genes as the first line of defense against *P. sojae* infection, selective pressure against this pathogen has led to the evolution of many of *P. sojae* races. In the USA, there have been at least 55 races of *P. sojae* that have been identified, with this number only continuing to grow due to use of *Rps* genes (Sugimoto et al., 2011).

Partial resistance consists of multiple genes with small and uniform effects (Fry, 1982). Unlike complete resistance, partial resistance has different degrees of effectiveness. Glover and Scott described partial resistance as the ability to survive infection without showing severe symptoms (e.g. death or yield loss) (Glover and Scott, 1998). Characteristics of partial resistance are fewer rotted roots and slower disease progression

(Sugimoto et al., 2011). The yield of soybean cultivars with strong partial resistance was not significantly different from those with *Rps* genes in the presence of *P. sojae*. Partial resistance was shown to be effective in environments with diverse *P. sojae* populations under moderate disease pressure. Furthermore, strong partial resistance did not negatively impact soybean yield (Dorrance et al., 2003). Partial resistance is a better long-term solution to the problem, as dependence on multiple genes would be broader and more durable than dependence on a single gene (i.e. complete resistance). Therefore, a cultivar that provides strong partial resistance would seldom need to be replaced, other than if a more resistant cultivar were to be bred. Implementing additional control measures such as improved soil drainage or fungicide seed treatment may be necessary to complement partial resistance (Sugimoto et al., 2011).

Partial resistance is a polygenic response to pathogen infection, as such, there are multiple components of the defense response that are split into two forms: physical and molecular. Plants contain two physical barriers to pathogen invasion in the root: the endodermis and the exodermis (De Coninck et al., 2015). Root-specific boundaries in the endodermis and exodermis referred to as "Casparian strips" contain the polyester, suberin. Suberin is also found in underground plant tissues and certain aboveground tissues (e.g. bundle sheath and phellem). This lipophilic macromolecule is a waterproof waxy substance consisting of a poly(aliphatic) and a poly(phenolic) domain (Bernards, 2002). The primary function of suberized cells is the prevention of water loss. It is also proposed to have a second function as a preformed and/or wound-induced barriers to fungal pathogens (Lulai and Corsini, 1998). It was shown there is a quantitative difference in the amount of suberin when comparing a resistant and susceptible cultivar of soybean during *P. sojae* infection (Thomas et al., 2007).

Soybeans produce a vast array of compounds and secondary metabolites that have been associated with the defense response. Yi et al found that when soybean hypocotyls were infected with *Phytophthora megasperma f.sp. glycinea* that synthesis and accumulation of B-1,3-glucanases and chitinases were induced (Yi and Hwang, 1996). These enzymes are proposed to be involved in plant defense through hydrolysing the cells walls of pathogens (Balasubramanian et al., 2012). Many phenylpropanoid compounds such as coumestrol,

coumarin, anthocyanins, flavonoids and isoflavonoids are induced upon biotic and abiotic stress in plants.

1.4 Isoflavonoids

Isoflavonoids are plant specialized metabolites that are synthesized from a branch of the general phenylpropanoid pathway. Previously thought to be exclusive to legumes, isoflavonoids have been found at extremely low levels in some non-legume plants (Lapcik, 2007). Isoflavonoids play two essential roles in soybean, root nodule formation and plant defense. Isoflavonoids confer positive benefits to human health through antimicrobial, antifungal, and antioxidant properties (Dakora and Phillips, 1996).

1.4.1 Role of isoflavonoids in plants

Isoflavonoids play two critical roles in soybeans: as preformed and induced defense compounds and as signalling molecules for nitrogen fixing bacteria.

Isoflavonoids act as phytoalexins and are associated with defense against pathogens. Phytoalexins are low molecular weight compounds that are induced upon pathogen attack. Studies have shown increased isoflavonoid accumulation in correspondence with pathogen attack in soybean (Abbasi and Graham, 2001; Lozovaya et al., 2004; Kubeš et al., 2014). Pre-existing isoflavones conjugates are hydrolysed, and conversely converted to more complex derivatives in response to infection (Graham et al., 1990). The soybean phytoalexin, glyceollin is found as three isomers, named glyceollin I, II and III. Glyceollin I is the predominant form in all tissues except for leaves, in which glyceollin III is the most abundant isomer. Infected hypocotyls tend to accumulate glyceollin I quickly, and then levels declined after 48 hours. However, glyceollin II and III continued to accumulate past 48 hours, but at much lower levels that glyceollin I. In leaves, levels of all isomers would rapidly accumulate and decline after 48 hours. However, in tissues bordering the site of inoculation, the three isomers would gradually accumulate (Bhattacharyya and Ward, 1986). Glyceollin accumulation was similar between soybean suspension cultures and seedling tissues indicating that wounding does not stimulate

glyceollin production (Ayers et al., 1976). These compounds were also produced by soybean in response to various abiotic stresses (Stössel, 1982; Graham and Graham, 1991). Glyceollins function by interfering with H+-transporting ATPases in the plasma membrane and tonoplast (Giannini et al., 1988), and mitochondrial electron transport by targeting site I between NADH dehydrogenase and coenzyme Q (Kaplan et al., 1980; Boydston et al., 1983).

Isoflavonoids also play a role in the establishment of a mutualistic interaction between soybeans and nitrogen-fixing bacteria. The isoflavone aglycones daidzein and genistein have been reported to induce transcription of nod genes in *Bradyrhizobium japonicum* (Kosslak et al., 1987; Smit et al., 1992). The necessity for isoflavonoids in this symbiotic relationship was proven when the biosynthetic gene, *isoflavone synthase (IFS)* was silenced by RNA interference, it led to reduced nodulation (Subramanian et al., 2006). Nitrogen is an essential nutrient to plants, and through the relationship brought upon by isoflavonoids, legumes are able to directly acquire nitrogen from the atmosphere.

1.4.2 The benefits of isoflavonoids to human health

Through clinical and epidemiological studies, isoflavonoids have also been associated with numerous health benefits in humans. They have been tied to prevention and treatment of certain cancers, osteoporosis and heart disease as well as alleviating symptoms of menopause (Messina, 2010a). Benefits to human health stem from isoflavonoid's estrogen-like structure. By competing for binding with estrogen receptors, isoflavonoids will either exert anti-estrogenic effects in high estrogen environments, or estrogenic effects in low estrogen environments (Messina, 1999; Messina, 2010a).

Interest in isoflavonoids began when a research group in Japan discovered that the isoflavone, genistein was capable of inhibiting the activity of tyrosine-specific protein kinases expressed in cancer cells (Akiyama et al., 1987). Following this finding, isoflavonoid research pertaining to its beneficial effects for human health spiked. Expectedly, the primary focus of this research was on cancer prevention and treatment. Correlating diet to cancer prevalence, Asian countries were found to have lower

prevalence of breast and prostate cancer, thus these two forms of cancer became the main focus of isoflavonoid research (Pisani et al., 2002). Studies have shown long-term intake of soy is correlated to reduced breast cancer risk (Lamartiniere et al., 1995; Duffy et al., 2007). Furthermore, soy consumption during childhood was linked to reduced risk of developing breast cancer (Shu et al., 2001; Korde et al., 2009). Similarly in terms of prostate cancer, isoflavone-rich diets were correlated with reduced progression (Pollard and Suckow, 2006) and invasion of prostate cancer cells (Xu et al., 2009).

Studies have also tied isoflavonoid and soy protein intake with reduced risk of cardiovascular disease, osteoporosis and menopausal symptoms. Studies have associated soy proteins intake with reductions in serum cholesterol levels, thus reducing chronic heart disease and mortality risk (Anderson et al., 1995; Zhan and Ho, 2005). Of these reduction, 60-70% were attributed to isoflavonoids (Anderson et al., 1995). To compound its benefits as a non-animal protein source, isoflavonoids in soyfood help to manage hypocholesterolemia through one's diet. A rodent study, later corroborated with a clinical study in hypercholesterolemic postmenopausal women found that soy proteins rich in isoflavonoids led to improved bone density (Arjmandi et al., 1996; Potter et al., 1998). Isoflavonoids have also been associated with modestly improving menopausal symptoms, such as reductions in hot flashes and night sweats (Howes et al., 2006; Cheng et al., 2007).

1.5 Isoflavonoid biosynthesis

The general phenylpropanoid pathway is responsible for the production of numerous secondary metabolites. The pathway shown in Figure 1.1 starts with phenylalanine, a product of the shikimate pathway, which is deaminated by phenylalanine ammonia-lyase (PAL) and results in cinnamic acid. Cinnamic acid is then hydrolyzed by a P450 enzyme, cinnamate 4-hydroxylase (C4H), and then converted to p-coumerate-CoA by 4-coumerate-CoA ligase (4CL). From here the phenylpropanoid pathway branches to produce a wide range of metabolites including anthocyanins, flavonoids, isoflavonoids, lignins, stilbenes and tannins.

The pathway branches towards flavonoid, isoflavonoid and anthocyanin production by the activity of chalcone synthase (CHS). In the presence of three molecules of malonyl-CoA, CHS condenses p-coumerate-coenzyme A (CoA) to naringenin chalcone. The *CHS* gene family consists of nine genes, *CHS1-9* (Akada and Dube, 1995; Tuteja and Vodkin, 2008) and dCHS1/ICHS1, a duplicate copy of CHS1 (Senda et al., 2002; Matsumura et al., 2005). Phylogenetic analysis revealed that CHS1-6 and CHS9 are closely related to one another, whereas, CHS7 and CHS8 cluster together in a separate clade (Matsumura et al., 2005). Comparison of gene expression between a high seed isoflavonoid cultivar (RCAT Angora) and a low seed isoflavonoid cultivar (Harovinton) during embryo development revealed that *CHS7* and *CHS8* were expressed at significantly higher levels in RCAT Angora. The expressions of these two genes increased in soybean embryos from 30 days after pollination (DAP) to 70 DAP, which correlated with an increased in isoflavonoid content over the same time period (Dhaubhadel et al., 2007). The involvement of *CHS7* and *CHS8* in isoflavonoid biosynthesis was confirmed when silencing of these two genes reduced isoflavonoid accumulation compared to wild type samples. Promoter sequence differences between *CHS7* and *CHS8* are responsible for differential tissue-specific expression and expression during different stages of development (Yi et al., 2010).

Chalcone reductase (CHR), an aldo/ketoreductase, co-acts with CHS and catalyzes the reduction of an intermediate produced by CHS to produce isoliquiritigenin chalcone. The *GmCHR* gene family is composed of 16 members (Sepiol, 2015). Silencing of *CHR* genes led to significant reduction of all 5-deoxyisoflavonoids, including the soybean aglycone, daidzein and further led to reduced resistance against *P. sojae* infection (Subramanian et al., 2006; Graham et al., 2007).

Both naringenin and isoliquiritigenin chalcone are further acted on by chalcone isomerase (CHI) to produce naringenin and liquiritigenin, respectively. There are two types of *CHI*s: type 1 is specific to legume plants while type 2 is found in all plant species. There are twelve members within the *CHI* gene family of soybean grouped into four subfamilies. Of the twelve members, only GmCHI1A, GmCHI1B1, GmCHI1B2 (Type 2)

and GmCHI2 (Type 1) are enzymatically capable of isomerizing chalcones to their corresponding flavones (Ralston et al., 2005; Dastmalchi and Dhaubhadel, 2015).

Naringenin acts as a substrate for flavonoid and isoflavonoid metabolism, whereas isoliquiritigenin acts as a substrate for only isoflavonoid metabolism. The combined activity of isoflavone synthase (IFS) and 2-hydroxyisoflavone reductase (2HID) produces the three isoflavone aglycones: daidzein, genistein and glycitein. There are two *IFS* genes in soybeans, *IFS1 and IFS2* (Jung et al., 2000). These two genes encodes proteins that differ by 14 amino acids and tissue-specific expression (Dhaubhadel et al., 2003). Silencing of *IFS* genes led to reduced isoflavone accumulation and increased susceptibility to *P. sojae* infection (Subramanian et al., 2005).

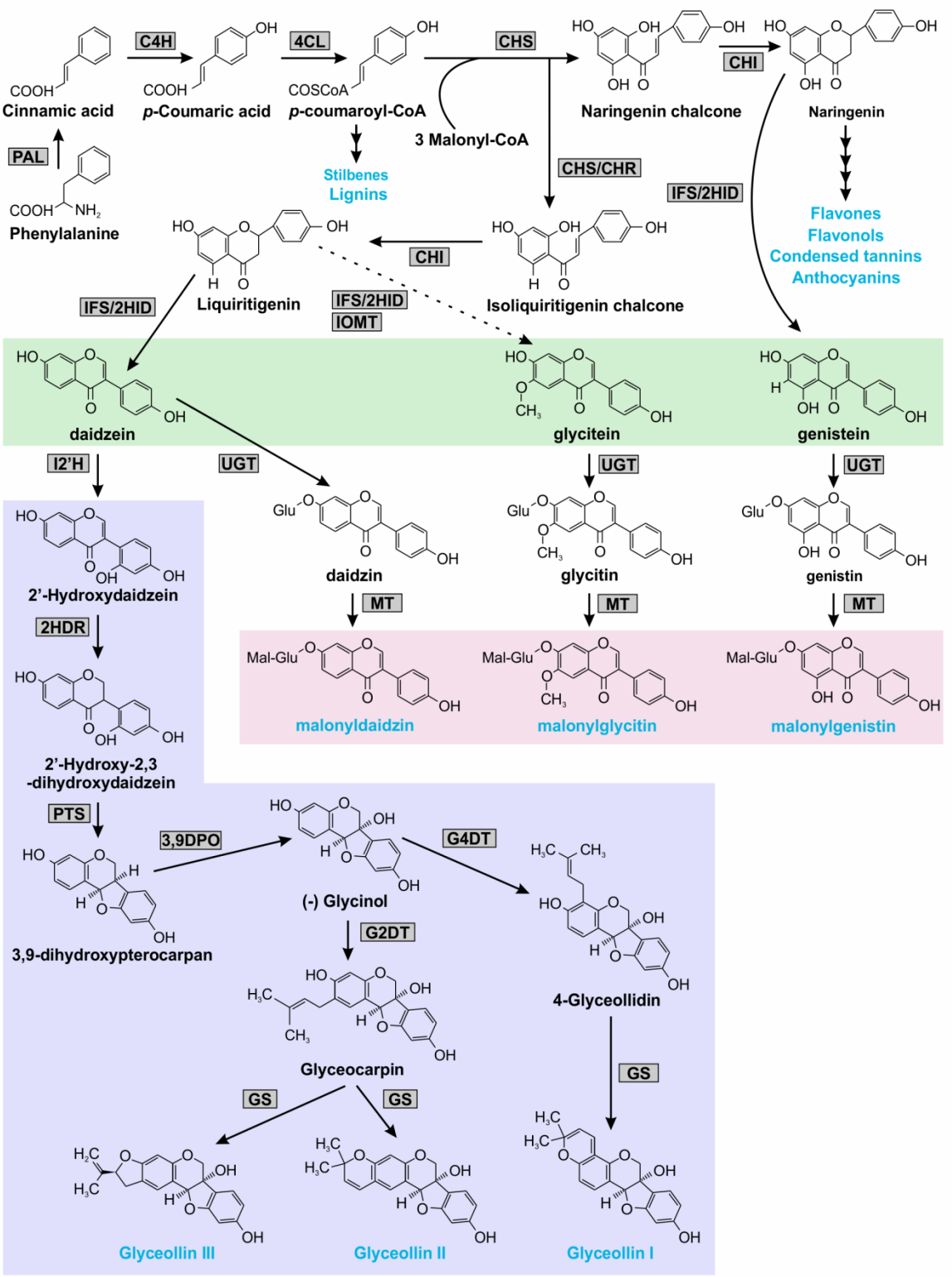
The isoflavone aglycones are normally conjugated to 7-*O*-glycosides (daidzin, genistin and glycitin) by uridine diphosphate glycosyltransferases (UGTs) and then further to a malonyl glycoside derivatives (6”-*O*-malonyldaidzin, 6”-*O*-malonylgenistin and 6”-*O*-malonylglycitin) by malonyltransferases (MTs) (Kudou et al., 1991; Dhaubhadel et al., 2008). The aglycones are reactive and toxic, but conjugation allows for storage in cell vacuoles though increased water solubility and improved chemical stability (Jones and Vogt, 2001).

Upon biotic or abiotic stress, daidzein acts as a substrate for further enzymatic activity that culminates in the production of the soybean phytoalexin, glyceollin. The first step towards glyceollin production is the hydroxylation of daidzein by isoflavone 2’-hydroxylase (Kochs and Grisebach, 1986; Akashi et al., 1998). This product is then subsequently reduced by 2’-hydroxydaidzein oxidoreductase (2HDR) (Fischer et al., 1990b) and then cyclized by pterocarpan synthase (PTS) (Fischer et al., 1990a). This compound is then converted to glycinol by 3,9-dihydroxypterocarpan 6a-monooxygenase (3,9DPO) (Schopfer et al., 1998).

Soybean produces three isomers of glyceollin: glyceollin I, glyceollin II and glyceollin III. The production of the three forms of glyceollin is the result of the differential prenylation of glycinol. The first instance of isoflavonoid-specific prenyltransferase

activity in soybean was found in 1979 by Zahringer et al. They demonstrated that when dimethylallyl pyrophosphate (DMAPP) was incubated with the particulate fraction of *P. megasperma var. sojae*-treated soybean cotyledons a new product was formed. Furthermore, the amount of this new product increased when glycinol was added to the incubation mixture. NMR spectrum of the product identified there were two compounds one with a prenylated C-4 position and one with a prenylated C-2 position, thereby, identifying a biosynthetic intermediate between glycinol and glyceollins (Zähringer et al., 1979). At the moment, the prenylation at the C-4 position by glycinol 4-dimethylallyltransferase (G4DT) has been confirmed (Akashi et al., 2009). Prenylation by G4DT produces 4-glyceollidin, whereas, prenylation of the C-2 position by G2DT (not yet identified) on glycinol is predicted to produce glyceocarpin. The final step in the pathway is catalyzed by the cytochrome P450 monooxygenase, glyceollin synthase (GS), which will cyclize 4-glyceollidin to produce glyceollin I and cyclize glyceocarpin to produce either glyceollin II or glyceollin III. The differential cyclization of glyceocarpin might indicate the presence of multiple GS, but this has not been tested (Welle and Grisebach, 1988).

Figure 1.1 The isoflavonoid biosynthetic pathway in soybeans. The pathway begins with the amino acid phenylalanine, and is converted into many metabolites (blue font). The pathway produces the three isoflavone aglycones (highlighted in green), which are stored as conjugated glycosides (highlighted in pink). Upon biotic or abiotic stress, further enzymatic reactions results in an induced portion of the pathway (highlighted in blue). PAL (phenylalanine ammonia-lyase); C4H (cinnamate-4-hydroxylate) 4CL (4-coumarate-CoA-ligase); CHS (chalcone synthase), 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)' PTS (pterocarpen synthase); 3,9 DPO (3,9-dihydroxypterocarpan 6a-monooxygenase); G4DT (glycinol 4-dimethylallyltransferase); G2DT (glycinol 2-Dimethylallyltransferase), GS (glyceollin synthase). Adopted from Anguraj Vadivel et al., 2015.


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
Flavones
Flavonols
Condensed tannins
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
2'-Hydroxy-2,3
-dihydroxydaidzein
malonyldaidzin
malonylglycitin
malonylgenistin
PTS
3,9DPO
G4DT
3,9-dihydroxypterocarpan
(-) Glycinol
G2DT
4-Glyceollidin
Glyceocarpin
GS
GS
GS
Glyceollin III
Glyceollin II
Glyceollin I

1.5.1 Isoflavonoid metabolon

In higher organisms, metabolic processes are compartmentalized at various levels (from tissues to organelles) to allow for increased efficiency and precision. The idea of metabolic complexes began in the 1970s, but the concept came into fruition when adopted by Paul Srere. In 1985, he coined the term "metabolon" in reference to a "supramolecular complex of sequential metabolic enzymes and cellular structural elements". In his vision, such a complex would allow channelling of substrates along components for increased efficiency (Srere, 1985). The complex is envisioned to be an area where high concentrations of intermediates are in direct contact with the active site of component enzymes, whilst being shielded from the cytoplasm. Advantages of such a system would include eliminating the transition time required for intermediates to move from one active site to the next, thereby, bypassing kinetic constraints and protecting labile intermediates by eliminating transport through the cytoplasm. The coordination could also lead to interaction with shared enzymes/intermediates of other pathways (Ralston and Yu, 2006).

The concept of a multi-enzyme complex was proposed by Helen Stafford in the 1970s as a way to control flux amongst the multiple branches of the phenylpropanoid pathway (Stafford, 1974). The model proposes that members of this pathway are spatially organized as endoplasmic reticulum (ER)-associated multiprotein complexes with the membranes-bound proteins serving as anchors points for the rest of pathway enzymes (Winkel, 2004, 2009).

Extensive support for metabolite channelling between the first two members of the phenylpropanoid pathway has been reported through the years. Beginning in 1975, PAL and C4H were reported to assemble on microsomal membranes (Czichi and Kindl, 1975), which went against knowledge that PAL exists as a soluble proteins. A substrate feeding experiment demonstrated C4H showed preference for cinnamate (produced by PAL) over [^{14}C]-cinnamate (exogenously supplied) (Czichi and Kindl, 1977). The interaction between the soluble PAL and the ER-bound C4H were supported by subsequent studies (Hrazdina and Wagner, 1985; Rasmussen and Dixon, 1999).

Additional studies to uncover evidence for a (iso)flavonoid metabolon have led to identified interactions between PAL and CHS (Schopker 1985), CHS, CHI and DFR (Burbulis and Winkel-Shirley, 1999) and CHS and CHI (Saslowsky and Winkel-Shirley, 2001). Interactions with soluble proteins (i.e. PAL, CHS, CHI and DFR) were abundant, but recently, a study looking into the isoflavonoid metabolon has demonstrated a membrane-tethered complex involving the aforementioned proteins (Dastmalchi et al., 2016).

1.6 Prenyltransferases

1.6.1 Structure and function

Prenyltransferases are a class of enzymes that transfer an allylic prenyl moiety to an acceptor molecule (Liang et al., 2002). The transfer of the prenyl group is an electrophilic alkylation or a Friedel-Craft alkylation.

Depending on the stereochemistry of the product these enzymes are split into two classes, *trans*- or *cis*-prenyltransferase. Within *trans*-prenyltransferases, there are two aspartate-rich motifs that are essential for substrate binding and catalytic activity. These aspartate-rich motifs are not found in *cis*-prenyltransferases. The number of prenyl groups added is regulated by residues upstream of the first aspartate-rich motif (FARM) in trans-prenyltransferases. Whereas, in *cis*-prenyltransferase, residues localized on the 3rd helix are responsible for chain length (Takahashi and Koyama, 2006).

Members of the UbiA superfamily of prenyltransferases are widespread in all life forms from bacteria to animals and function numerous biological processes and diseases (Li, 2016). The namesake of the superfamily comes from the *E. coli* prenyltransferase, 4-hydroxybenzoic acid oligoprenyltransferase (UbiA), shown to be crucial for the production of ubiquinone (Young et al., 1972). Genes with this domain typically have prenyltransferase activity, and are commonly referred to as a whole as the UbiA superfamily. In plants, these are commonly integral membrane proteins that contain at least one aspartate-rich motif (NDxxDxxxD) and require the presence of a divalent cation (e.g. Mg^{2+}) for activity (Saleh et al., 2009).

Prenylation of aromatic compounds is critical to the production of a variety of secondary metabolites in plants. Different compounds arise from the prenylation position, length of the prenyl chain and from further modification (e.g. cyclization and hydroxylation) (Tahara and Ibrahim, 1995). The first confirmed plant prenyltransferases [p-hydroxybenzoate:geranyltransferase (*LePGT-1* and *LePGT-2*)] shown to transfer a prenyl moiety to an aromatic substrate were identified *in Lithospermum erythrorhizon* (Yazaki et al., 2002). These enzymes are predicted to have 8 or 9 transmembrane domains and are involved in the biosynthesis of shikonin. The first flavonoid-specific prenyltransferase was naringenin 8-dimethylallyltransferase (SfN8DT-1), which is involved in the production of the flavonoid, sophoraflavanone G (SFG) (Sasaki et al., 2008). It was identified to be a membrane-bound protein with a divalent cation requirement, similar to other UbiA prenyltransferases. SfN8DT-1 localizes to plastids and was predicted to have nine transmembrane α-helices and the prenyltransferase motif (NQxxDxxxD). Furthermore, the expression of *SfN8DT-1* was upregulated in cell culture upon application of yeast extract, methyl jasmonate or salicylic acid, suggesting a potential role in pathogen defense.

1.6.2 Glycinol-4-dimethylallyltransferase

During glyceollin biosynthesis, glycinol is prenylated at either the C-4 or C-2 position to produce the immediate precursor to the glyceollin isomers. Identification and characterization of the C-4 prenyltransferase was performed by Akashi et al. who named the enzyme, glycinol-4-dimethylallyltransferase (G4DT) (Akashi et al., 2009). Similar to other UbiA prenyltransferases, G4DT is predicted to be composed of nine transmembrane domains and the two conserved motifs (important for catalytic activity) are on the cytosolic face of the plastid membrane (Figure 1.2) (Omasits et al., 2014). The enzyme localizes to the plastid, and contains two conserved prenyltransferase motifs: the universally conserved motif (NQxxDxxxD) and another motif specific to flavonoid and homogentisate prenyltransferases (KD(I/L)xDx(E/D)GD) (Akashi et al., 2009). The structure of the protein was predicted using Protter, which displays that it is imbedded in the plastid membrane through its nine transmembrane domains

In the presence of glycinol, DMAPP and Mg^{2+}, G4DT catalyzed the prenylation of glycinol to produces 4-glyceollidin (Figure 1.3). Elicitor-induced phytoalexin production identified that *G4DT* is upregulated starting at 5 to 24 hours after elicitation similar to the expression of other isoflavonoid biosynthesis genes (Akashi et al., 2009).

1.7 Soybean genome

The soybean (*G. max* var. Williams 82) genome contains 1.1 gigabases assembled into 20 chromosomes, with nearly 56,000 protein-coding genes and 89,000 transcripts (Schmutz et al., 2010). Soybean is a palaeopolyploid whose genome has undergone two whole genome duplications at approximately 59 and then 13 million years ago resulting in 75% of genes within the genome in multiple copies (Schmutz et al., 2010).

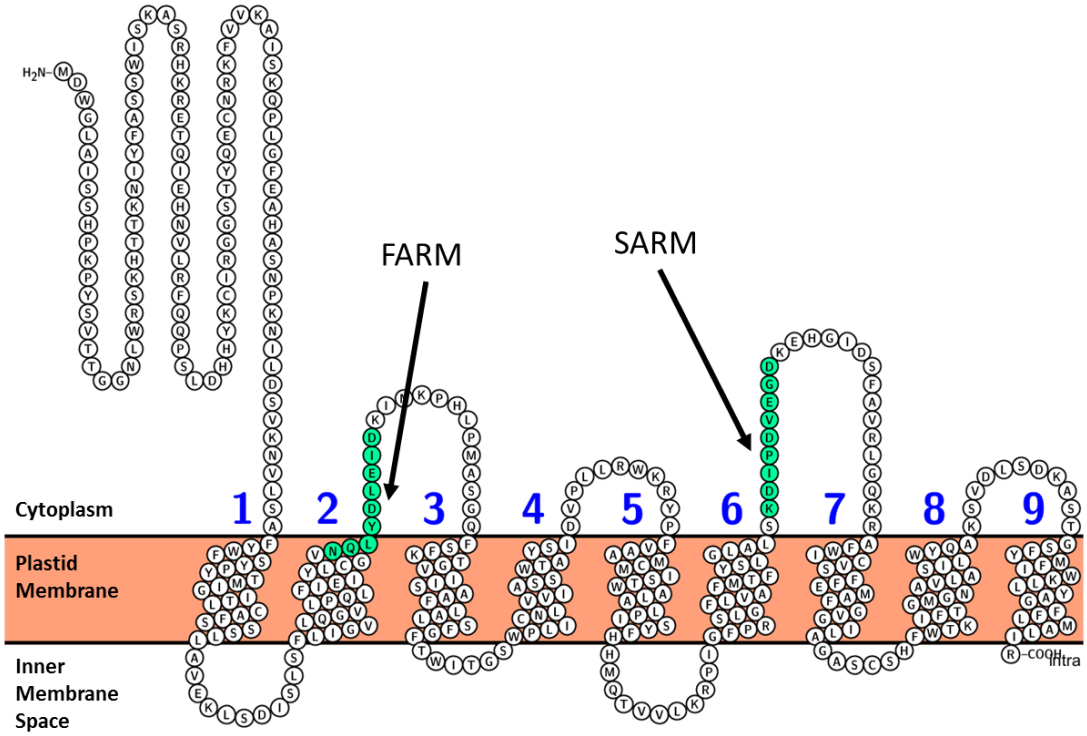


Figure 1.2 Predicted folding of G4DT in the plastid membrane. The protein is shown as an integral plastid membrane protein with nine transmembrane domains, numbered 1 through 9. The two conserved motifs (green residues) correspond to the first and second aspartate-rich motif, FARM and SARM, respectively.

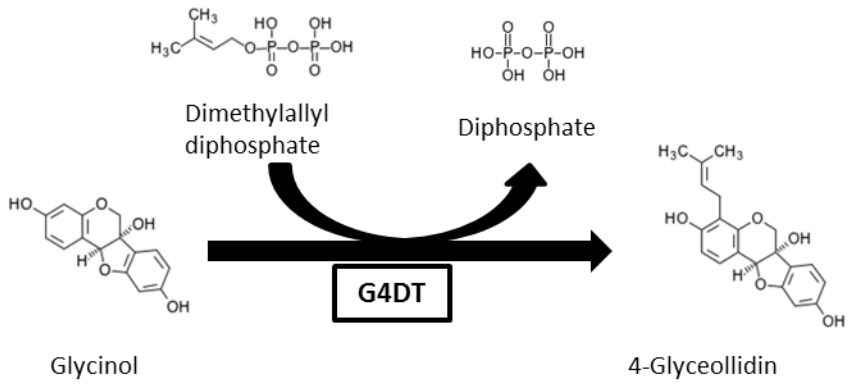


Figure 1.3Preylation of glycinol by glycinol-4-dimethylallyltransferase (G4DT). G4DT transfers a prenyl moiety from DMAPP to the C-4 position of glycinol. This reaction is commonly referred to as a Friedel-Craft alkylation.

1.8 Hypothesis and objectives of the study

Many isoflavonoid biosynthesis genes such as *CHS* (Tuteja and Vodkin, 2008), *CHR* (Sepiol, 2015), *CHI* (Dastmalchi and Dhaubhadel, 2015) and *IFS* (Jung et al., 2000) are members of large gene families. Therefore, I hypothesize that there are multiple prenyltransferase genes responsible for glyceollin biosynthesis in soybean, which are induced upon pathogen infection.

The overall objective of my study is to identify and characterize the gene family encoding prenyltransferases responsible for glyceollin biosynthesis.

To fulfill this objective, I will look to

A. Identify members of the isoflavonoid-specific *GmPT* gene family in soybean
B. Determine the expression profile of candidate *GmPTs* upon pathogen attack
C. Determine the basal tissue-specific expression of candidate *GmPTs*
D. Analyze the promoter region of candidate *GmPTs* genes
E. Determine the subcellular localization of candidate GmPTs
F. Functionally characterize candidate GmPTs through an enzyme assay
G. Identify QTL(s) and QTL markers linked to *P. sojae* resistance that are associated with candidate *GmPTs*

Particular focus will be in characterizing those prenyltransferases that are responsive to pathogen attack and offer to serve as candidates for further genetic manipulation to alter the accumulation of isoflavonoids. Increased isoflavonoid production would offer to strengthen the partial resistance within soybeans, and offer as a long-term mechanism to protect against stem and root rot disease caused by *P. sojae*

Chapter 2: Materials and Methods

2.1 Plant materials and growth conditions

Nicotiana benthamiana were grown on PRO-MIX® BX MYCORRHIZAE™ soil (Rivière-du-Loup, Canada) in a growth chamber. The growth chamber was set to a 16 hour light at 25°C and 8 hour dark at 20°C cycle, with 60-70% relative humidity and light intensity of 100-150 µmol $m^{-2}s^{-1}$. *N. benthamiana* seeds were sprinkled onto the surface of wet soil and grown for one week. At this stage, seedlings were transferred to sterilized pots and watered regularly. Once per week, a nutrient mixture of nitrogen, phosphorous and potassium (20-20-20) was applied.

Soybean (*G. max L. Merr.*) cv. Harosoy63 (H63) seeds were surface sterilized with 70% ethanol containing 3% H_2O_2 for 2 minutes and then rinsed with water prior to planting. Seeds were planted in soaked vermiculite and grown for 6 days in a growth chamber. The growth chamber was set to 25°C with no light and 60-70% relative humidity.

2.2 Bacterial and yeast strains

For transient expression pertaining to determination of subcellular localization, *Escherichia coli* strain DH5α (Invitrogen) and *Agrobacterium tumefaciens* strain GV3101 were used.

For transient protein expression, *E. coli* TOP10F' (Invitrogen) and *Saccharomyces cerevisiae* strain BJ2168 (*MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2*) (ATCC® 208277™) were used.

2.3 *In silico* and phylogenetic analysis

To identify putative prenyltransferases (PTs) involved in isoflavonoid biosynthesis, the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html) was mined to first identify all the soybean prenyltransferases (GmPTs). A keyword search using the words "prenyltransferase" and "dimethylallyltransferase" was performed in the annotated *G. max* Wm82.a2.v1 genome, in the following databases: Panther, Pfam, GO, KOG, and

KEGG. To ensure no *GmPTs* were missed in the keyword search, each PT identified in the soybean genome was used as a query for a Protein BLAST (BLASTP) search.

A phylogenetic tree was constructed by using MEGA 6 (Tamura et al., 2013). The protein sequences were first aligned with ClustalW, prior to constructing a neighbour-joining tree with 1000 bootstrap replications. The presence of transmembrane domains was predicted using TMHMM (Krogh et al., 2001) and the subcellular localization was predicted using WoLF PSORT (Horton et al., 2007). Chloroplast signal peptide presence and cleavage site was predicted using ChloroP (Emanuelsson et al., 1999).

Pairwise nucleotide and amino acid comparison was performed using the percent identity function of ClustalW.

2.3.1 Promoter analysis

To identify promoter elements in the candidate genes, the region 1000 bp upstream of the translational start site of each candidate *GmPT* was acquired from the soybean genome on Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). Analysis was performed by MatInspector version 8.0 (Catharius et al., 2005) in the Genomatix suit of programs (http://www.genomatix.de; Genomatix Software, Munich, Germany).

2.4 Pathogen infection and $AgNO_3$ treatment

Seven-day old *G. max* cv. L76-1988 plants were inoculated with *P. sojae* Race 7 or water for control. The inoculated stems of these plants were collected at 24, 48 and 72 hours post-innoculation.

$AgNO_3$ has been reported to induce plant defence mechanisms, and therefore, it was used to mimic pathogen attack (Kubeš et al., 2014). $AgNO_3$ treatment of soybean was performed according to Ward et al. (1979). Briefly, *G. max* cv. H63 were grown in water soaked vermiculite at 25°C in the dark. After 6 days, the etiolated seedlings were transferred from pots to glass trays as shown in Figure 2.1. Five droplets (10 µL each) of either water (control) or 1 mM $AgNO_3$ were placed on the hypocotyl of each seedling. The roots were wrapped in water-soaked cheese cloth and the glass trays were sealed to ensure that the seedlings do not dehydrate. Glass trays containing the treated seedlings were stored in the dark at 25°C and hypocotyl tissue was collected at 6, 12, 24, 48 and 72 hours post-treatment.

Figure 2.1 $AgNO_3$ treatment of soybean hypocotyls. Soybean cv. H63 was germinated in the dark at 25°C for 6 days in water-soaked vermiculite prior to transfer to glass trays. Seedlings were placed onto the place tray and held in place with plastic brackets. The roots were wrapped in water-soaked cheese cloth to prevent dehydration. The hypocotyls were treated with either water (control) or with 1 mM $AgNO_3$ (treatment).

2.5 RNA extraction, RT-PCR and qPCR

Total RNA was extracted from 100 mg of stem/hypocotyl tissue using the RNeasy Plant Mini kit (Qiagen). On-column DNA digestion was performed using DNase I (Promega). RNA was quantified with a NanoDrop 1000 spectrophotometer (ThermoScientific), and evaluation of quality was performed analyzing its A_{260}/A_{280} ratio. Further quality confirmation was performed by gel electrophoresis of the extracted RNA on a RedSafe (iNtROn Biotechnology) strained 1% agarose gel in 1X TAE buffer.

Total RNA from field grown soybean tissues (roots, stems, leaves, flower buds, flowers, embryos [30, 40, 50, 60 and 70 DAP], seed coats and pod walls) was extracted following a protocol adapted from Wang and Vodkin (1994). Potential DNA contaminants were removed using the TURBO DNA-free™ kit (Life Technologies). RNA quantity and quality was determined as mentioned above. DNase-treated RNA (1 µg) was used to synthesize cDNA through the ThermoScript™ RT-PCR System (Life Technologies).

For reverse-transcription polymerase chain reaction (RT-PCR), reactions were set using gene-specific primers (Table 2.1) for amplification. *CONS4* was used as a loading control.

For real-time PCR (qPCR), SsoFast EvaGreen Supermix kit (BioRad) and gene-specific qPCR primers were used (Table 2.1). Reactions were analyzed in a Bio-Rad C1000 Thermal Cycler with the CFX96™ Real-Time PCR System. All reactions were performed as three technical replicates. A reaction without template, no template control (NTC), was used as a negative control. *CONS4* was used as a reference gene to normalize gene expression. The data were analyzed using CFX manager (BioRad).

2.6 Statistical analysis

Statistical differences in gene expression were determined by a Two-Way ANOVA calculated using SAS® 9.1 for Windows. Statistical significance was set at $p < 0.05$.

2.7 Cloning and plasmid construction

For subcellular localization, gateway compatible primers were designed with *attB1* adaptor sequence (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT-3') for the forward primer and *attB2* adaptor sequence (5'-GGGG AC CAC TTT GTA CAA GAA AGC TGG GT-3') for reverse primer to allow for recombination into Gateway entry vector, pDONR/Zeo (Invitrogen) (Table 2.1). PCR amplification using gene-specific primers was performed to amplify full length *GmPTs*. PCR products were run on a RedSafe (iNtROn Biotechnology) strained 1% agarose gel in 0.5X TBE buffer. Gels were subsequently visualized on a Bio-Rad Gel Doc. PCR amplicons were extracted using EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc.) and quantified using NanoDrop 1000 spectrophotometer (ThermoScientific). Purified amplicons were run on a RedSafe (iNtROn Biotechnology) strained 1% agarose gel in 0.5X TBE buffer to confirm successful purification.

The purified PCR amplicon was recombined into pDONR/Zeo using Gateway BP Clonase® II Enzyme mix (Invitrogen). The BP reaction was then transformed into *E. coli* strain DH5α and then plated on a lysogeny broth (LB) agar plate containing zeocin (50 μg/mL) to allow for selection. Transformation of bacterial cells was carried out via electroporation for 4-6 milliseconds at 1.8 kV in a Gene Pulser® Cuvette (Bio-Rad Laboratories, Inc.) with a 0.1 cm electrode gap by MicroPulser^{TM} (Bio-Rad Laboratories, Inc.). Positive colonies were ascertained by performing colony PCR using vector-specific M13F and M13R primers. The positive colonies were grown overnight for plasmid DNA (pDONR/Zeo-GmPT) extraction using the EZ-10 Spin Column Plasmid DNA kit (Bio Basic Inc.), quantified using NanoDrop 1000 spectrophotometer (ThermoScientific) and sequence confirmed. All sequencing of DNA constructs was carried out at the London Research and Development Center, Agriculture and Agri-Food Canada in London, Ontario.

The pDONR/Zeo-GmPT constructs were recombined into destination vector pEarleyGate101 (pEG101) using Gateway LR Clonase® II Enzyme mix (Invitrogen), which would produce a translational fusion of each GmPT fused to yellow fluorescent protein (YFP). The LR reaction was then transformed into *E. coli* strain DH5α by

electroporation and then plated on a LB agar plate containing kanamycin (50 μg/mL) to allow for selection. Positive colonies were ascertained by performing colony PCR using gene-specific primers (Table 2.1). Plasmid DNA was extracted using the EZ-10 Spin Column Plasmid DNA kit (Bio Basic Inc.) and quantified using NanoDrop 1000 spectrophotometer (ThermoScientific).

The pEG101-GmPT constructs were transformed into *A. tumefaciens* GV3101 using electroporation. Transformation of bacterial cells was carried out via electroporation as described previously at 2.18 kV. The transformed *Agrobacterium* was plated on a LB agar plates containing rifampicin (10 μg/mL), gentamycin (50 μg/mL), and kanamycin (50 μg/mL) to allow for selection. Colony PCR was performed using gene-specific primers (Table 2.1) to verify whether the *Agrobacterium* colony contained the pEG101-*GmPT* construct.

For protein expression, truncated *GmPTs* were amplified lacking their signal peptides using primers listed in Table 2.1. The truncated *GmPT* were then ligated into yeast expression vector, pYES2.1/V5-His-TOPO® (Invitrogen) and transformed into *E. coli* TOP10F' cells (Invitrogen) using the pYES2.1 TOPO® TA Expression Kit (Invitrogen). Transformation of *E. coli* TOP10F' was carried out according to the manufacturer's instructions with slight modifications. Selection of transformants was performed using the antibiotic, ampicillin and positive colonies was ascertained by colony PCR. Purified plasmid constructs from positive colonies were sent for sequencing to ensure sequence identity and insert orientation. Upon confirmation, plasmid was transformed into *S. cerevisiae* BJ2168 and screened on minimal SD/-ura plates (Clontech). Preparation of competent *S. cerevisiae* BJ2168 cells and subsequent transformation of set cells were carried out using the Frozen-EZ Yeast Transformation II™ kit (Zymo Research).

E. coli, A. tumefaciens and *S. cerevisiae* were grown at 37°C, 28°C and 30°C, respectively.

Table 2.1 List of primers used for subcellular localization, qPCR and protein expression.

Gene	Primer Name	Sequence (5'→3')	Amplicon size (bp)
Subcellular Localization			
GmPT01	GmPT01F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGATTCGGGGCTTGCTATATC	1179
	GmPT01R	GGGGACCACTTTGTACAAGAAAGCTGGGTATCTAATTAAAGCCACGAGGAAGTATGATG	
GmPT03	GmPT03F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGCTTCATTGATTCTTATATCTTCTCC	1239
	GmPT03R	GGGGACCACTTTGTACAAGAAAGCTGGGTAATGACATCCTTGGCTAATTAAAGGC	
GmPT10a	GmPT10aF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGATTGGGGGCTTGCTATA	1230
	GmPT10aR	GGGGACCACTTTGTACAAGAAAGCTGGGTATCTAATTAATGCCATGAGAAAGAACCC	
GmPT10b	GmPT10bF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGAATACAATCTTTTGAGGTTCCA	1080
	GmPT10bR	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCTGATTAAAGGTATGAGGAAGTATGATG	
GmPT10d	GmPT10dF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATCGTGCCCTTGTTATGTCTT	1209
	GmPT10dR	GGGGACACTTTGTACAAGAAAGCTGGGTCTCTAATGAAAGGTAAGAGAAAGTACGCTGC	
GmPT11a	GmPT11aF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGATTCGATGCATGTTATTATGTC	1227
	GmPT11aR	GGGGACCACTTTGTACAAGAAAGCTGGGTATCTGATTAAAGGCATGAGGAAGTATGC	
GmPT11b	GmPT11bF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGTTCGGTGCCTCTTTTTTC	1223
	GmPT11bR	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCTAATGAAAGCCAAGAGAAAGTACGCT	
GmPT11c	GmPT11cF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGTATGGAGGGCTTGTTATGTCTT	1227
	GmPT11cR	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCTGATTAAAGGCATGAGGAGGTATG	
GmPT20	GmPT20F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGATTCGGGGTCTGTTATATC	1227
	GmPT20R	GGGGACCACTTTGTACAAGAAAGCTGGGTATCTAATTAAAGCCATGAGAAAGTATGCT	
qPCR			
GmPT01	PT01-QF1	CGGCAATGTGCATGGTTTCA	159
	PT01-QR1	TTGCCAATGCCATACCCACA	

GmPT10a	PT10a-QF1	AATCGCTTCATCGTGGACGG	157
	PT10a-QR1	TGGCCTCTTCAACACAACGG	
GmPT10d	PT10d-QF2	GCCCCTGCTATTCCATCTTTTG	159
	PT10d-QR2	TGCAAGGAGAGACGCGGAAA	
	PT10d-QF3	GGTGGGAAATCATATCGTAAGTGGT	132
	PT10d-QR3	GTTGTATCCTGTTTTCCACCGTCT	
GmPT11a	PT11a-QF2	AAATGGCTTTCGGAGTTGCC	155
	PT11a-QR2	GGATCTAGTGGAAGCTGGGC	
	Pt11a-QF3	TGTAGGTGGAAATCTCTGGCGC	159
	Pt11a-QR3	TATGTAGGCCCTCCTTCATTGC	
GmPT20	PT20-QF1	CCTTGTGCTTCCAATCCCGA	144
	PT20-QR1	TGATGGTTTCTCCACCGCAA	
Protein Expression			
GmPT01	PT01-d21F	ATGGCATCAAAAGCTTCGCAGCACAC	1116
	PT01-R-Stop	TTATCTAATTAAAGCCACGAGGAAGTATGA	
GmPT10a	PT10a-d44F	AAAATGCACAAAAGGGAAACTCAAATAGAACATAATG	1086
	PT10a-d41F	ATGGCTTCACGACACAAAAGGGAAACT	1110
	PT10a-R-Stop	TCCTCATCTAATTAATGCCATGAG	
GmPT10d	PT10d-d14F	ATGGTCACAAATGGTGGAAATCTCTGG	1167
	PT10d-R-Stop	GCGTACTTTCTCTTACCTTTCATTAGATAA	
GmPT11a	PT11a-d11F	ATGGCTGGTTCAGTTACAACTGGT	1233
	PT11a-R-Stop	TTATCTAATGAAAGCCAAGAGAAAGTACGC	
GmPT20	PT20-d36F	ATGTGTGCGTCAAAAGCTTCGAAATACA	1122
	PT20-R-Stop	TCCTTATCTAATTAAAGCCATGAGAAAGTA	

2.8 Subcellular localization

2.8.1 Transient expression of GmPTs in *N. benthamiana*

A single *A. tumefaciens* colony containing a pEG101-*GmPT* construct was grown at 28˚C in infiltration culture media (LB broth containing 10 mM 2-N-morpholino-ethanesulfonic acid (MES) pH 5.6 and 100 μM acetosyringone with kanamycin (50 μg/mL), rifampicin (10 μg/mL), and gentamycin (50 μg/mL)) until an OD_{600} of 0.5 - 0.8 was reached. The *Agrobacterium* was pelleted by centrifugation at 3000 rpm for 30 minutes, and then resuspended in Gamborg's solution (3.2 g/L Gamborg's B5 and vitamins, 20 g/L sucrose, 10 mM MES pH 5.6, and 200 μM acetosyringone) to obtain a final OD_{600} of 1.0. To activate the virulence genes, *Agrobacterium* possessing the pEG101-*GmPT* constructs were incubated at room temperature for 1 hour with gentle agitation.

Leaves of 4-6 week old *N. benthamiana* were infiltrated with bacterial culture by placing the tip of the syringe barrel against the underside of the leaf and applying moderate pressure. To validate the subcellular localization, the *Agrobacterium* infiltration culture containing pEG101-GmPT would be co-infiltrated with a plastid organelle marker translationally fused to cyan fluorescent protein (CFP) at a 1:1 ratio. Plants were then returned back to the growth room set to normal conditions as described in section 2.1.

2.8.2 Confocal microscopy

Following the 48 hour incubation, the epidermal cell layers of the infiltrated *N. benthamiana* leaves were visualized using a Leica TCS SP2 inverted confocal microscope using a 63X water immersion objective lens. For YFP visualization, the excitation wavelength was set to 514 nm and emission was collected at 530-560 nm. For CFP visualization, the excitation wavelength was set to 434 nm and emission was collected at 470-500 nm. To visualize the co-localization of the YFP and CFP signals, the 'Sequential Scan' tool was utilized.

2.9 Yeast microsome preparation

Protein induction for microsomal preparation was performed following the protocols used by Akashi et al. (2009) and Endo et al. (2011) with some modifications. A seed culture was set using a single colony of *S. cerevisiae* BJ2168 containing the pYES2.1-GmPT construct in 3 mL synthetic dextrose (SD)/-ura media and grown overnight at 30°C. The seed culture was then used to inoculate a 250 mL SD/-ura culture, and then was grown under similar condition until an OD_{600} of 0.5-1 was reached. The cells were then harvested by centrifugation, and protein induction was performed by growing the cells in 250 mL of SLI broth containing 2% galactose. The cultures were grown at 30°C for 24 hours.

Yeast cells were harvested and then resuspended in 25 mL of potassium phosphate buffer (KPB) (pH 7.6). Cells were homogenized using a French Laboratory Press set to 1000-1500 psi. Cell lysates were then centrifuged at 10,000 × *g* to remove cell debris. The supernatant was subjected to ultracentrifugation at 100,000 × *g* for microsome preparation. Microsomes were resuspended in KPB buffer and quantitated by Bradford assay (Bradford, 1976).

2.10 Glycinol production

Production of glycinol was adapted from protocols developed by Akashi et al. (2009) and Boue et al. (2009). *G. max* cv. H63 seeds were surface sterilized with 70% EtOH followed by soaking in water for 4-5 hours prior to treatment. Soybean seeds were halved and treated with $AgNO_3$ or water (control). Seeds were treated with a combination of $AgNO_3$ and isoprenoid pathway inhibitors, lovastatin (Sigma Aldrich) and fosmidomycin (Fisher Scientific). Samples were incubated at 25°C in the dark in "treatment chambers". Chambers consisted of a petri dish (100x15mm) lined with two layers of autoclaved filter paper and 2.5 g of seeds.

To determine the optimal incubation time, seeds were imbibed for 4-5 hours before transfer to the chambers. Seeds were cut in half and treated with 3 mL of 10 mM $AgNO_3$

or water and stored at 25°C in the dark. Samples were collected at 24, 48 and 72 hours post-treatment.

To determine the effectiveness and optimal concentration of isoprenoid inhibitors, seeds were imbibed for 4 hours before transfer to the chambers. Seeds were cut in half and treated with 10 mM $AgNO_3$ plus varied concentration of inhibitors (100 µM, 250 µM, 500 µM and 1 mM) or water and stored at 25°C in the dark.

Extraction was performed by grinding the seeds to a fine powder and immersing the powder in methanol (2 mL/g tissue). Samples were thoroughly vortexed, followed by sonication for 15 minutes in a Cole Palmer 8893 Ultrasonic cleaner. Following sonication, the sample was centrifuged to eliminate plant debris. The supernatant was isolated and concentrated using a rotary evaporator, then the dried material was resuspended in methanol. The concentrated sample was analyzed by high performance liquid chromatography (HPLC). Initial screening was performed with an Agilent 1100 HPLC System equipped with a multi-wavelength diode array detector, using a Poroshell 120, EC-C18 column (4.6 x 100 mm, 2.7 µm; Agilent). The column temperature was maintained at 32°C. The injection volume was 10 µL at a flow rate of 1 mL/min. Samples were run under the following solvent system: A) water + 0.1% formic acid (FA), B) acetonitrile (ACN) + 0.1% FA. The gradient system was as follows: 0% B for 2 min, then a linear gradient to 100% B in 13 min, followed by a 2 min wash with 100% B and then 3 min of equilibration at 0% B prior to the next injection. Spectra were detected at 210, 254 and 280 nm.

The solvent system was adapted to prevent targeted metabolite co-elution. The revised solvent system was: A) water + 0.1% FA, B) ACN + 0.1% FA. The gradient system was as follows: 15% B for 2 min, then a linear gradient to 50% B in 15 min, followed by a linear gradient to 100% B in 4 min, then a 2 min wash with 100% B, then back to 15% B in 1 min, and 2 min of equilibration at 15% B prior to the next injection. Spectra were detected at 210, 254 and 280 nm.

Glycinol isolation was performed using an Agilent 1200 HPLC system with a semipreparative Gemini® C18 110 Å column (150 x 4.6 mm, 5 µm; Phenomenex). The

column temperature was maintained at 32°C. The injection volume was 100 µL at a flow rate of 5 mL/min in the following solvent system: A) water B) ACN. The gradient system was as follows: 0-2 min, 15% B; 2-17 min, 50% B; 17-21 min, 100% B; 21-23 min, 100% B; 23-24 min, 15% B; 24-26 min, 15%B prior to the next injection. Spectra were detected at 254, 280 and 285 nm.

Putative glycinol confirmation was performed by UV-visible spectrophotometry and mass spectrometry.

2.11 Enzyme assay

Enzyme assay conditions used in this study were: 40 µg of recombinant yeast microsome in 250 µL of KPB incubated with 400 µM glycinol, 400 µM DMAPP (Sigma-Aldrich) and 10 mM $MgCl_2$. For a control the following reaction was set in the absence of glycinol. Each reaction was incubated at 30°C for 10 min, followed by ethyl acetate extraction. The solvent was evaporated using N_2 gas, and the dried pellet was resuspended in 100 µL methanol. The reactions were analyzed using an Agilent 1100 HPLC system equipped with a multi-wavelength diode array detector using a Poroshell 120, EC-C18 column (4.6 x 100 mm, 2.7 µm; Agilent). The column temperature was maintained at 32°C. The injection volume was 10 µL at a flow rate of 1mL/min. Samples were run under the following solvent system: A) water + 0.1% FA, B) ACN + 0.1% FA. The gradient system was as follows: 15% B for 2 min, then a linear gradient to 34% B in 8 min, followed by a linear gradient to 100% B in 0.5 min, then a 2.5 min wash with 100% B, then back to 15% B in 0.5 min, and 2.5 min of equilibration at 15% B prior to the next injection. Spectra were detected at 210, 254 and 280 nm.

HPLC grade acetonitrile was purchased from Caledon. Water was obtained from a Millipore system.

2.12 QTL analysis

To determine whether any candidate *GmPT* is located within any QTLs associated with *P. sojae* resistance, previously reported QTLs and QTL markers were identified from literature. Furthermore, the SoyBase and Soybean Breeder's Toolbox database (Grant et

al., 2010) was mined. The relative position of QTLs and QTL markers were compared with the transcribed *GmPTs*.

Chapter 3: Results

3.1 The soybean (iso)flavonoid prenyltransferase gene family contains 11 members

Identification of *GmPTs* was performed through keyword searches within the annotated *G. max* Wm82.a2.v1 genome on the Phytozome database. Using the terms "dimethylallyltransferase" and "prenyltransferase", 63 putative *GmPTs* were identified. Each identified *GmPT* was then used as a query for a BLASTP search to ensure no *GmPTs* was missed by the keyword search, but no additional *GmPTs* were identified.

Prenyltransferases are involved in numerous biosynthetic pathways within plants such as homogentisate, ubiquinone, shikonin and flavonoid biosynthesis (Li, 2016). To identify *GmPTs* involved specifically with isoflavonoid biosynthesis, a phylogenetic analysis was performed. A neighbour joining tree was produced from the amino acid sequences of 63 putative GmPTs and previously characterized PTs from various plant species (Figure 3.1). Out of 63, 17 putative GmPTs were divided across various known clades such as: flavonoids, homogentisate (tocotrienol), homogentisate (tocopherol), homogentisate (plastiquinone), phloroglucinol, PHB (shikonin), PHB (ubiquinone), whereas 46 did not cluster with any of the previously known groups. The flavonoid clade contained 11 putative GmPTs, which clustered with previously characterized flavonoid PTs such as N8DT from *Sophora flavescens* (Sasaki et al., 2008) and PT1 from *Lupinus albans* (Shen 2012).

The candidate gene, *GmPT10a* (Glyma.10G295300) encodes the previously identified and characterized G4DT (Akashi et al., 2009). The list of putative *GmPTs* with their coding sequence length, number of splice variants, predicted number of transmembrane domains (TMHMM v. 2.0), and molecular mass (ExPASy) are shown in Table 3.1. Pairwise transcript sequence comparison between candidate genes range from 71 to 95% identity whereas, pairwise amino acid sequence comparison range from 56 to 92% (Table 3.2)

The structure of UbiA PTs has been shown to be an integral membrane protein with 7-9 transmembrane α-helices (Ohara et al., 2009). All candidate GmPTs were predicted to contain 8 or 9 transmembrane domains except for GmPT08 and GmPT10c that had five predicted domains, which did not fit the criteria of other UbiA prenyltransferases.

3.1.1 Conservation of catalytic sites in candidate GmPTs

As mentioned in section 1.6.1, all UbiA prenyltransferases contain at least one aspartate-rich motif (Saleh et al., 2009). Additionally, flavonoid/homogentisate prenyltransferases possess a second conserved sequence (Akashi et al., 2009). Structural and mutagenesis studies of prenyltransferases have identified residues critical for enzymatic activity (Bräuer et al., 2008; Ohara et al., 2009).

To ensure that the 11 candidate GmPTs potentially possess catalytic activity, a critical residue search was performed. The amino acid sequence of the 11 candidate GmPTs was aligned with other previously characterized PTs (Figure 3.2). Search for the aspartate-rich motifs within the sequences identified that GmPT10c was missing an aspartate residue at the 310 position (D310R). When this residue was mutated in LePGT1, the enzyme exhibited extremely weak activity (Ohara et al., 2009). Within GmPT20, a glutamate residue at the 305 position, which is conserved in the other 10 GmPTs was substituted with a lysine (E305K), however, this residue has not been proven to be essential for enzyme activity.

Since GmPT08 and GmPT10c contain a fewer number of transmembrane domain to qualify as UbiA prenyltransferase and GmPT10c lacked the critical residue for enzymatic activity, these two candidates were eliminated as potential isoflavonoid-specific prenyltransferases.

Figure 3.1 Phylogenetic analysis of soybean prenyltransferases. The protein sequences of the soybean prenyltransferases and other known prenyltransferases were aligned using ClustalW, following by construction of a neighbour-joining tree with 1000 bootstrap replications using MEGA 6 software. G4DT is highlighted within the black rectangle. Colours indicate separate clades as labelled. ApVTE2-1 (*Allium porrum*; DQ231057); AtVTE2-1 (*Arabidopsis thaliana*, AY089963); AtVTE2-2 (*A. thaliana*, DQ231060); CpVTE2-1 (*Cuphea pulcherrima*, DQ231058); CrHST (*Chlamydomonas reinhardtii*, AM285678); GmVTE2-1 (*Glycine max*, DQ231059); GmVTE2-2 (*G. max*, DQ231061); HI-PT1 (*Humulus lupulus*, AB543053); HvHGGT (*Hordeum vulgare*, AY222860); LaPT1 (*Lupinus albus*, JN228254); LePGT1 (*Lithospermum erythrorhizon*, AB055078); LePGT2 (*L. erythrorhizon*, AB055079); OsHGGT (*Oryza sativa*, AY222862); OsPPT1 (*O. sativa*, AB263291); SfiLDT (*Sophora flavescens*, AB604223); SfL17a (*S. flavescens*, AB371287); SfL17b (*S. flavescens*, AB370329); SfN6DT (*S. flavescens*, BAK52291); SfN8DT-1 (*S. flavescens*, AB325579); SfN8DT-2 (*S. flavescens*, AB370330); SfN8DT–3, (*S. flavescens*, AB604222); TaHGGT (*Triticum aestivum*, AY222861); TaVTE2-1 (*T. aestivum*, DQ231056); ZmVTE2-1(*Zea mays*, DQ231055)

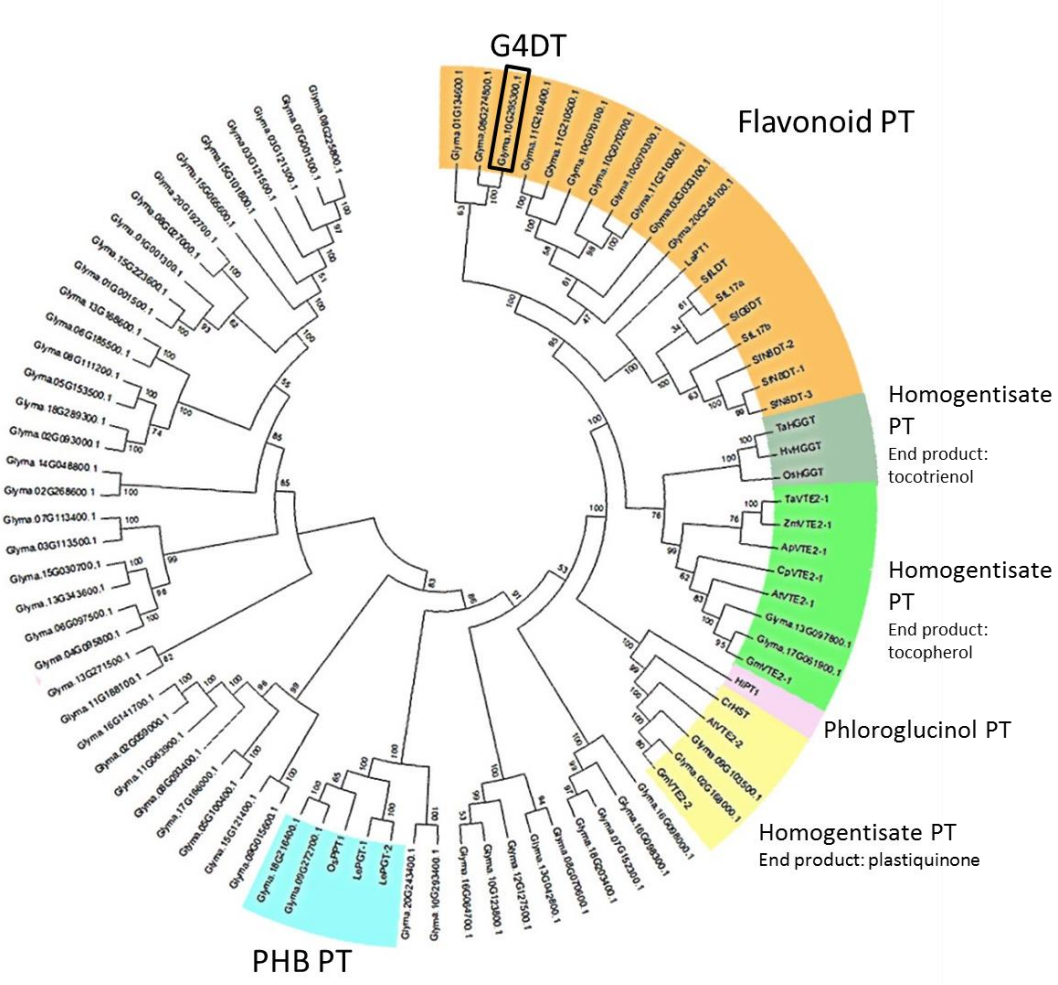

G4DT
Flavonoid PT
Homogentisate PT
End product: tocotrienol
Homogentisate PT
End product: tocopherol
Phloroglucinol PT
Homogentisate PT
End product: plastiquinone
PHB PT

Table 3.1 Characteristics of candidate *GmPTs*

Gene Name	Locus Name	Predicted Mol. Weight (kDa)	Coding Sequence Length (bp)	Splice Variants	# of Transmembrane Domains
GmPT01	Glyma.01G134600	43.66	1182	6	9
GmPT03	Glyma.03G033100	46.72	1239	1	9
GmPT08	Glyma.08G274800	38.17	1029	1	5
GmPT10a	Glyma.10G295300	46.02	1230	1	9
GmPT10b	Glyma.10G070100	41.07	1080	1	9
GmPT10c	Glyma.10G070200	37.47	897	1	5
GmPT10d	Glyma.10G070300	45.5	1209	1	8
GmPT11a	Glyma.11G210300	46.32	1266	3	9
GmPT11b	Glyma.11G210400	46.32	1233	3	9
GmPT11c	Glyma.11G210500	46.24	1227	1	9
GmPT20	Glyma.20G245100	45.84	1227	1	9

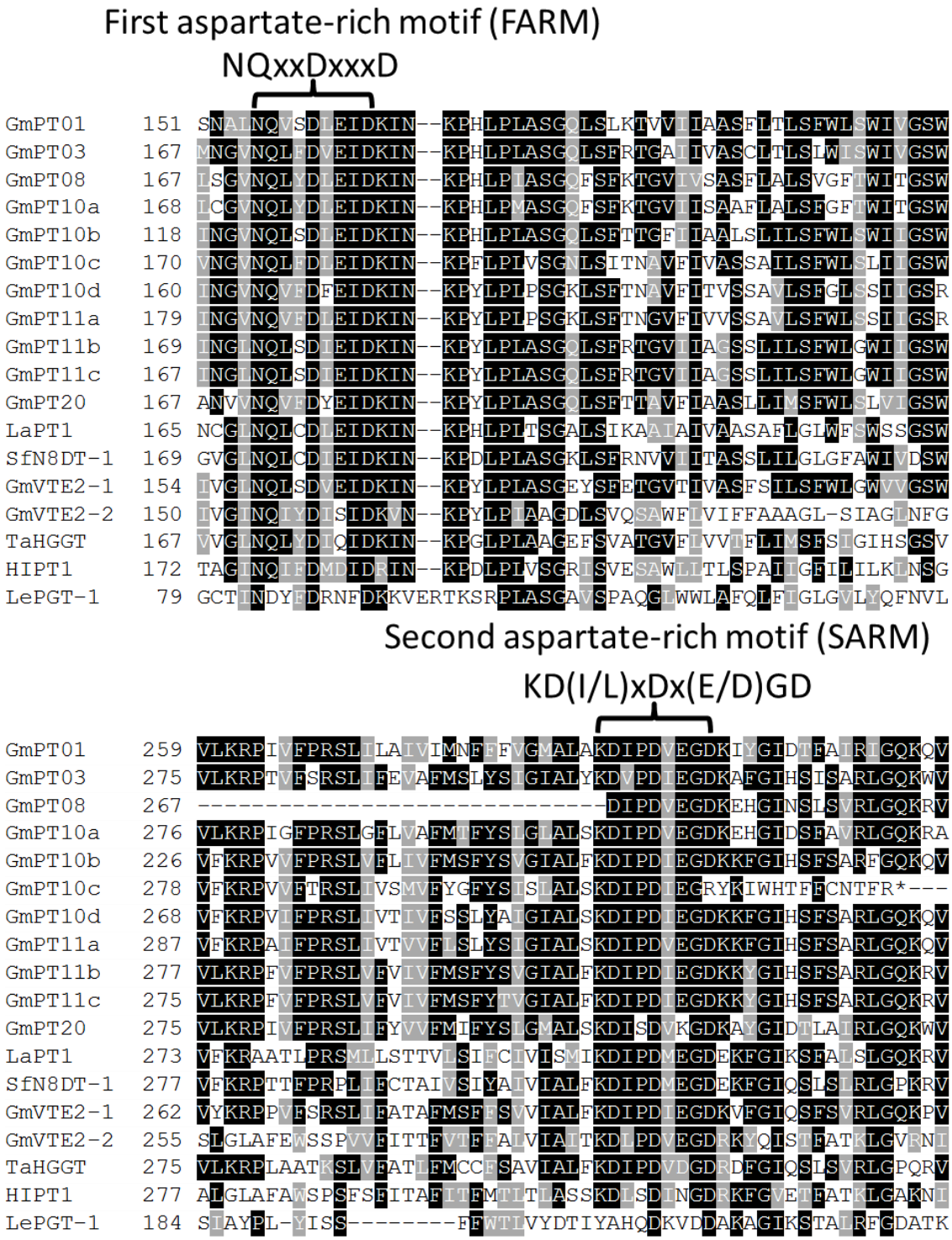
bp: base pairs, kDa: kilo-Dalton.

Table 3.2.2 Pairwise comparison of nucleotides and amino acid sequences of isoflavonoid-specific candidate GmPTs

		Protein										
		GmPT01	GmPT03	GmPT08	GmPT10a	GmPT10b	GmPT10c	GmPT10d	GmPT11a	GmPT11b	GmPT11c	GmPT20
Nucleotide	GmPT01		61.30	65.62	63.28	60.72	57.48	57.47	57.73	61.40	60.84	63.80
	GmPT03	74.15		65.40	63.48	69.08	62.15	64.60	66.50	69.36	69.88	66.18
	GmPT08	75.87	78.19		78.95	59.39	55.94	55.94	58.70	60.70	61.54	61.88
	GmPT10a	75.45	78.14	86.58		61.56	56.00	57.62	59.85	62.01	62.72	63.97
	GmPT10b	75.28	81.02	74.69	76.67		64.13	72.27	69.75	83.01	84.12	63.79
	GmPT10c	71.15	77.42	74.07	74.21	79.28		70.49	67.28	64.00	62.73	60.62
	GmPT10d	74.41	79.76	76.02	76.48	81.28	83.89		86.75	70.28	70.57	62.53
	GmPT11a	73.57	79.40	74.75	76.08	80.84	80.13	90.11		67.49	68.24	63.79
	GmPT11b	74.94	81.34	75.28	76.48	89.35	77.99	80.98	80.33		91.89	67.16
	GmPT11c	74.77	81.08	76.80	77.66	89.81	78.21	81.41	79.57	94.70		66.42
	GmPT20	77.24	79.49	78.24	79.05	78.43	75.84	78.89	77.16	79.34	79.93	

Nucleotide (bottom triangle) and amino acid (top triangle) sequence identity (%) between eleven candidate prenyltransferases; nucleotide sequences corresponding to the coding region and the amino acid sequence for each gene was acquired the Phytozome *G. max* genome; alignment of sequence and percentage identities calculations were performed by ClustalW.

Figure 3.2 Identification of catalytic site residues in isoflavonoid-specific candidate GmPTs. Multiple sequence alignment was performed using ClustalW to align candidate GmPTs with previously characterized PTs. Abridged version of alignment is shown to highlight both the first aspartate-rich motif (FARM) and second aspartate-rich motif (SARM).


First aspartate-rich motif (FARM)
NQxxDxxxD
Second aspartate-rich motif (SARM)
KD(I/L)xDx(E/D)GD

3.2 *GmPT* expression analysis

Isoflavonoid biosynthetic gene expression and glyceollin production was induced by pathogen infection (Ayers et al., 1976; Ward et al., 1979; Kimpel and Kosuge, 1985; Bhattacharyya and Ward, 1986). *GmPT* gene expression is expected to be upregulated in conjunction with biotic or abiotic stress. This point has already been verified for G4DT (referred to in this study as *GmPT10a*), which was upregulated following elicitation by yeast extract (Akashi et al., 2009). Therefore, the expression of the other candidate *GmPTs* upon pathogen infection was determined.

3.2.1 Candidate *GmPTs* are induced during *P. sojae* infection

To identify isoflavonoid-specific *GmPTs*, *G. max* cv. L76-1988 stems were infected with *P. sojae* Race 7 and samples were collected at 24, 48 and 72 hours post-inoculation. Gene expression analysis was performed by RT-PCR of the nine candidate *GmPTs* (Figure 3.3). Comparison of gene expression in control and *P.* sojae-infected stems identified five *GmPTs (GmPT01, GmPT10a, GmPT10d, GmPT11a* and *GmPT20*) that showed induced expression upon infection.

GmPT01, GmPT10a and *GmPT20* all displayed similar expression profiles. They were expressed under control conditions, but were upregulated in *P.sojae*-inoculated samples, increasing from 24 to 72 hours. *GmPT10d* and *GmPT11a* showed weak to no expression under control conditions. However, it showed gradual increase in expression from 24 to 72 hours.

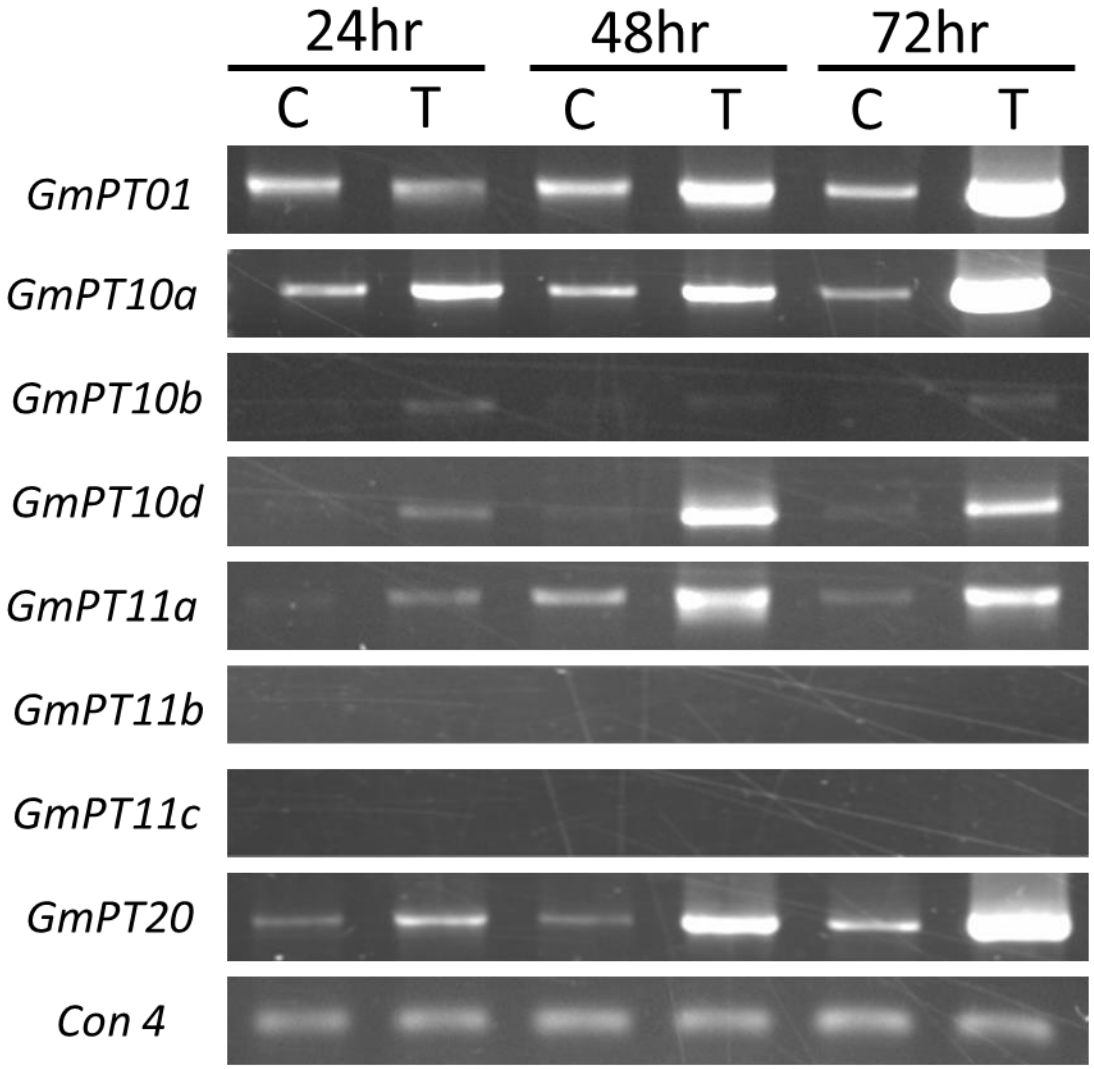


Figure 3.3 Gene expression of *GmPTs* in response to *P. sojae* infection. RT-PCR using total RNA extracted from either *P. sojae*-inoculated (T) or water-treated (C) soybean stems was performed. *CONS4* was used as a loading control.

3.2.2 $AgNO_3$-treatment mimics expression pattern of *P. sojae* infection

Silver nitrate has been used previously to induce glyceollin production (Stössel, 1982), and thus has been used to mimic plant response to pathogen attack (Bhattacharyya and Ward, 1986; Moy et al., 2004). Therefore, I treated etiolated soybean hypocotyls with 10 mM $AgNO_3$ or water (control) and samples were collected at 6, 12, 24, 48 and 72 hours post-treatment. Gene expression was performed by RT-PCR of the 9 candidate *GmPTs*. Upon comparison of control and $AgNO_3$-treated hypocotyls, as expected, I observed that the same genes induced by *P. sojae* infection were induced by $AgNO_3$ treatment (Figure 3.5). Furthermore, the $AgNO_3$-treated hypocotyls possessed brown lesions at the site of $AgNO_3$ treatment, which is physiologically similar to *P. sojae* infection symptoms (Figure 3.4) (Ward et al., 1979). Therefore, this experiment corroborated the use of $AgNO_3$ to mimic pathogen infection.

Expression of the $AgNO_3$-induced genes was further quantified by qPCR (Figure 3.5). *GmPT01* gene expression peaked early and plateaued across the time course. *GmPT10a*, *GmPT10d* and *GmPT11a* were all rapidly induced and began to decline 12 hours post-treatment. *GmPT20* gradually increased over the time course, with peak expression occurring at 48 hours and was still present at high levels at 72 hours. Two-way ANOVA performed using SAS software demonstrated that treatment of $AgNO_3$ had a significant effect on the gene expression of all candidate genes except *GmPT20*. Furthermore, the interaction of time and treatment had a significant effect on *GmPT10a* and *GmPT10d* gene expression.

GmPT10b, *GmPT11b* and *GmPT11c* were not induced by either biotic or abiotic stress. As a result of an inability to be induced upon stress, these genes are unlikely to play a role in glyceollin production, and were therefore, removed as candidate genes.

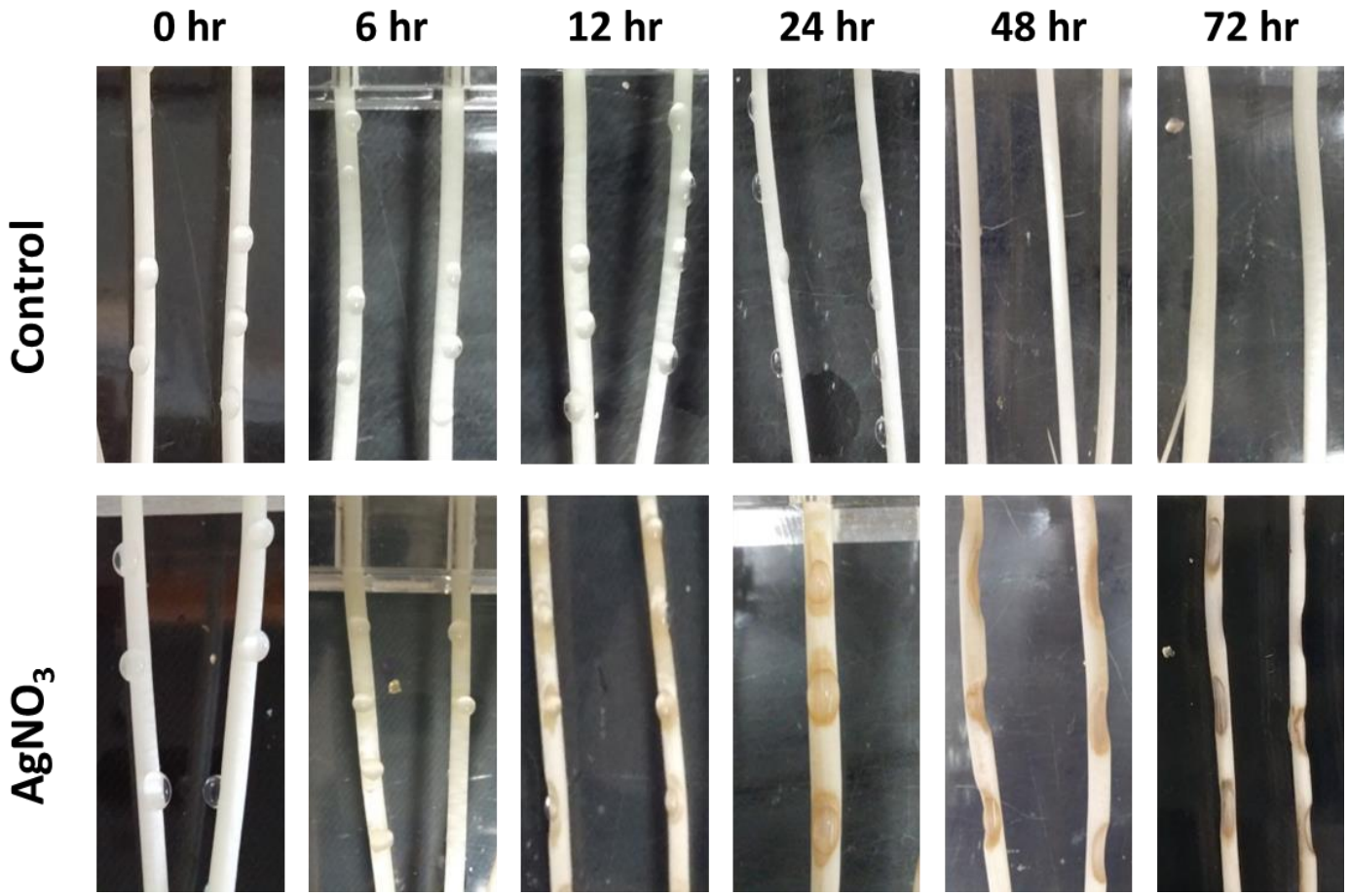


Figure 3.4 Soybean hypocotyl physiology during $AgNO_3$ treatment. Hypocotyls were treated with either 1 mM $AgNO_3$ or water (control). Pictures were taken at the indicated times of the time course. Brown lesions that worsened over the time course on the $AgNO_3$-treated hypocotyls are absent in the control.

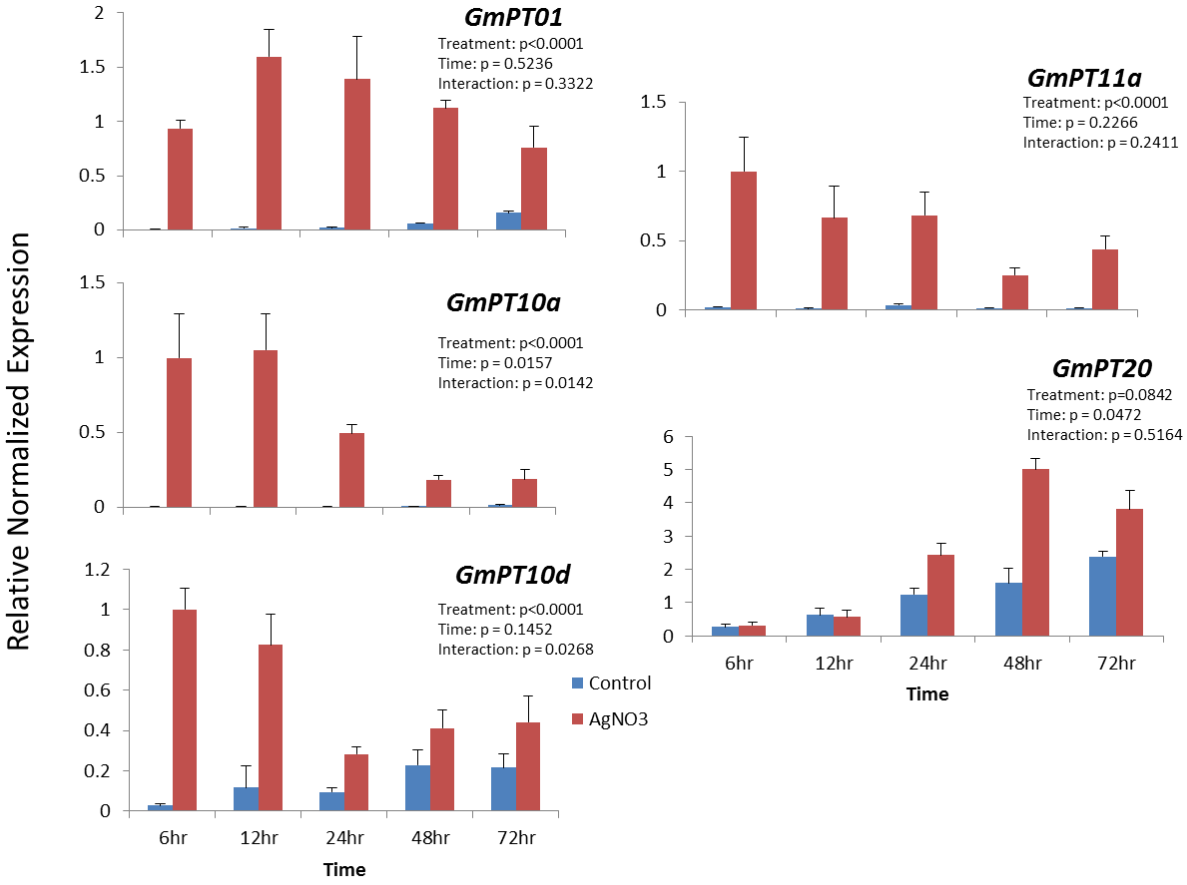


Figure 3.5 Gene expression of *GmPTs* in response to $AgNO_3$ treatment. qPCR was performed using total RNA extracted from $AgNO_3$-treated (red) and water-treated (control) hypocotyls (blue). Error bars denote standard error of the mean (SEM) of two biological and three technical replicates. Values were normalized against the reference gene, *CONS4*. Two-way ANOVA results are noted for each gene study with $p<0.05$ denoting significance.

3.3 Candidate *GmPTs* show tissue-specific expression

As seen in the previous experiment, certain induced *GmPTs* were also expressed under normal conditions (Figure 3.5). To determine the spatiotemporal gene expression of the candidate *GmPTs* during soybean development, RNA was extracted from twelve soybean tissues and qPCR was performed.

As shown in Figure 3.6, *GmPT01* was expressed predominantly in the root, leaf and flower. *GmPT10a* was highly expressed in the roots and mature embryo (60 and 70 DAP) before they start approaching desiccation. In addition, it was weakly expressed in the leaf and flower (Figure 3.6). As well, *GmPT10d* displayed the largest distribution of expression. It was expressed in all the tissues under the study except in stems, but showed highest expression in the flower bud, seed coat and late embryo stages. *GmPT11a* had the most unique expression profile with relatively weak expression in comparison to other *GmPTs*; concentrated in the root and flower. *GmPT20* was mostly expressed in the roots and leaf.

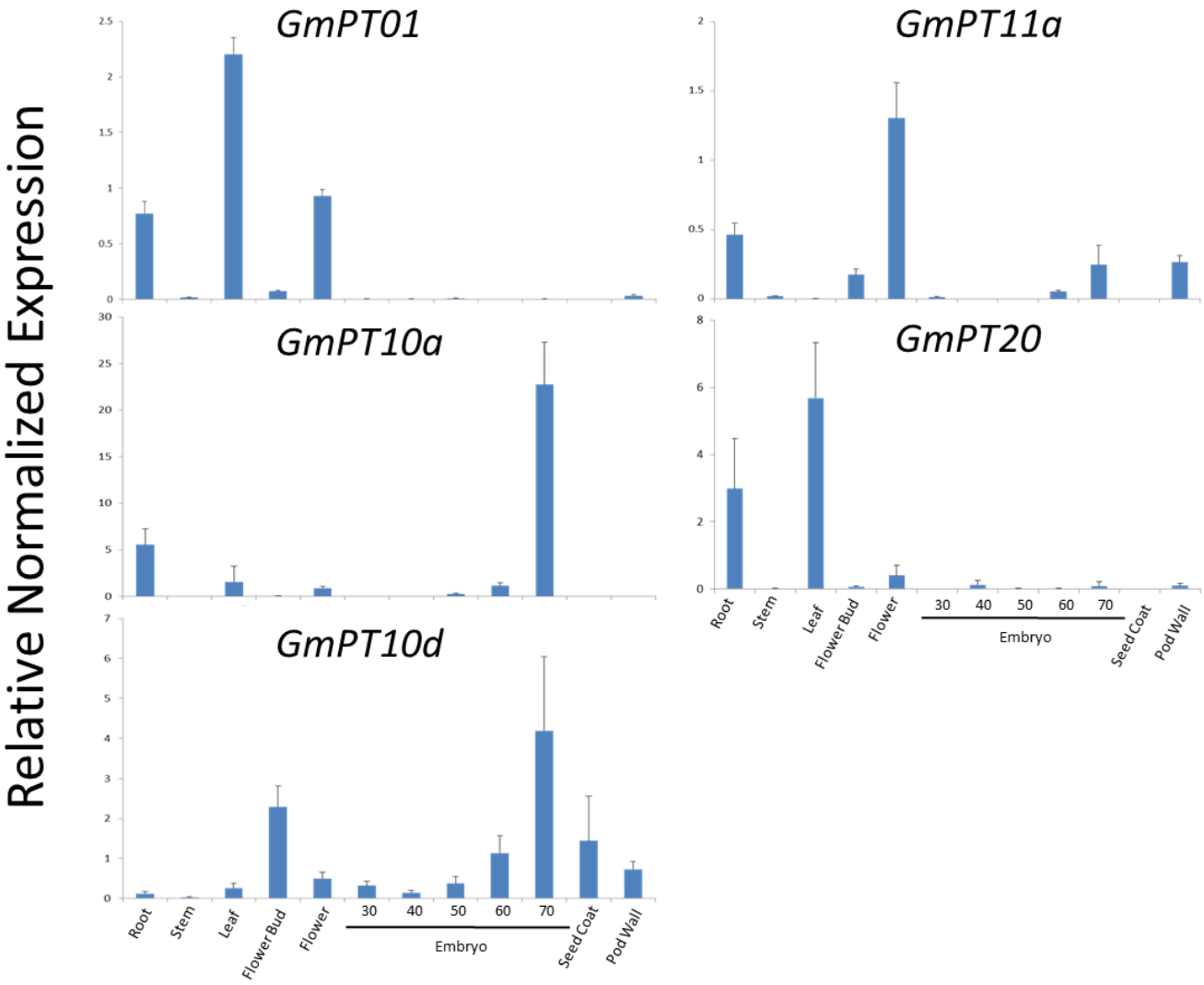


Figure 3.6 Tissue-specific expression of candidate *GmPTs* during normal growth and development. qPCR was performed using total RNA extracted from twelve soybean tissues: root, stem, leaf, flower bud, flower, embryo (30, 40, 50, 60, and 70 DAP), seed coat and pod wall. Error bars denote standard error of the mean (SEM) of two biological and three technical replicates. Values were normalized against the reference gene, *CONS4*.

3.4 Promoter analysis

Cis-activing elements are segments of non-coding DNA that regulate gene transcription. To determine whether the induced GmPTs contain any promoter elements that would explain its expression pattern, *cis*-acting elements within the promoter region of the induced candidate *GmPT* were identified by the MatInspector program. This program identified approximately 150 regulatory motifs per gene promoter. The results were screened to identify motifs of interest, from which, 11 different motif categories were selected (Table 3.5). CCAAT-box, TATA box, TATA box binding proteins (TBP) and transcription regulation elements are involved in transcription of these genes. Ethylene, jasmonate, and salicylic acid have all been shown to be involved in plant response to biotic and abiotic stress. Pathogen response elements (e.g. TEIL (tobacco EIN3-like)) and WRKY transcription factors play a role in plant defense. MYB transcription factors have already been shown to regulate isoflavonoid biosynthesis. Lastly, nodulin genes are involved in the symbiotic relationship between legumes and nitrogen fixing bacteria, and are one of the important functions of isoflavonoids in plants. These eleven motifs were found in most of the induced *GmPTs*, which may point towards a potential role of these genes in plant defense against pathogens.

Table 3.3 Promoter elements found upstream of induced *GmPTs*

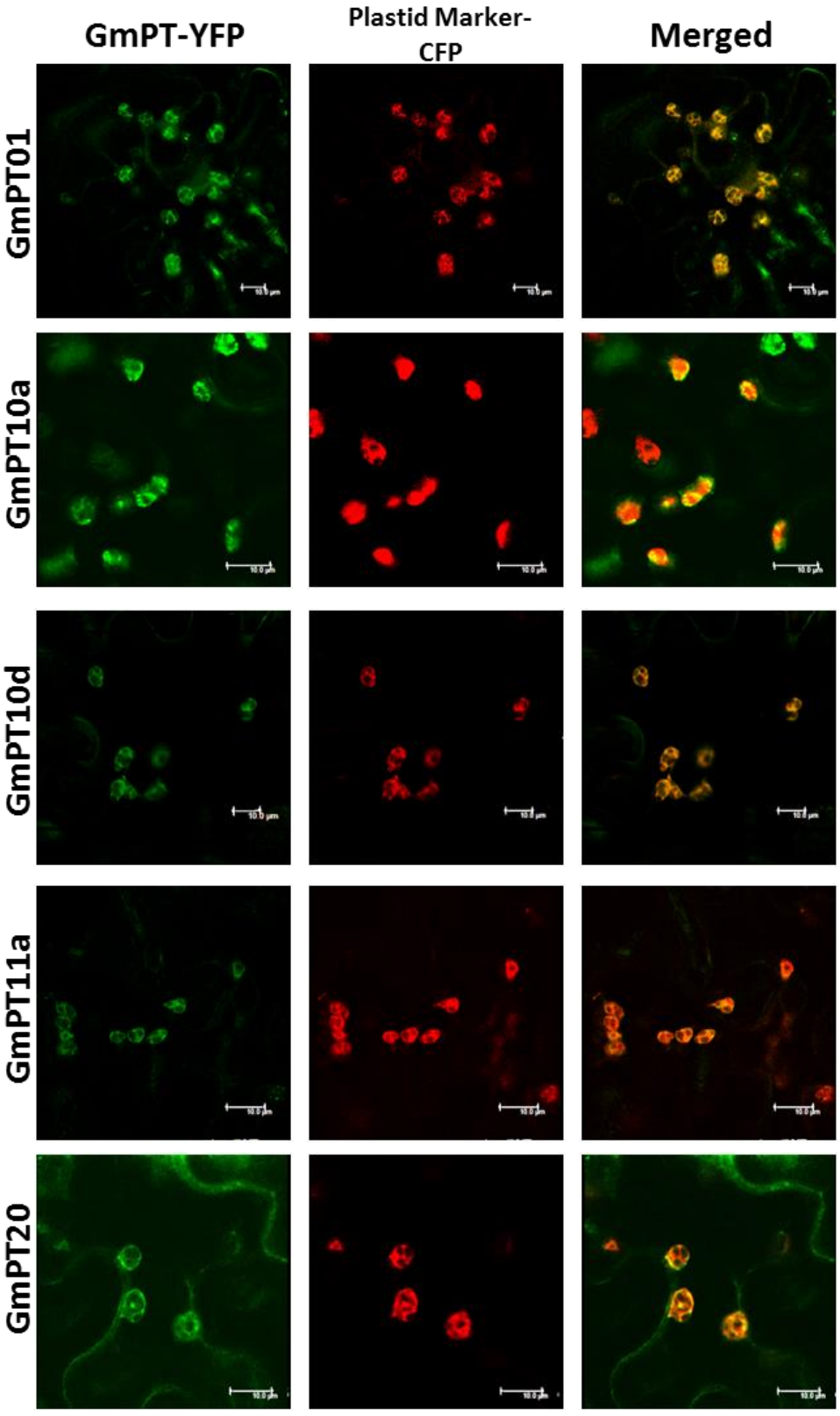
Promoter Elements	*GmPT01*	*GmPT10a*	*GmPT10d*	*GmPT11a*	*GmPT20*
CCAAT-box	4	0	1	1	0
Ethylene responsive	6	4	4	2	5
Jasmonate related	1	2	4	1	1
MYB-related	14	15	14	13	27
Pathogen response	2	1	2	4	0
SA related	1	1	0	1	1
TATA box	22	9	20	15	25
TATA binding proteins (TBP)	2	9	6	3	17
Transcription regulation	1	1	2	0	0
WRKY domain	1	5	1	4	8
Nodulin	0	0	0	3	4
Total:	54	47	54	47	88

3.5 Candidate GmPTs localize to plastids

Determining the subcellular localization of the candidate GmPTs would further knowledge of glyceollin biosynthesis and the isoflavonoid metabolon. Previously, GmG4DT was shown to localize to plastids (Akashi et al., 2009). Using the software WoLF PSORT, the candidate GmPTs are predicted to possess a plastid transit peptide. Verification of this prediction was performed by production of a translational fusion of the full length candidate GmPTs to YFP. These constructs were transiently expressed in the leaf epidermal cells of *N. benthamiana* and visualized by confocal microscopy. Confirmation of the localization was performed by co-expressing the GmPT-YFP construct with a plastid localization signal fused to CFP. All five induced GmPTs show plastid localization (Figure 3.7), with GmPT10a replicating a previous finding (Akashi et al., 2009). GmPT20, uniquely, was localized to both the cytosol and the chloroplast.

Referring to expression analysis data on Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html), *GmPT03* is expressed at extremely low levels in roots and the shoot apical meristem. This gene was unable to be amplified from various soybean tissues and under stress conditions, therefore, this gene was removed from the list of candidate genes.

Figure 3.7 Subcellular localization of the candidate GmPTs. A translational fusion of GmPT-YFP was transiently expressed in the leaf epidermal cells of *N. benthamiana* and visualized with confocal microscopy. Confirmation of localization was performed through co-localization of GmPT-YFP fusion with a plastid marker fused to CFP. Merged signal was obtained by sequential scanning of the two channels. Scale bar represents 10 µM.


GmPT-YFP
Plastid Marker-CFP
Merged
GmPT01
GmPT10a
GmPT10d
GmPT11a
GmPT20

3.6 Functional analysis of GmPTs

To perform an *in vitro* enzyme assay to test the enzymatic function of candidate GmPTs, the coding sequence, lacking the region encoding the signal peptide, were cloned in a yeast expression vector and yeast microsomal fractions containing GmPTs were prepared. Enzymatic reactions using yeast microsomal fractions containing a full-length prenyltransferase were unable to prenylate glycinol (Akashi et al., 2009).

3.6.1 Production and isolation of glycinol

As described in Section 1.5, glycinol is the substrate for the candidate GmPTs. However, one difficulty with this work was that pure glycinol is not commercially available, and thus it was incumbent that it be purified from soybean. The isolation of glycinol was heavily reliant on literature (Boue et al., 2009; Simons et al., 2011; Aisyah et al., 2013). Consistent throughout literature was that under similar extraction protocols, solvent, and gradient conditions, glycinol eluted off the HPLC column in between the glycosides, daidzin and genistin. Therefore, standards of these two compounds were used as markers for the area in which glycinol would potentially elute.

Glyceollin production was induced in $AgNO_3$-treated seed samples but was absent in water-treated seed samples (control) (Figure 3.8). Glyceollins are the last peaks eluted in the $AgNO_3$-treated chromatograms and are only visible at the 280 nm signal. The chromatograms of these samples detected at 254 nm can be found at Appendix C. Furthermore, glyceollin accumulation peaked at 48 hours and then declined by 72 hours. Therefore, the 48 hour time point was selected as the optimal incubation time for seed treatment.

Lovastatin and fosmidomycin have been reported to prevent prenylation of glycinol, and therefore, was used to accumulate glycinol (Akashi et al., 2009). To test the efficacy of these chemical inhibitors, seeds were treated with $AgNO_3$ and inhibitors at different concentrations. Upon comparison of the various treatments, the chemicals appear to have been successful in inhibiting glyceollin production (Figure 3.9). Furthermore, there is no increased efficacy when the concentration of inhibitors was increased. Therefore, 100 μM is sufficient to prevent glyceollin accumulation.

Referring back to literature, glycinol is expected to elute between daidzin and genistin (Boue et al., 2009), but the absence of a peak corresponding to glycinol within any treatment suggested the possibility of peak co-elution. To test this theory, samples were run under a new gradient system. From visual inspection of the chromatograms, new peaks were present, therefore, indicating that co-elution was previously occurring (Figure 3.10). One of the new peaks is present in the treated conditions and not in the control. Furthermore, this peak is only visualized at 280 nm, which may indicate that this peak is glycinol.

Following this, the $AgNO_3$-treated samples were screened for a compound with glycinol's mass-to-charge (273.07) and eluted between the two standards, daidzin and genistin (Figure 3.11). Under these criteria, a compound was putatively identified as glycinol.

To confirm the identity of this compound, the putative glycinol peak (boxed in red) was isolated using semi-preparative HPLC (Figure 3.12a). Following isolation, the sample was analyzed with LC-MS/MS with increasing collision energy (Figure 3.12b). Based on manual interpretation of the MS^2 results, the compound was putatively identified as glycinol. A NMR confirmation will have to be performed to verify the identity of this compound.

Figure 3.8 Metabolite profile of control and $AgNO_3$-treated soybean seeds. Soybeans seeds were imbibed, halved, and then treated with $AgNO_3$ or water (control). Treated seeds were incubated in the dark for 24, 48 or 72 hours in the dark at 25°C. Metabolites were extracted with methanol and samples analyzed with HPLC. Samples are detected at 280 nm with absorbance being measured in mAU. Daidzin (D) and genistin (G) standards were also run.

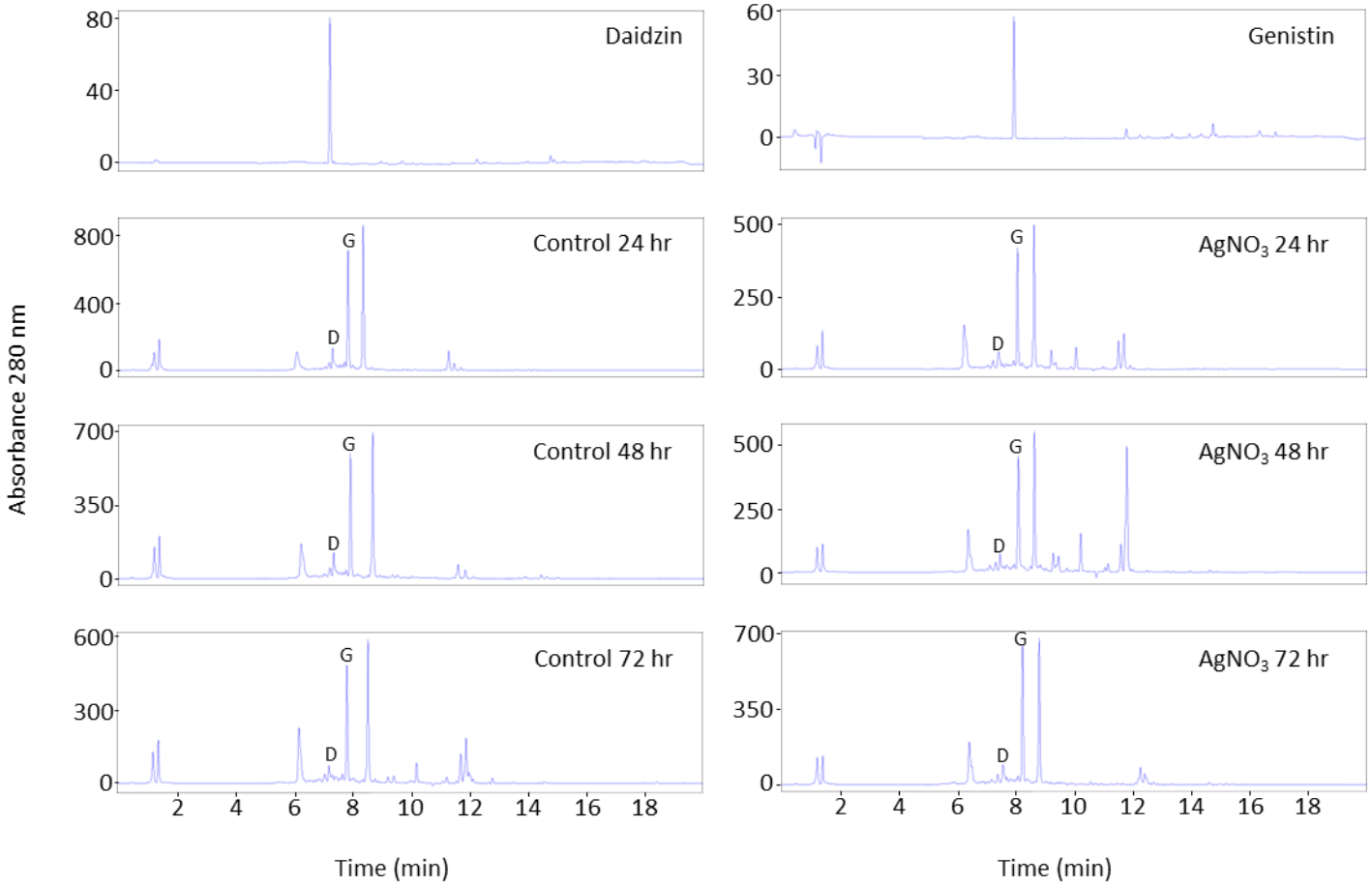

Daidzin
Genistin
Control 24 hr
AgNO3 24 hr
Control 48 hr
AgNO3 48 hr
Control 72 hr
AgNO3 72 hr
G
D
Absorbance 280 nm
Time (min)
Time (min)

Figure 3.9 Effect of inhibitors on metabolite profile. Soybean seeds were imbibed, halved and treated with $AgNO_3$ and varied concentrations of inhibitors. Treated seeds were incubated in the dark for 48 hours at 25°C. Seeds were extracted with methanol and samples analyzed with HPLC. Samples are detected at 280 nm with absorbance being measured in mAU. Daidzin (D) and genistin (G) standards were also run. Red rectangles indicate the peaks for glyceollins.

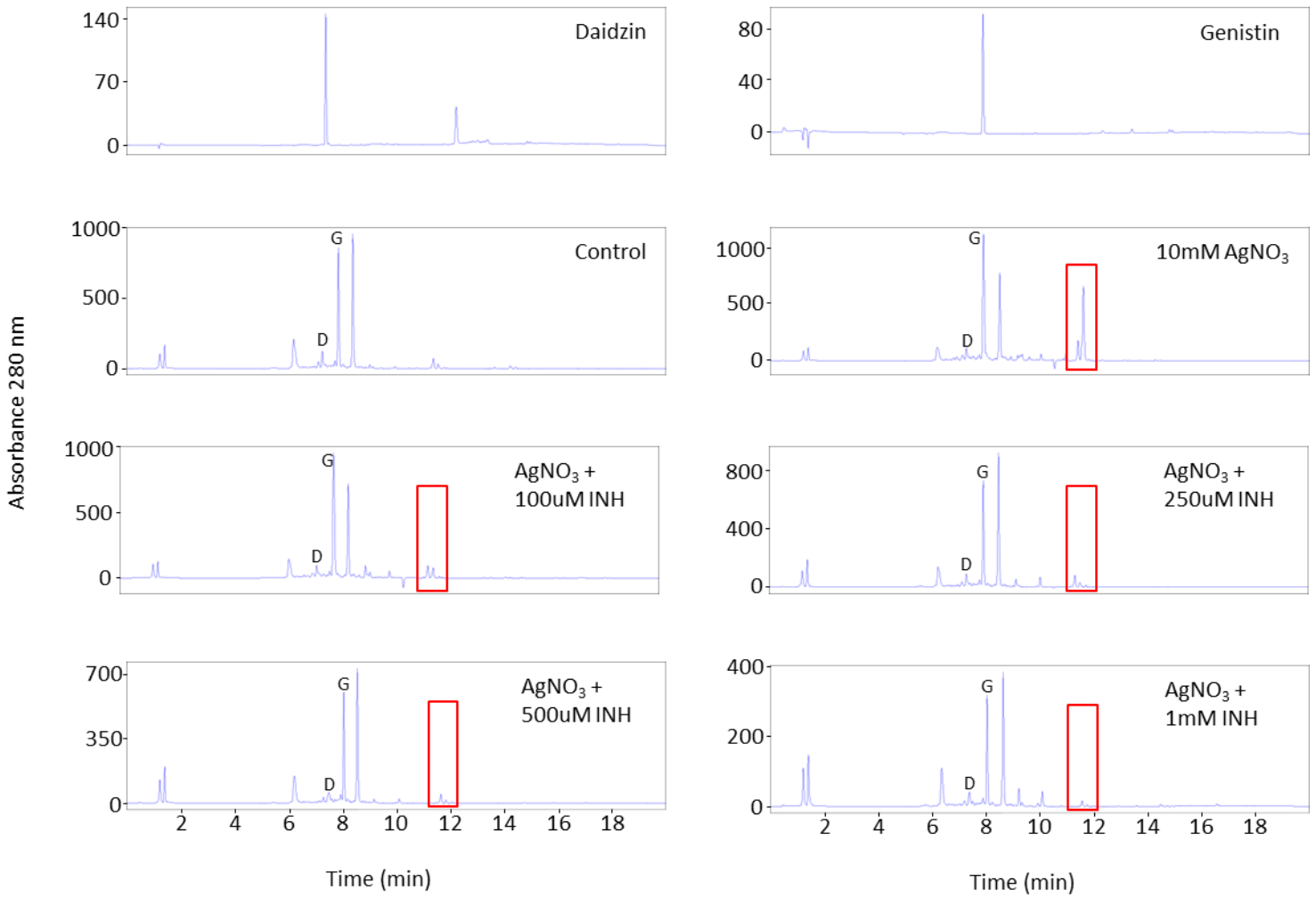

Absorbance 280 nm
Daidzin
Genistin
Control
10mM AgNO3
AgNO3 + 100uM INH
AgNO3 + 250uM INH
AgNO3 + 500uM INH
AgNO3 + 1mM INH
G
D
Time (min)

Figure 3.10 Altering the solvent gradient alleviates co-elution of metabolites. Samples were run under an altered solvent gradient to lengthen the portion in between daidzin and genistin. Samples are detected at 254 and 280 nm with absorbance being measured in mAU. Daidzin (D) and genistin (G) standards were also run. Red box encloses the peak predicted to be glycinol.

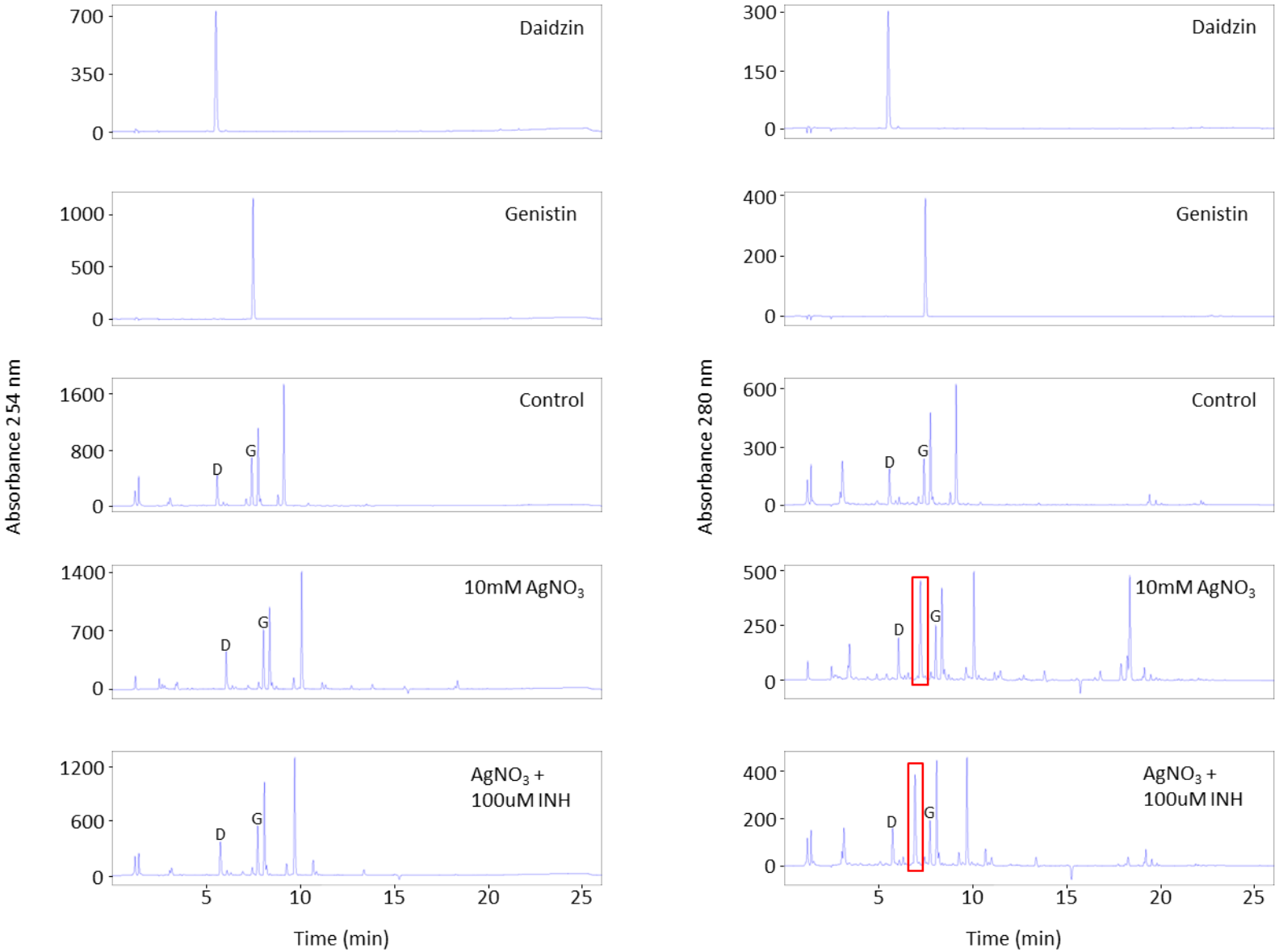

Absorbance 254 nm
Absorbance 280 nm
Daidzin
Genistin
Control
10mM AgNO3
AgNO3 + 100uM INH
D
G
Time (min)

Figure 3.11 LC-MS of $AgNO_3$-treated soybean seeds. $AgNO_3$-treated seeds and standards were analyzed with LC-MS. Top three plots depict the raw chromatograms of daidzin, genistin and $AgNO_3$-treated samples. The $AgNO_3$-treated chromatogram depicts metabolites with a mass-to-charge of 273. The bottom three plots depict the matching MS spectra.

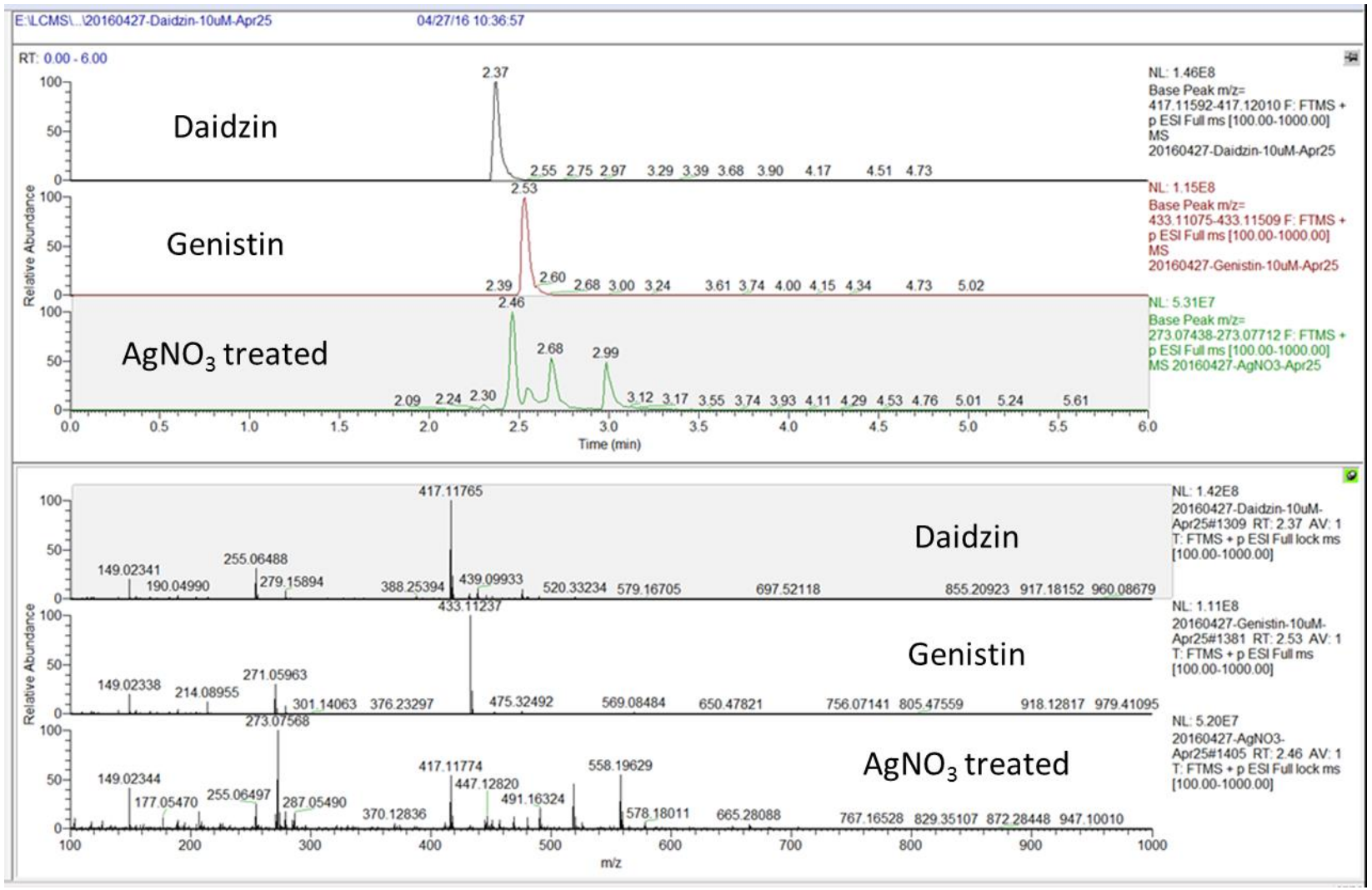
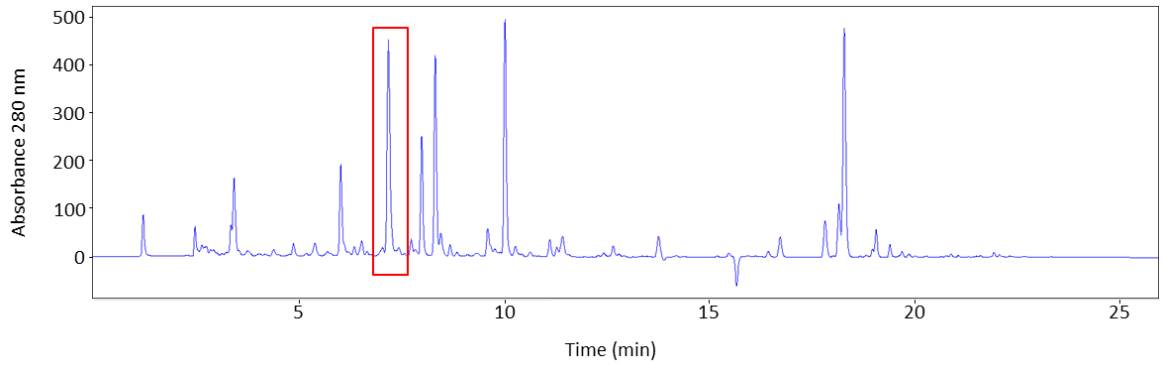
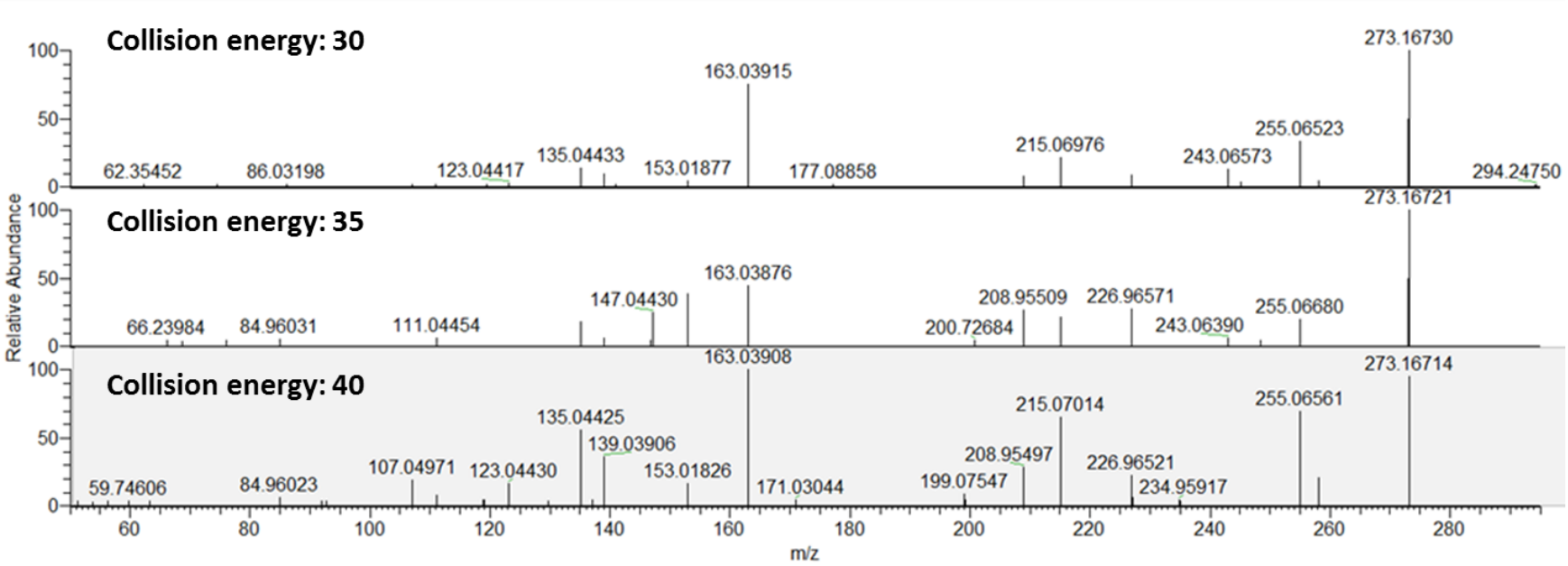

E:\LCMS\...\20160427-Daidzin-10uM-Apr25
04/27/16 10:36:57
RT: 0.00 - 6.00
Daidzin
Genistin
AgNO3 treated
Relative Abundance
Time (min)
NL: 1.46E8
Base Peak m/z= 417.11592-417.12010 F: FTMS + p ESI Full ms [100.00-1000.00] MS 20160427-Daidzin-10uM-Apr25
NL: 1.15E8
Base Peak m/z= 433.11075-433.11509 F: FTMS + p ESI Full ms [100.00-1000.00] MS 20160427-Genistin-10uM-Apr25
NL: 5.31E7
Base Peak m/z= 273.07438-273.07712 F: FTMS + p ESI Full ms [100.00-1000.00] MS 20160427-AgNO3-Apr25
NL: 1.42E8
20160427-Daidzin-10uM-Apr25#1309 RT: 2.37 AV: 1 T: FTMS + p ESI Full lock ms [100.00-1000.00]
NL: 1.11E8
20160427-Genistin-10uM-Apr25#1381 RT: 2.53 AV: 1 T: FTMS + p ESI Full ms [100.00-1000.00]
NL: 5.20E7
20160427-AgNO3-Apr25#1405 RT: 2.46 AV: 1 T: FTMS + p ESI Full lock ms [100.00-1000.00]
m/z

Figure 3.12 Purification and Confirmation of Glycinol. A) Chromatogram of $AgNO_3$-treated samples at 280 nm. Red box encloses the peak predicted to be glycinol. B) LC-MS/MS chromatograms for putative glycinol compound with increasing collision energies.

A)


Absorbance 280 nm
500
400
300
200
100
0
5
10
15
20
25
Time (min)

B)


Collision energy: 30
273.16730
163.03915
255.06523
215.06976
243.06573
62.35452
86.03198
123.04417
135.04433
153.01877
177.08858
294.24750
Collision energy: 35
273.16721
163.03876
147.04430
208.95509
226.96571
255.06680
66.23984
84.96031
111.04454
200.72684
243.06390
Collision energy: 40
163.03908
273.16714
135.04425
215.07014
255.06561
139.03906
208.95497
107.04971
123.04430
153.01826
226.96521
59.74606
84.96023
171.03044
199.07547
234.95917
Relative Abundance
m/z
60
80
100
120
140
160
180
200
220
240
260
280

3.6.2 Enzyme assay of GmPT10a

In vitro enzyme assay was performed to test whether induced GmPTs could prenylate glycinol in the presence of DMAPP and $MgCl_2$. GmPT10a (G4DT) has been shown to have C-4 prenyltransferase activity (Akashi et al., 2009), thus was used as a positive control. Microsomal fractions containing truncated GmPT10a were incubated with purified glycinol, DMAPP and $MgCl_2$. Following incubation, the ethyl acetate extract was analyzed by HPLC. Analysis of the chromatogram for the reaction extract did not show a new peak corresponding to the absence of a product compound (Figure 3.13). Furthermore, the retention time for the peak in the glycinol and the reaction mixture chromatogram are the same, indicating that glycinol was not prenylated. The reaction is expected to work and therefore, the protocol will require further optimization.

3.7 QTL analysis

Detailed search of the Soybase database (Grant et al., 2010) and literature identified 72 QTLs associated with *P. sojae* resistance in soybean (Appendix A). Of the induced *GmPTs*, only *GmPT01* was found within a QTL for resistance to *P. sojae* infection. The chromosome number, parental lines, LOD scores and genomic location are noted (Table 3.3). Phytoph 13-3 has an LOD score of 3.7, which would associate it with a high probability of genetic linkage or phenotypic association.

A literature search for QTL markers associated with *P. sojae* resistance identified a total of 109 markers (Appendix B). Two QTL markers (Sat_414 and Satt439) were found to flank *GmPT01* and three QTL markers (Satt094, Satt0345 and Satt420) were found to flank *GmPT10d*. The gene loci, marker loci and the type of marker are described in Table 3.4.

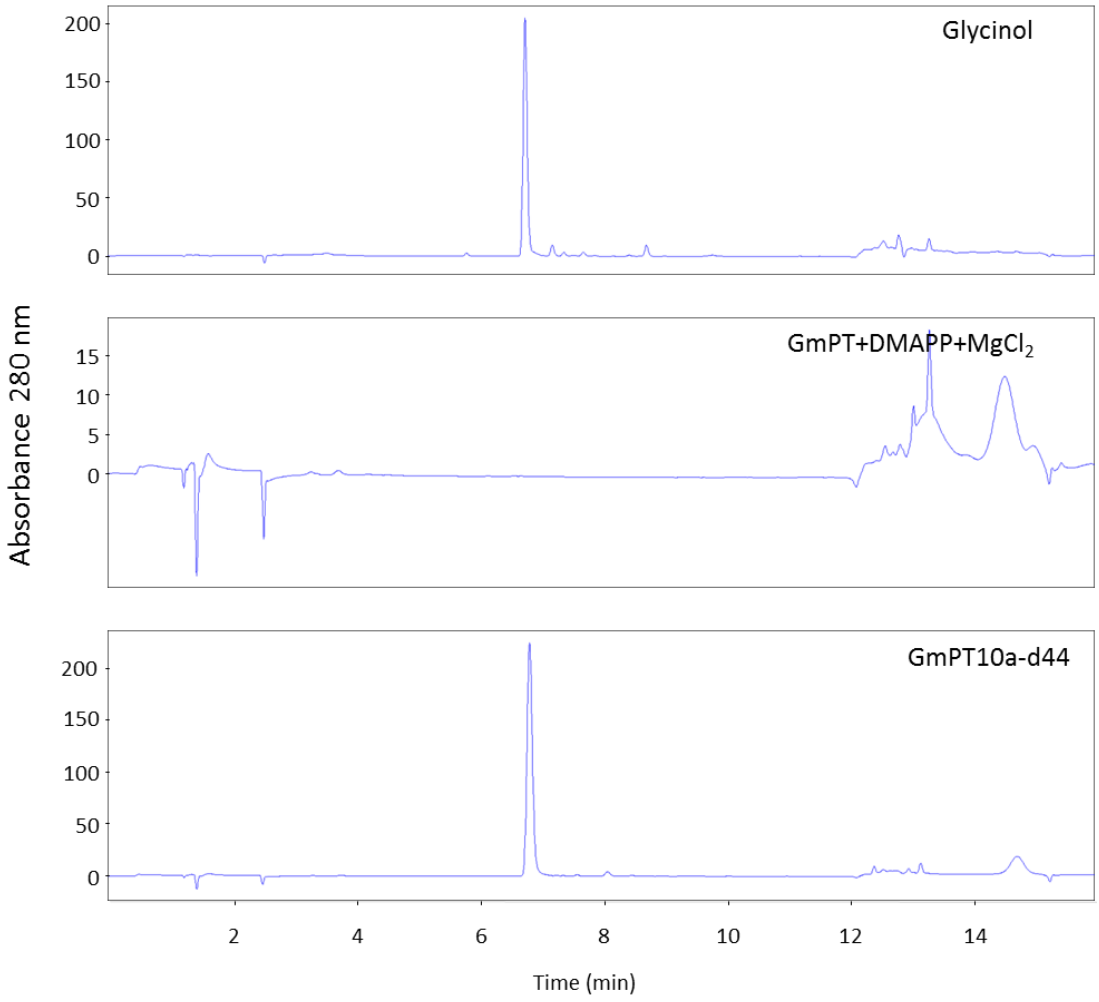


Figure 3.13 Enzyme assay of GmPT10a. Microsomal fraction containing truncated GmPT10a in KPB was incubated with glycinol, DMAPP and $MgCl_2$. Ethyl acetate extract was analyzed by HPLC.

Table 3.4 QTL marker associated with *P. sojae* resistance contains *GmPT01*

QTL	Chromosome	QTL Marker Interval (physical map)	*GmPT* location	Parents	Heritability	LOD	Reference
Phytoph 13-3	01	42,423,585 - 46,773,166	*GmPT01* 45619439 - 45625938 reverse	**Parent 1:** Conrad **Parent 2:** Sloan	0.87	3.7	Wang et al., 2012

Table 3.5 QTL markers that flanks induced *GmPTs*

GmPT	QTL Marker	QTL Marker location	Type of Marker
GmPT01 45619439 - 45625938 reverse	Sat_414	51,449,361-51,449,420	SSR
	Satt439	48,828,582-48,828,608	SSR
GmPT10d 7023173 – 7029710 forward	Satt094	12,407,171 – 12,407,229	SSR
	Satt345	12,172,070-12,172,150	SSR
	Satt420	9,907,940 – 9,907,987	SSR

Chapter 4: Discussion

An essential part of the plant defense response, isoflavonoid biosynthesis culminates in glyceollin production. Prenyltransferase activity plays an important role in the second last step of glyceollin biosynthesis. After the hydroxylation of daidzein, the pathway follows a linear progression until the prenylation of glycinol, which branches the pathway, leading to the production of the different glyceollin isomers. The prenyltransferase reaction was sparsely studied with the two prominent studies being published in 1979 and then in 2009 (Zähringer et al., 1979; Akashi et al., 2009). Current knowledge of the pathway has one of the two proposed prenyltransferases being identified and characterized (G4DT), and the second prenyltransferase simply being a proposed enzyme (G2DT). The soybean genome contains numerous multi-gene families that resulted from two whole-genome duplications (Schmutz et al., 2010). With research all but stalled on this critical enzyme to soybean innate defense, I identified the gene family of soybean prenyltransferases that are induced in response to pathogen infection. The isoflavonoid-specific *GmPT* family members display differential spatial expression patterns under normal conditions, and differential temporal expression in response to pathogen infection. The GmPTs localize to plastids and possess promoter motifs associated with response to biotic and abiotic stress. Furthermore, one *GmPT* is localized within a QTL linked to *P. sojae* resistance. The present research furthers knowledge of isoflavonoid biosynthesis and its metabolon, and further proposes a mechanism to connect prenyltransferase expression to stress response.

4.1 Soybean (iso)flavonoid prenyltransferase family contains 11 members

Extensive data mining of the annotated soybean genome (*G. max* Wm82.a2.v1) identified 63 putative *GmPTs*. This number represents the current number of annotated prenyltransferases; however, this value is subject to change with additional deposits to the database. Of these 63, 11 *GmPTs* clustered with previously characterized (iso)flavonoid prenyltransferases in the phylogenetic tree (Figure 3.1). One of these 11 was G4DT, a previously characterized GmPT proven to be involved in glyceollin

biosynthesis (Akashi et al., 2009). The other ten genes have yet to be characterized and offer the possibly of putative prenyltransferase activity related to glyceolin biosynthesis. The pathway proposes the existence of two genes, one C-4 prenyltransferase (the aforementioned G4DT) and another C-2 prenyltransferase (yet to be identified). Therefore, there should be at least two isoflavonoid-specific *GmPTs*, however, the number of genes found in my research contrasts with this value. This finding is quite common in plant metabolic pathways, especially in soybean. There are other isoflavonoid biosynthetic genes that exist as multi-gene families such as *CHS* (Tuteja and Vodkin, 2008), *CHR* (Sepiol, 2015), *CHI* (Dastmalchi and Dhaubhadel, 2015) and *IFS* (Jung et al., 2000).

The large gene families are predicted to be the result of tandem duplications and whole genome duplication of the soybean genome. Following tandem duplication, stress responsive genes have a higher probability of retention (Hanada et al., 2008). The benefit exists in the event of a null mutation, the phenotypic effects will be smaller with multigene families in comparison to those controlled by single genes (Hanada et al., 2009).

A paper recently published following the start of this study identified ten *GmPTs* that were mentioned in this study (Wang et al., 2015). This study examined the evolution of genes in the UbiA prenyltransferase gene family found in various plant species. They identified 24 soybean UbiA prenyltransferases, of which, 12 were classified as flavonoid and homogentisate prenyltransferases. The study identified numerous paralogous and orthologous gene pairs in soybean. The genes labelled as flavonoid prenyltransferases were found to exhibit high nucleotide substitution rates and frequent sequence exchange events, which would result in the accumulation of genetic variation. This high rate of genetic variation is predicted to be driven by positive selection. The proposed model suggests that after legumes split from dicots, the common ancestral gene produced two copies through transposition (Lavin 2005). When *G. max* and *Phaseolus vulgarius* split, tandem duplication of one gene resulted in three copies (four total). Two of the three recently duplicated genes continued to undergo tandem duplication and duplicative transposition, resulting in five copies (six total). The third gene was lost. Following the

second whole genome duplication, all but one gene underwent duplication (eleven total). The last step was the tandem duplication of one gene to produce two copies, which left the genome with twelve copies of flavonoid or homogentisate prenyltransferases.

Here, I identified 11 *GmPTs*, with *GmPT08* being previously undiscovered. Nine genes were characterized based on their expected enzymatic activity due to the number of transmembrane domains and possession of critical residues. Though the eliminated *GmPTs* are structurally different from other UbiA prenyltransferase, this does not rule out the possibility of having enzymatic activity. They may possess weak prenyltransferase activity or have evolved to perform a new function.

4.2 Five *GmPTs* are induced upon infection

Many isoflavonoid biosynthesis genes are upregulated in response to pathogen attack (Moy et al., 2004; Vega-Sánchez et al., 2005; Akashi et al., 2009). The candidate *GmPTs* are proposed to be involved in glyceollin biosynthesis, therefore, they too are expected to be upregulated in response to stress. Semi-quantitative RT-PCR identified 5 (*GmPT01, GmPT10a, GmPT10d, GmPT11a* and *GmPT20*) out of the 9 candidate genes that showed upregulation of gene expression in infected samples in comparison to control (Figure 3.5).

Accumulation of glyceollin in elicited hypocotyls was decreased by 50% in the presence of an ethylene biosynthesis inhibitor, aminoethoxyvinylglycine (AVG) (Kimpel and Kosuge, 1985). Furthermore, when cotyledons were removed prior to elicitation, the presence or absence of AVG did not impact glyceollin accumulation. Therefore, ethylene was proposed not to initiate glyceollin production, but rather may function as a signal to promote metabolite flux into the pathway. In the absence of ethylene, ethylene receptor, ETR1 is involved with negative regulation of ethylene signaling. $AgNO_3$ has been proposed to interfere with ETR1 activity, thereby promoting ethylene signalling (Rodríguez et al., 1999). Through $AgNO_3$ treatment, this mechanism may function to indirectly promote glyceollin production.

Further treatment with $AgNO_3$, which is capable of eliciting glyceollin production, was used to mimic pathogen attack. Quantitative analysis was performed with these samples to view the expression profile across 72 hours. Analysis of this expression profile supported the use of $AgNO_3$ as a mimic to pathogen attack and the upregulation of the same five *GmPTs*, but showed that these genes differ temporally. *GmPT10a* (G4DT), *GmPT10d* and *GmPT11a* are all induced relatively quickly in response to treatment. This expression is very similar to glyceollin I (the result of prenylation by G4DT), which was found to rapidly accumulate and then decline following stress. *GmPT01* and *GmPT20,* however, showed gradual increase in transcript level throughout the time course. This is similar to glyceollin II and III (the result of prenylation by G2DT) that accumulate slowly following stress (Bhattacharyya and Ward, 1986).

4.3 *GmPTs* are spatially distributed under normal conditions

Prior to glyceollin accumulation at the site of infection, precursor molecules accumulate at the site (Hahn et al., 1985; Graham et al., 1990). Basal-level expression of the candidate *GmPTs* would offer to further understanding of the plant defense response. Quantitative analysis of the transcript expression of the candidate *GmPTs* was measured in twelve soybean tissues and the results indicate that the *GmPT* differ spatially (Figure 3.6). The differential expression across all the soybean tissues, may offer to elaborate which genes are expressed upon infection. Basal level expression of all *GmPTs* was found in most soybean tissues, demonstrating that these induced enzymes are expressed under normal conditions. A potential mechanism may exist where biosynthetic enzymes are present under normal conditions, which could trigger rapid phytoalexin production in the event of pathogen infection.

GmPT10a and *GmPT20* were expressed at higher level in roots relative to other *GmPTs,* suggesting a more prominent role of these two genes during *P. sojae* infection. Under such a model, *GmPTs* would be expressed at basal level in different soybean tissues and respond according to wherever stress is occurring. All *GmPTs* are expressed in flowers, which may reflect a synchronous response to stress of this tissue. *GmPT20* would be expressed the most in leaves in response to infection, whereas, *GmPT01* and *GmPT10a*

would be weakly expressed. *GmPT10a* and *GmPT10d* would be expressed in the seeds during embryo development, and *GmPT10d* would be expressed in the soybean pods.

This spatial distribution of a gene family was similarly reported in a recent study characterizing the cation/proton antiporter 1 (CPA1) gene family in grapevine (*Vitis vinifera*) in response to osmotic and salt stress (Ma 2015). Subfunctionalization of two members in this family resulted in these genes functioning in different tissues during different stages of stress. From analysis of tissue-specific expression, *CHS* gene family members demonstrate potential subfunctionalization by their unique expression profiles (Tuteja et al., 2004). Following this model, an ancestral isoflavonoid prenyltransferase may have evolved to form the present *GmPT* gene family, during which, subfunctionalization of this gene produced *GmPTs* with altered tissue expression. Distribution of the expression of the *GmPTs* across the various tissues would eliminate the necessity to express all *GmPTs* in all tissues.

4.4 All GmPTs localize to plastids

The isoflavonoid metabolon proposes a precise and efficient mechanism to rapidly produce isoflavonoids and eventually, glyceollins (Dastmalchi et al., 2016). With isoflavonoids playing an important role in plant defense, the ability to quickly produce glyceollins would be essential to preventing and eliminating the spread of infection. Important findings by Dastmalchi et al. have established the core of an isoflavone metabolon, anchored to the ER by two cytochrome P450 monooxygenases (Dastmalchi et al., 2016). Through transient expression of GmPTs in *N. benthamiana*, the GmPTs appear to localize to plastids (Figure 3.7). This is in line with expectations of other flavonoid UbiA prenyltransferases, which also show plastidic localization.

In plants, there are two biosynthetic pathways for isoprenoids. The classical pathway is the melvonate (MVA) pathway that operates in the cytosol. The alternative is the methylerythritol phosphate pathway (MEP) localized to plastids (Lichtenthaler et al., 1997; Lohr et al., 2012). To determine the origin of the prenyl moiety that is transferred by prenyltransferase, a tracer experiment was performed. The study identified dimethylallyl pyrophosphate (DMAPP) from the MEP pathway as the source of the

prenyl group (Akashi et al., 2009). Similarly in peanuts (*Arachis hypogaea)*, a stillbenoid prenyltransferase, resveratrol 4-dimethylallyltransferase transferred a prenyl moiety from DMAPP to resveratrol (Yang et al., 2016). GmPTs would operate similarly, by recruiting DMAPP molecules from inside plastids to prenylate glycinol.

Unlike the other candidate GmPTs, GmPT20 was localized to plastids and within the cytosol. This localization was not previously identified by other (iso)flavonoid prenyltransferases, however, examples exists in literature within other pathways. ATP sulfurylase 2 (ATPS2) involved with sulfur assimilation in *A. thaliana* was found to encode plastidic and cytosolic localizing isoforms. The differential localization resulted from translational initiation at different start codons within its transit peptide. Initiation at AUG^{Met1} produces a plastid-localizing enzymes, whereas, initiation at AUG^{Met52} or AUG^{Met58} produces a cytosol-localizing enzyme (Bohrer et al., 2015). An alternate model could arise from a study looking at tyrosine metabolism in soybean, which found evidence for the existence of parallel cytosolic and plastid pathways (Schenck et al., 2015).

4.5 Evolution of isoflavonoid-specific GmPTs

Genes involved with plant secondary metabolites have long been suggested to have been derived from genes involved with primary metabolism (Panchy et al., 2016). The neighbouring clade to the GmPTs and the previously characterized (iso)flavonoid prenyltransferase function to produce tocopherols and tocotrienols (vitamin E) (Figure 3.1). Vitamin E is a primary metabolite, and the HGA prenyltransferases involved in its biosynthesis also show plastidic localization (Tian et al., 2007). This offers the theory that neofunctionalization of HGA prenyltransferases may have led to the creation of isoflavonoid prenyltransferases (Wang et al., 2015). Stress is proposed to place selective pressure on duplicate genes to gain novel functions (Zou et al., 2009). In the presence of stress, gene duplication would have allowed for adaptation to the surrounding biotic and abiotic environments (Panchy et al., 2016). Neofunctionalization of genes involved in primary metabolism may have resulted in the formation of legume-specific isoflavonoid biosynthetic genes. The isoflavonoid biosynthetic pathway may have been an adaptation to biotic and abiotic stress (Wang et al., 2015).

4.6 QTL and promoter analysis

The goal of this study was to determine whether these candidate *GmPTs* play a role in strengthening soybean's partial resistance through manipulation of isoflavonoid production, and therefore, breed cultivars resistant to *P. sojae* infection. Screening for QTL and QTL markers has been a large component of breeding for resistance. In comparison of QTL and QTL marker loci and *GmPT* loci, *GmPT01* is situated in a QTL associated with *P. sojae* resistance and also flanks two QTL markers (Table 3.3 and 3.4). *GmPT10d* was also flanked by QTL markers, thus making these two genes promising candidates for promoting partial resistance.

Analysis of the promoter region of candidate *GmPTs* attempted to identify any promoter elements that may elaborate the role of these genes as induced components of the plant defense response. The promoter region of the candidate *GmPTs* has numerous transcription factor binding sites associated with biotic and abiotic stress. In particular ethylene, jasmonate, and salicylic acid related elements were found in the promoter region. These compounds have all been shown to be involved with stress responses. Furthermore, WRKY transcription factors and WRKY binding sites (W-box) were found in the upstream region of these genes. These elements may function towards the induced expression of the genes in response to stress. Though the five induced *GmPTs* genes share many of the same promoter motifs, they vary in the distribution of each motif (Table 3.5). The differences in their transcriptional regulation could explain their unique spatiotemporal expression profiles. Paralogous MADS-domain transcription factors function differently as result of difference in their promoter and coding region (McCarthy et al., 2015). Differences in the promoter region of *CHS7* and *CHS8* were proposed to explain the differences in gene expression at different stages of plant development (Yi et al., 2010).

4.7 Update on the isoflavonoid metabolon

The existing model of the isoflavonoid metabolon proposes that soluble isoflavonoid enzymes are tethered to the ER by at least two P450 enzymes, C4H and IFS (Dastmalchi et al., 2016). The current model, however, does not contain the enzymes downstream of IFS that are involved in glyceollin biosynthesis. Based on the assumption that the downstream portion of the pathway also interacts with the metabolon, we obtain the model shown in Figure 4.1. To the current model we add the three P450s, I2'H, 3,9DPO and GS. The soluble enzymes, 2HDR and PTS are expected to be tethered by these ER proteins. Lastly, this study adds the G4DT/G2DT enzymes as integral membrane proteins localized to the plastid, which has their catalytic site facing the cytosol. Confirmation of interactions between the newly added members is first required before interactions can be shown.

The model proposes a shuttling of substrates across the enzymes of the metabolon, however, the localization of G4DT to plastids, would suggest a more complicated mechanism may be in place. Tocopherols are compounds that are synthesized in plastids, which impact ER fatty acid desaturation, suggesting a plastid-ER interaction (Mehrshahi et al., 2013). Mutants lacking the plastid-localized enzymes for tocopherol biosynthesis were complemented with wild type enzymes targeted to the ER in a procedure named "transorganellar complementation". The ER-localized enzymes were able to access plastid metabolite pools, suggesting interorganellar continuity between the plastid and ER (Mehrshahi et al., 2013). It was further suggested that contact points between both organelles termed plastid-associated membranes (PLAMs) (Andersson et al., 2007).

The isoflavonoid metabolon may be situated near a PLAM to mediate the shuttling of glycinol from the preceding enzyme, 3,9DPO to G4DT or G2DT. The prenyltransferase has its catalytic site on the cytosolic face of the plastid membrane. Following prenylation of glycinol, the metabolite would be shuttled to glyceollin synthase (sitting in the ER) to catalyze the last step of glyceollin biosynthesis.

Additionally, ADT6, an shikimate pathway enzyme that is upstream of the phenylpropanoid pathway was demonstrated to interact with IFS2 (Dastmalchi et al.,

2016). The shikimate pathway is responsible for the biosynthesis of various aromatic amino acids, including phenylalanine. The interaction of ADT6 and IFS2 was proposed as a mechanism for metabolite flux towards isoflavonoid biosynthesis. ADT6 is a plastidic enzyme, and may interact with IFS through a plastid-ER interaction mentioned above.

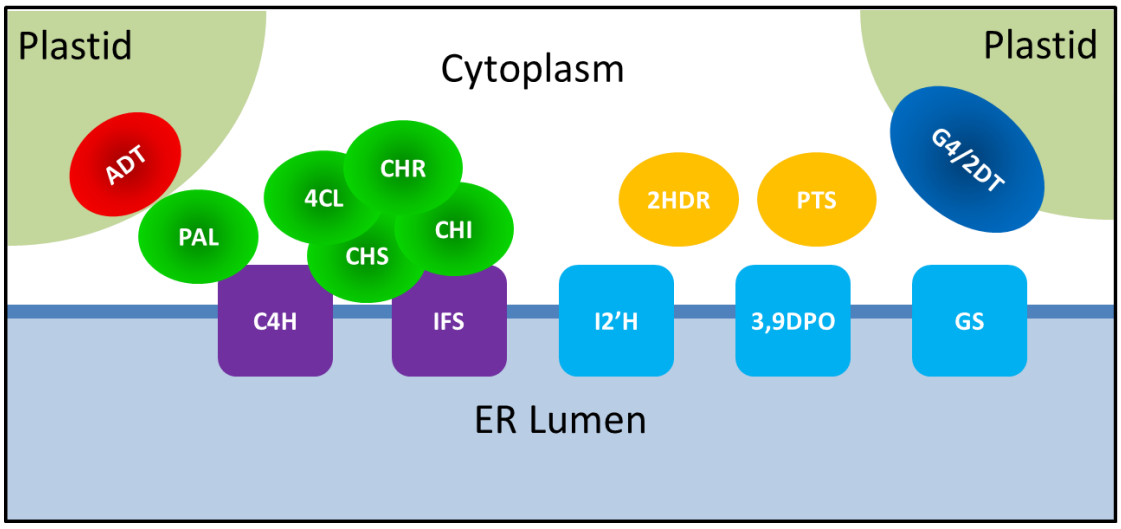


Figure 4.1 Generalized model of the isoflavonoid metabolon. The model proposed by Dastmalchi et al (2016), consists of the twin ER-anchors, C4H and IFS (dark purple) tethering the four soluble biosynthesis enzymes, 4CL, CHS, CHR and CHI (green). Enzymes upstream of isoflavonoid biosynthesis are highlighted in red. New additions to the model are the following: three new P450s sitting in the ER, I2'H, 3,9DPO and GS (light blue), two new soluble enzymes, 2HDR and PTS (yellow), and the prenyltransferases, G4DT or G2DT sitting in the plastid membrane (dark blue). ADT (arogenate dehydratase); PAL (phenylalanine ammonia lyase); C4H (cinnamate 4-hydroxylase); 4CL (4-coumarate-CoA ligase); CHS (chalcone synthase); CHR (chalcone reductase); CHI (chalcone isomerase); IFS (isoflavone synthase); I2'H (isoflavone 2'-hydroxylase); 2HDR (2'-hydroxydaidzein oxidoreductase); PTS (pterocarpan synthase); 3,9DPO (3,9-dihydroxypterocarpan 6a-monooxygenase), G4DT (glycinol-4-dimethylallyltransferase), G2DT (glycinol-2-dimethylallyltransferase), GS (glyceollin synthase)

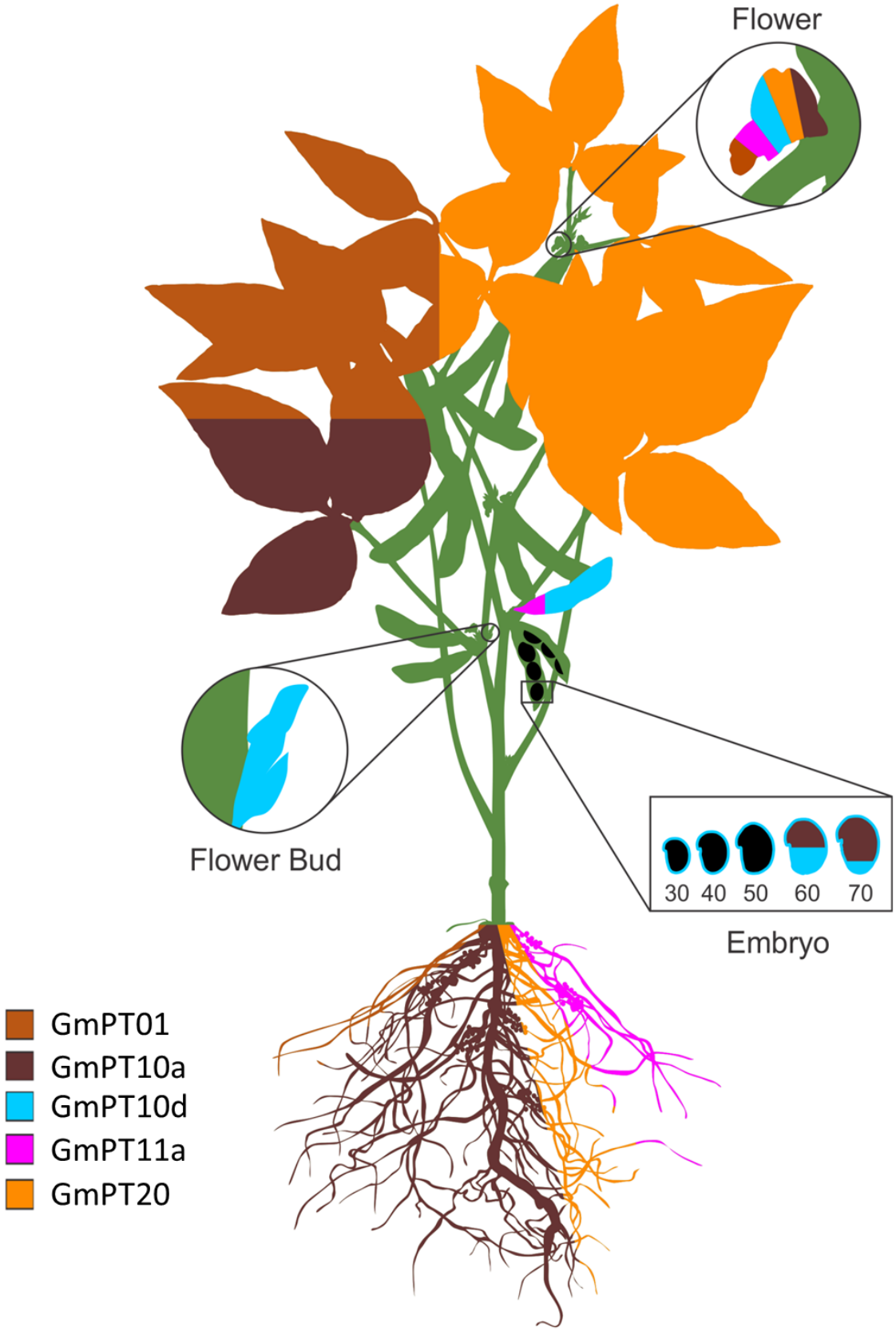
4.8 Isoflavonoid biosynthesis in response to pathogen infection

The culmination of the research undertaken by this project has widened the image for isoflavonoid biosynthesis in plant defense. The already known model is upon pathogen infection, an elicitor molecule is secreted by the pathogen (Shan et al., 2004), which is proposed to initiate the plant defense response (i.e. isoflavonoid production). The existence of an isoflavonoid metabolon (Dastmalchi et al., 2016) would represent a mechanism to rapidly and efficiently produce metabolites. As a result, isoflavonoids would work to prevent the spread of infection.

This study identified five *GmPTs* proposed to function in glyceollin biosynthesis. The interesting finding being that they show varied spatial distribution within the plant (Figure 3.6). This distribution may be a mechanism by which the plant selectively responds to an infection. The isoflavonoid metabolon would be expected to be localized ubiquitously throughout the plant; however, the *GmPTs* that comprise these metabolons would differ according to the tissue. The proposed model is visually represented in Figure 4.2 where the distribution across twelve soybean tissues is displayed. No *GmPTs* were expressed basally in stems, embryo 30, embryo 40 or embryo 50. However, as demonstrated in the pathogen infection study, *GmPTs* were found to be expressed in stems upon infection. Corroborating findings in past studies that indicate that glyceollins accumulate at the site of infection (Hahn et al., 1985). During infection, through an unknown mechanism, transcribed *GmPTs* would be transported to the site of infection to produce glyceollin.

An interesting correlation was that glyceollin II and III was shown to gradually accumulation during pathogen infection (Bhattacharyya and Ward, 1986), furthermore, *GmPT20* similarly gradually increased across the infection time course (Figure 3.5). Glyceollin III is the most predominant isomer present in soybean leaves upon infection, which further correlates to the expression pattern of *GmPT20*, with its highest accumulation in leaf compared to other tissues. This association may present *GmPT20* as a likely candidate for G2DT.

Figure 4.2 Isoflavonoid-specific prenyltransferase distribution in soybean. The proposed model suggests that isoflavonoid-specific prenyltransferases are spatially distributed and respond according to the proximity of the stress. In this model, specific GmPTs would be expressed predominantly in the designated sites. The model highlights twelve soybean tissues: root, stem, leaf, flower bud, flower, embryo 30 to 70 days after pollination (DAP), seed coat, and pod wall. Presence and relative abundance is highlighted by the correspond colour associated in the legend. Stem and embryo 30-50 DAP remains uncoloured indicating no basal-level expression of any *GmPT*.


Flower
Flower Bud
30 40 50 60 70
Embryo
GmPT01
GmPT10a
GmPT10d
GmPT11a
GmPT20

4.9 Future directions

Soybeans play a well-established role in world today; they are grown to feed humans and feedstock, and they are used as material for industrial products. Soybeans comprise a large component of the economy for countries producing it. The importance of this crop cedes no argument, therefore, it is crucial that mitigation of crop loss be a top priority. Taking advantage of current molecular biological and functional genomic techniques, novel methods to improve understanding of plant defence and pathogen attack could be developed. Through increased knowledge of critical biosynthetic and signalling pathways, research can be undertaken to manipulate plant metabolism to produce a more resistant crop. In regard to soybeans, our underlying hypothesis is that increased isoflavonoid production should strengthen soybean's partial resistance, and promote a long-term solution to a variety of biotic and abiotic stresses.

In attempting to further knowledge pertaining to isoflavonoid biosynthesis, this study has identified and characterized the gene family of prenyltransferases proposed to be involved in this pathway. Eleven *GmPTs* were identified in this study, of these, five GmPTs were localized to plastids and are upregulated in response to pathogen infection. One of these was G4DT, a GmPT already proven to prenylate the C-4 position of glycinol. Enzymatic assay of the five candidate GmPTs was not successful, but confirmation of enzymatic activity is essential to characterization of this gene family. *GmPT20* shows strong similarity to the glyceollin III expression profile. Whereas, *GmPT01* and *GmPT10d* are associated with QTL and QTL markers linked to *P. sojae* resistance in soybean. These three genes are interesting candidates to be putatively involved in isoflavonoid biosynthesis.

To corroborate models proposed in this study, upon confirmation of enzymatic activity, soybean resistance to *P. sojae* infection should be tested in *GmPT* silenced soybean plants. Furthermore, tissue-specific expression of *GmPTs* under biotic or abiotic stress would offer to corroborate the model proposed in Figure 4.2. Protein-protein interaction should be performed to confirm interactions of new added members of the isoflavonoid metabolon. Furthermore, interaction between of GmPTs and an ER-embedded P450s (e.g. IFS or I2'H) would further strengthen the theory of transorganellar channelling.

Better understanding of GmPT localization, expression and catalytic activity can result in more precise manipulation of this gene family. Altered expression designed to increase glyceollin production would offer to produce a more resistant cultivar of soybean. Such a cultivar would offer a long-term solution to crop loss by pathogen infection. Furthermore, with increased knowledge of other soybean specialized metabolism (e.g. suberin, coumarin, ect.) a collaborative approach can be undertaken to manipulate multiple biosynthetic pathways to produce a cultivar with strong partial resistance.

Appendix

Appendix A. QTLs associated with resistance to *P. sojae* infection

QTL	Chromosome	Parents	Reference
Phytoph 1-1	13	Parent 1: Conrad Parent 2: Harosoy	(Tucker et al., 2010)
Phytoph 1-2	2	Parent 1: Conrad Parent 2: Harosoy	(Tucker et al., 2010)
Phytoph 2-1	13	Parent 1: Conrad Parent 2: Sloan	(Burnham et al., 2003)
Phytoph 2-2	2	Parent 1: Conrad Parent 2: Sloan	(Burnham et al., 2003)
Phytoph 3-1	13	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Phytoph 3-2	2	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Phytoph 4-1	13	Parent 1: Conrad Parent 2: Williams	(Burnham et al., 2003)
Phytoph 4-2	2	Parent 1: Conrad Parent 2: Williams	(Burnham et al., 2003)
Phytoph 5-1	6	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Phytoph 5-2	15	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Phytoph 5-3	10	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Phytoph 6-1	13	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-2	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-3	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-4	8	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-5	11	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-6	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-7	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-8	6	Parent 1: Conrad	(Li et al., 2010)

		Parent 2: Hefeng 25	
Phytoph 7-1	16	Parent 1:Conrad Parent 2:OX760-6-1	(Weng et al., 2007)
Phytoph 8-1	16	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
Phytoph 8-2	20	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
Phytoph 8-3	18	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
Phytoph 8-4	13	Parent 1:V71-370 Parent 2:PI407162	(Tucker et al., 2010)
Phytoph 9-1	12	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2010)
Phytoph 9-2	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2010)
Phytoph 9-3	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2010)
Phytoph 9-4	14	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2010)
Phytoph 9-5	17	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2010)
Phytoph 9-6	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2010)
Phytoph 10-1	13	Parent 1: Conrad Parent 2: OX760-6-1	(Han et al., 2008)
Phytoph 10-2	13	Parent 1: Conrad Parent 2: OX760-6-1	(Han et al., 2008)
Phytoph 10-3	2	Parent 1: Conrad Parent 2:OX760-6-1	(Han et al., 2008)
Phytoph 11-1	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-2	8	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-3	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-4	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-5	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-6	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)

Phytoph 11-7	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-8	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-9	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-10	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-11	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-12	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-13	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-14	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-15	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-16	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-17	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-18	N/A	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-19	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-20	17	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-21	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-22	14	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 12-1	13	Parent 1: S99-2281 Parent 2:PI 408105A	(Nguyen et al., 2012)
Phytoph 12-2	17	Parent 1: S99-2281 Parent 2: PI 408105A	(Nguyen et al., 2012)
Phytoph 13-1	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 13-2	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 13-3	1	Parent 1: Conrad	(Wang et al., 2012)

		Parent 2: Sloan	
Phytoph 13-4	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 13-5	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 14-1	1	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-2	13	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-3	18	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-4	2	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-5	3	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-6	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-7	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-8	7	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-9	15	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-10	20	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)

Appendix B QTL markers linked to resistance to *P. sojae* infection

QTL Marker	Chromosome	Parents	Reference
Sat_414	1	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt439	1	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Sat_089	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Sat_423	2	Multiple accession were used	(Sun et al., 2014)
Satt266	2	Multiple accession were used	(Sun et al., 2014)
Satt266	2	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Satt274	2	Parent 1:Conrad Parent 2:OX760-6-1	(Han et al., 2008)
Satt351	2	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt542	2	Multiple accession were used	(Sun et al., 2014)
Satt579	2	Parent 1 : Conrad Parent 2: Williams	(Burnham et al., 2003)
Satt600	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt634	2	Multiple accession were used	(Sun et al., 2014)
Satt009	3	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
A078_1	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt578	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
AW734043	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Staga001	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt079	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt100	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)

Satt100	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_246	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_251	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt307	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt307	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt316	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt322	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt365	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt376	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt520	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt557	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt463	7	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
GMA2_OSU19	8	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt437	8	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt632	8	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Sat_196	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_242	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_273	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_274	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_321	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt094	10	Parent 1 : Su88-M21	(Wu et al., 2011)

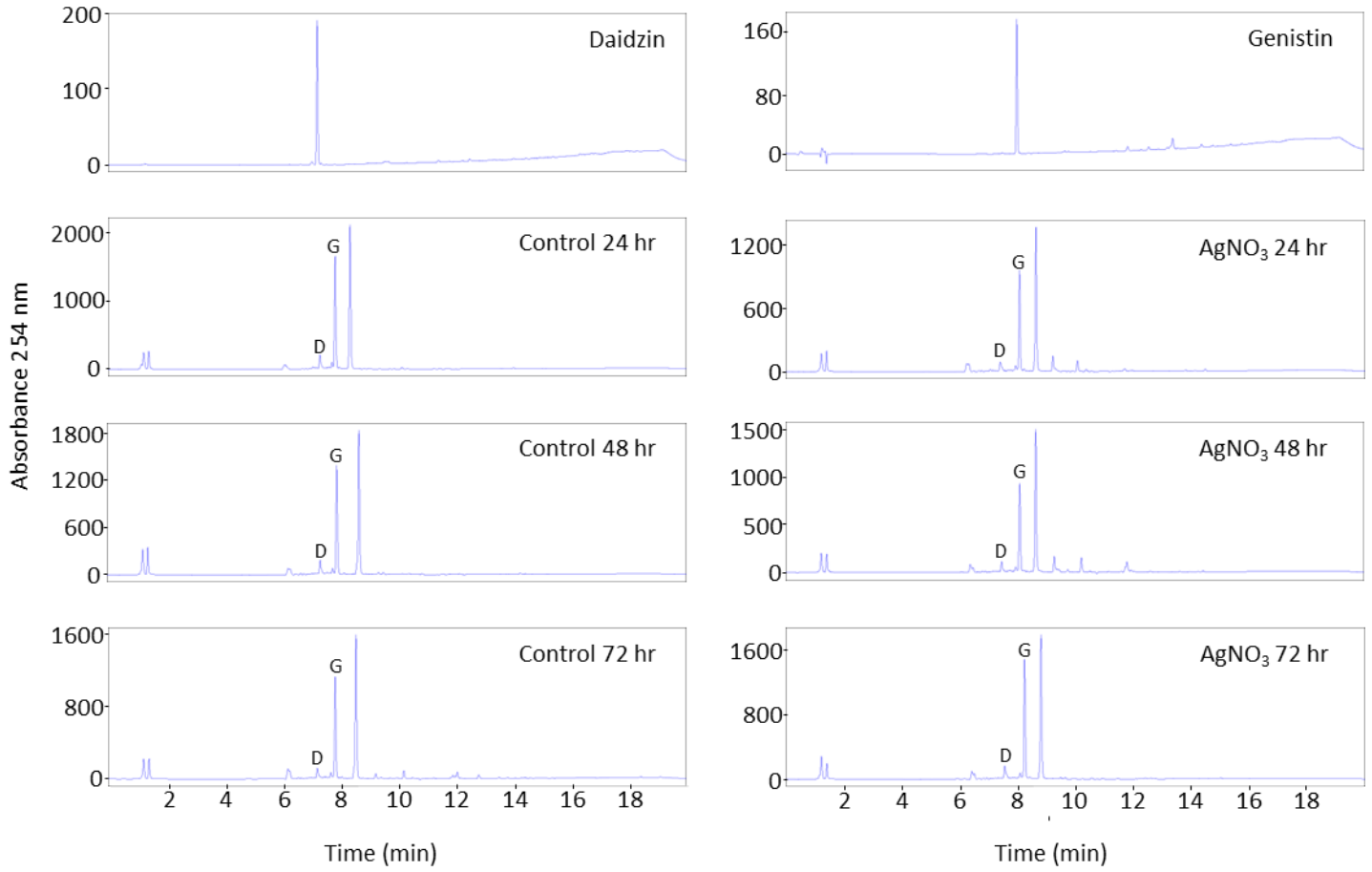
		Parent 2 : Xinyixiaoheidou	
Satt188	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt241	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt345	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt358	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt420	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt445	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt479	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt550	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt576	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt592	10	Parent 1 : Su88-M21	(Wu et al., 2011)
		Parent 2 : Xinyixiaoheidou	
Satt453	11	Parent 1: Conrad	(Li et al., 2010)
		Parent 2: Hefeng 25	
A036_1	12	Parent 1: Conrad	Wang et al. 2010
		Parent 2: Sloan	
GMH_OSU31	12	Parent 1: Conrad	(Wang et al., 2012)
		Parent 2: Sloan	
F424_294	13	Parent 1: Conrad	(Wang et al., 2012)
		Parent 2: Sloan	
K644_1	13	Parent 1: V71-370	(Tucker et al., 2010)
		Parent 2: PI407162	
Sct_033	13	Parent 1: Conrad	Wang et al. 2010
		Parent 2: Sloan	
Sct_033	13	Parent 1: S99-2281	(Nguyen et al., 2012)
		Parent 2: PI 408105A	
Sat_133	13	Parent 1: OX20-8	(Lee et al., 2013)
		Parent 2: PI 398841	
Satt030	13	Parent 1:Conrad	(Han et al., 2008)
		Parent 2:OX760-6-1	
Satt160	13	Parent 1: Conrad	Wang et al. 2010
		Parent 2: Sloan	

Satt252	13	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Satt252	13	Parent 1 : Conrad Parent 2: Williams	(Burnham et al., 2003)
Satt252	13	Parent 1: Conrad Parent 2: Sloan	(Burnham et al., 2003)
Satt252	13	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Satt343	13	Parent 1:Conrad Parent 2:OX760-6-1	(Han et al., 2008)
Satt343	13	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt304	14	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010
Sat_380	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt204	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt267	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt369	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt384	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt384	15	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt403	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt491	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt598	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt651	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt720	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt414	16	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
Satt596	16	Parent 1: Conrad Parent 2: OX760-6-1	(Weng et al., 2007)
Satt_222	17	Multiple accession were used	(Sun et al., 2014)

Satt226	17	Multiple accession were used	(Sun et al., 2014)
Satt300	17	Multiple accession were used	(Sun et al., 2014)
Satt301	17	Multiple accession were used	(Sun et al., 2014)
Satt543	17	Parent 1: S99-2281 Parent 2: PI 408105A	(Nguyen et al., 2012)
Satt543	17	Multiple accession were used	(Sun et al., 2014)
Satt574	17	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010
Satt574	17	Multiple accession were used	(Sun et al., 2014)
Satt615	17	Multiple accession were used	(Sun et al., 2014)
A671_1	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARC-039397-07314	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_18_1707	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_18_1777	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1393	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
GM19_OSU10	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
OPAD08	18	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt472	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt678	18	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
BARC-021321-04035	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARC-03997-07624	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARC-064609-18739	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)

BARCSOYSSR_18_1793	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1243	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1473	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1494	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1532	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
GML_OSU42	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt527	19	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010
Sat_267	20	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)

Appendix C. Metabolite profile of control and $AgNO_3$-treated soybean seeds at 254nm.



References

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

Aisyah, S., Gruppen, H., Madzora, B., and Vincken, J.P. (2013). Modulation of isoflavonoid composition of Rhizopus oryzae elicited soybean (Glycine max) seedlings by light and wounding. J. Agric. Food Chem. **61,** 8657-8667.

Akada, S., and Dube, S.K. (1995). Organization of soybean chalcone synthase gene clusters and characterization of a new member of the family. Plant Mol. Biol. **29,** 189-199.

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 Lett. **431,** 287-290.

Akashi, T., Sasaki, K., Aoki, T., Ayabe, S.-i., and Yazaki, K. (2009). Molecular Cloning and Characterization of a cDNA for Pterocarpan 4-Dimethylallyltransferase Catalyzing the Key Prenylation Step in the Biosynthesis of Glyceollin, a Soybean Phytoalexin. Plant Physiol. **149,** 683-693.

Akiyama, T., Ishida, J., Nakagawa, S., Ogawara, H., Watanabe, S., Itoh, N., Shibuya, M., and Fukami, Y. (1987). Genistein, a specific inhibitor of tyrosine-specific protein kinases. J. Biol. Chem. **262,** 5592-5595.

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

Andersson, M.X., Goksör, M., and Sandelius, A.S. (2007). Optical Manipulation Reveals Strong Attracting Forces at Membrane Contact Sites between Endoplasmic Reticulum and Chloroplasts. J. Biol. Chem. **282,** 1170-1174.

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

Arjmandi, B.H., Alekel, L., Hollis, B.W., Amin, D., Stacewicz-Sapuntzakis, M., Guo, P., and Kukreja, S.C. (1996). Dietary soybean protein prevents bone loss in an ovariectomized rat model of osteoporosis. J. Nutr. **126,** 161-167.

Ayers, R.A., Jürgen, E., Finelli, R., Berger, N., and Peter, A. (1976). Host-Pathogen Interactions: IX. Quantitative Assays of Elicitor Activity and Characterization of the Elicitor Present in the Extracellular Medium of Cultures of Phytophthora megasperma var. sojae. Plant Physiol. **57,** 751-759.

Balasubramanian, V., Vashisht, D., Cletus, J., and Sakthivel, N. (2012). Plant β-1,3-glucanases: Their biological functions and transgenic expression against phytopathogenic fungi. Biotechnol. Lett. **34,** 1983-1990.

Bernards, M.A. (2002). Demystifying suberin. Can. J. Bot. **80,** 227-240.

Bhattacharyya, M.K., and Ward, E.W.B. (1986). Resistance, susceptibility and accumulation of glyceollins I–III in soybean organs inoculated with Phytophthora megasperma f. sp. glycinea. Physiol. Mol. Plant Pathol. **29,** 227-237.

Bohrer, A.S., Yoshimoto, N., Sekiguchi, A., Rykulski, N., Saito, K., and Takahashi, H. (2015). Alternative translational initiation of ATP sulfurylase underlying dual localization of sulfate assimilation pathways in plastids and cytosol in Arabidopsis thaliana. Front. Plant Sci. **5**.

Boue, S.M., Tilghman, S.L., Elliott, S., Zimmerman, M.C., Williams, K.Y., Payton-Stewart, F., Miraflor, A.P., Howell, M.H., Shih, B.Y., Carter-Wientjes, C.H., Segar, C., Beckman, B.S., Wiese, T.E., Cleveland, T.E., McLachlan, J.A., and Burow, M.E. (2009). Identification of the potent phytoestrogen glycinol in elicited soybean (glycine max). Endocrinology **150,** 2446-2453.

Boydston, R., Paxton, J.D., and Koeppe, D.E. (1983). Glyceollin: a site-specific inhibitor of electron transport in isolated soybean mitochondria. Plant Physiol. **72,** 151-155.

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

Bräuer, L., Brandt, W., Schulze, D., Zakharova, S., and Wessjohann, L. (2008). A structural model of the membrane-bound aromatic prenyltransferase UbiA from E. coli. ChemBioChem **9,** 982-992.

Burbulis, I.E., and Winkel-Shirley, B. (1999). Interactions among enzymes of the Arabidopsis flavonoid biosynthetic pathway. Proc. Natl. Acad. Sci. U. S. A. **96,** 12929-12934.

Catharius, K., Frech, K., Groet, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., and Werner, T. (2005). MatInspector and beyond:promoter analysis based on transcription factor binding sites. Bioinformatics **21,** 2933-2942.

Cheng, G., Wilczek, B., Warner, M., Gustafsson, J.A., and Landgren, B.M. (2007). Isoflavone treatment for acute menopausal symptoms. Menopause **14,** 468-473.

Czichi, U., and Kindl, H. (1975). Formation of p-coumaric acid and o-coumaric acid from l-phenylalanine by microsomal membrane fractions from potato: Evidence of membrane-bound enzyme complexes. Planta **125,** 115-125.

Czichi, U., and Kindl, H. (1977). Phenylalanine ammonia lyase and cinnamic acid hydroxylases as assembled consecutive enzymes on microsomal membranes of cucumber cotyledons: Cooperation and subcellular distribution. Planta **134,** 133-143.

Dakora, F.D., and Phillips, D.A. (1996). Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins. Physiol. Mol. Plant Pathol. **49,** 1-20.

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. Plant J. **85,** 689-706.

De Coninck, B., Timmermans, P., Vos, C., Cammue, B.P.A., and Kazan, K. (2015). What lies beneath: belowground defense strategies in plants. Trends Plant Sci. **20,** 91-101.

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

Dhaubhadel, S., McGarvey, B.D., Williams, R., and Gijzen, M. (2003). Isoflavonoid biosynthesis and accumulation in developing soybean seeds. Plant Mol. Biol. **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 Physiol. **143,** 326-338.

Dong, S., Yin, W., Kong, G., Yang, X., Qutob, D., Chen, Q., Kale, S.D., Sui, Y., Zhang, Z., Dou, D., Zheng, X., Gijzen, M., M. Tyler, B., and Wang, Y. (2011). *Phytophthora sojae* Avirulence Effector Avr3b is a Secreted NADH and ADP-ribose Pyrophosphorylase that Modulates Plant Immunity. PLoS Pathog. **7,** e1002353.

Dorrance, A.E., McClure, S.A., and St. Martin, S.K. (2003). Effect of partial resistance on *Phytophthora* stem rot incidence and yield of soybean in Ohio. Plant Dis. **87,** 308-312.

Dou, D., Kale, S.D., Wang, X., Chen, Y., Wang, Q., Wang, X., Jiang, R.H.Y., Arredondo, F.D., Anderson, R.G., Thakur, P.B., McDowell, J.M., Wang, Y., and Tyler, B.M. (2008). Conserved C-terminal motifs required for avirulence and suppression of cell death by Phytophthora sojae effector Avr1b. Plant Cell **20,** 1118-1133.

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

Emanuelsson, O., Nielsen, H., and Von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. **8,** 978-984.

Endo, A., Kimura, M., Kawakami, N., and Nambara, E. (2011). Functional Analysis of Abscisic Acid 8′-Hydroxylase. In Seed Dormancy, A.R. Kermode, ed (Humana Press), pp. 135-147.

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

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

Fry, W.E. (1982). Principles of Plant Disease Management. (Academic Press).

Giannini, J.L., Briskin, D.P., Holt, J.S., and Paxton, J.D. (1988). Inhibition of plasma membrane and tonoplast H+-transporting ATPases by glyceollin. Phytopathology **78,** 1000-1003.

GlobalSoybeanProduction.com. (2016). Global Soybean Production http://www.globalsoybeanproduction.com <accessed 2016/07/22>.

Glover, K.D., and Scott, R.A. (1998). Heritability and phenotypic variation of tolerance to Phytophthora root rot of soybean. Crop Sci. **38,** 1495-1500.

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 Res. **40,** D1178-1186.

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. Mol. Plant-Microbe Interact. **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 Physiol. **144,** 728-740.

Grant, D., Nelson, R.T., Cannon, S.B., and Shoemaker, R.C. (2010). SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res. **38,** D843-D846.

Hahn, M.G., Bonhoff, A., and Grisebach, H. (1985). Quantitative Localization of the Phytoalexin Glyceollin I in Relation to Fungal Hyphae in Soybean Roots Infected with Phytophthora megasperma f. sp. glycinea. Plant Physiol. **77,** 591-601.

Han, Y., Teng, W., Yu, K., Poysa, V., Anderson, T., Qiu, L., Lightfoot, D.A., and Li, W. (2008). Mapping QTL tolerance to Phytophthora root rot in soybean using microsatellite and RAPD/SCAR derived markers. Euphytica **162,** 231-239.

Hanada, K., Zou, C., Lehti-Shiu, M.D., Shinozaki, K., and Shiu, S.H. (2008). Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiol. **148,** 993-1003.

Hanada, K., Kuromori, T., Myoga, F., Toyoda, T., Li, W.-H., and Shinozaki, K. (2009). Evolutionary persistence of functional compensation by duplicate genes in Arabidopsis. Genome Biol. Evol.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: Protein localization predictor. Nucleic Acids Res. **35,** W585-W587.

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

Hrazdina, G., and Wagner, G.J. (1985). Metabolic pathways as enzyme complexes: Evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. Arch. Biochem. Biophys. **237,** 88-100.

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

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 Nat. Biotechnol. **18,** 559-559.

Kaplan, D.T., Keen, N.T., and Thomason, I.J. (1980). Studies on the mode of action of glyceollin in soybean incompatibility to the root knot nematode, Meloidogyne incognita. Physiol. Plant Pathol. **16,** 319-325.

Karamat, F., Olry, A., Munakata, R., Koeduka, T., Sugiyama, A., Paris, C., Hehn, A., Bourgaud, F., and Yazaki, K. (2014). A coumarin-specific prenyltransferase catalyzes the crucial biosynthetic reaction for furanocoumarin formation in parsley. Plant J. **77,** 627-638.

Kaufmann, M.J., and Gerdemann, J.W. (1958). Root and stem rot of soybean caused by *Phytophthora sojae* n. sp. Phytopathology **48,** 201-208.

Kimpel, J.A., and Kosuge, T. (1985). Metabolic Regulation during Glyceollin Biosynthesis in Green Soybean Hypocotyls. Plant Physiol. **77,** 1-7.

Kochs, G., and Grisebach, H. (1986). Enzymic synthesis of isoflavones. Eur J Biochem **155,** 311-318.

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. Epidemiol. Biomarkers. Prev. **18,** 1050-1059.

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

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E.L.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. **305,** 567-580.

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). Agric. Biol. Chem. **55,** 2227-2233.

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.

Li, W. (2016). Bringing Bioactive Compounds into Membranes: The UbiA Superfamily of Intramembrane Aromatic Prenyltransferases. Trends Biochem. Sci.

Liang, P.-H., Ko, T.-P., and Wang, A.H.J. (2002). Structure, mechanism and function of prenyltransferases. Eur. J. Biochem. **269,** 3339-3354.

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

Lohr, M., Schwender, J., and Polle, J.E.W. (2012). Isoprenoid biosynthesis in eukaryotic phototrophs: A spotlight on algae. Plant Sci. **185-186,** 9-22.

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 Physiol. Biochem. **42,** 671-679.

Lulai, E.C., and Corsini, D.L. (1998). Differential deposition of suberin phenolic and aliphatic domains and their roles in resistance to infection during potato tuber (Solanum tuberosum L.) wound-healing. Physiol. Mol. Plant Pathol. **53,** 209-222.

MacKay, T.F.C., Stone, E.A., and Ayroles, J.F. (2009). The genetics of quantitative traits: Challenges and prospects. Nature Reviews Genetics **10,** 565-577.

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. Theor. Appl. Genet. **110,** 1203 - 1209.

McCarthy, E.W., Mohamed, A., and Litt, A. (2015). Functional divergence of APETALA1 and FRUITFULL is due to changes in both regulation and coding sequence. Front. Plant Sci. **6**.

Mehrshahi, P., Stefano, G., Andaloro, J.M., Brandizzi, F., Froehlich, J.E., and DellaPenna, D. (2013). Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. Proc. Natl. Acad. Sci. U. S. A. **110,** 12126-12131.

Messina, M. (2010a). A brief historical overview of the past two decades of soy and isoflavone research. J. Nutr. **140,** 1350S-1354S.

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

Messina, M., and Messina, V. (2010). The role of soy in vegetarian diets. Nutrients **2,** 855-888.

Messina, M.J. (1999). Legumes and soybeans: Overview of their nutritional profiles and health effects. Am. J. Clin. Nutr. **70,** 439s-450s.

Mideros, S., Nita, M., and Dorrance, A.E. (2007). Characterization of Components of Partial Resistance, Rps2, and Root Resistance to Phytophthora sojae in Soybean. Phytopathology **97,** 655-662.

Morris, P.F., and Ward, E. (1992). Chemoattraction of zoospores of the soybean pathogen, *Phytophthora sojae*, by isoflavones. Physiol. Mol. Plant Pathol. **40,** 17-22.

Moy, P., Qutob, D., Chapman, P., Atkinson, I., and Gijzen, M. (2004). Patterns of gene expression upon infection of soybean plants by Phytophthora sojae. Mol. Plant-Microbe Interact. **17,** 1051-1062.

Ohara, K., Muroya, A., Fukushima, N., and Yazaki, K. (2009). Functional characterization of LePGT1, a membrane-bound prenyltransferase involved in the geranylation of p-hydroxybenzoic acid. Biochem. J. **421,** 231-241.

OMAFRA. (2014). Ontario Ministry of Agriculture, Food and Rural Affairs. Crops: Soybean Production in Ontario http://www.omafra.gov.on.ca/english/crops/field/soybeans.html <accessed 2016/07/22>.

Omasits, U., Ahrens, C.H., Müller, S., and Wollscheid, B. (2014). Protter: Interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics **30,** 884-886.

Panchy, N., Lehti-Shiu, M.D., and Shiu, S.-H. (2016). Evolution of gene duplication in plants. Plant Physiol.

Pflieger, S., Lefebvre, V., and Causse, M. (2001). The candidate gene approach in plant genetics: a review. Mol. Breed. **7,** 275-291.

Pisani, P., Bray, F., and Parkin, M.D. (2002). Estimates of the world-wide prevalence of cancer for 25 sites in the adult population. Int. J. 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. Comp. Med. **56,** 461-467.

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

Ralston, L., and Yu, O. (2006). Metabolons involving plant cytochrome P450s. Phytochem. Rev. **5,** 459-472.

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 Physiol. **137,** 1375-1388.

Rasmussen, S., and Dixon, R.A. (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. Plant Cell **11,** 1537-1551.

Rodríguez, F.I., Esch, J.J., Hall, A.E., Binder, B.M., Eric Schaller, G., and Bleecker, A.B. (1999). A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. Science **283,** 996-998.

Saleh, O., Haagen, Y., Seeger, K., and Heide, L. (2009). Prenyl transfer to aromatic substrates in the biosynthesis of aminocoumarins, meroterpenoids and phenazines: The ABBA prenyltransferase family. Phytochemistry **70,** 1728-1738.

Sasaki, K., Mito, K., Ohara, K., Yamamoto, H., and Yazaki, K. (2008). Cloning and characterization of naringenin 8-prenyltransferase, a flavonoid-specific prenyltransferase of Sophora flavescens. Plant Physiol. **146,** 1075-1084.

Saslowsky, D., and Winkel-Shirley, B. (2001). Localization of flavonoid enzymes in Arabidopsis roots. Plant J. **27,** 37-48.

Schenck, C.A., Chen, S., Siehl, D.L., and Maeda, H.A. (2015). Non-plastidic, tyrosine-insensitive prephenate dehydrogenases from legumes. Nat. Chem. Biol. **11,** 52-57.

Schmitthenner, A.F. (1985). Problems and progress in control of *Phytophthora* root rot of soybean. Plant Dis. **69,** 362-368.

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.

Schopfer, C.R., Kochs, G., Lottspeich, F., and Ebel, J. (1998). Molecular characterization and functional expression of dihydroxypterocarpan 6a-hydroxylase, an enzyme specific for pterocarpanoid phytoalexin biosynthesis in soybean (Glycine max L.). FEBS Lett. **432,** 182-186.

Senda, M., Jumonji, A., Yumoto, S., Ishikawa, R., Harada, T., Niizeki, M., and Akada, S. (2002). Analysis of the duplicated CHS1 gene related to the suppression of the seed coat pigmentation in yellow soybeans. Theor. Appl. Genet. **104,** 1086-1091.

Sepiol, C. (2015). Planting the Chalcone Reductase Family Tree: Identification and Characterization of Chalcone Reductase Genes in Soybean. In Department of Biology (*Electronic Thesis and Dissertation Repository*: Western University).

Shan, W., Cao, M., Leung, D., and Tyler, B.M. (2004). The Avr1b locus of Phytophthora sojae encodes an elicitor and a regulator required for avirulence on

soybean plants carrying resistance gene Rps1b. Mol. Plant-Microbe Interact. **17,** 394-403.

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. Epidemiol. Biomarkers. Prev. **10,** 483-488.

Simons, R., Vincken, J.P., Bohin, M.C., Kuijpers, T.F.M., Verbruggen, M.A., and Gruppen, H. (2011). Identification of prenylated pterocarpans and other isoflavonoids in *Rhizopus* spp. elicited soya bean seedlings by electrospray ionisation mass spectrometry. Rapid Commun. Mass Spectrom. **25,** 55-65.

Smit, G., Puvanesarajah, V., Carlson, R.W., Barbour, W.M., and Stacey, G. (1992). Bradyrhizobium japonicum nodD1 can be specifically induced by soybean flavonoids that do not induce the nodYABCSUIJ operon. J. Biol. Chem. **267,** 310-318.

SoyCanada. (2016). Canada's Growing Soybean Industry http://soycanada.ca/industry/industry-overview/ <accessed 2016/07/22>.

Srere, P.A. (1985). The metabolon. Trends Biochem. Sci. **10,** 109-110.

Stafford, H.A. (1974). The Metabolism of Aromatic Compounds. Annu. Rev. Plant Physiol. **25,** 459-486.

Stössel, P. (1982). Glyceollin Production in Soybean. J. Phytopathol. **105,** 109-119.

Subramanian, S., Stacey, G., and Yu, O. (2006). Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. Plant J. **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 Physiol. **137,** 1345-1353.

Sugimoto, T., Kato, M., Yoshida, S., Matsumoto, I., Kobayashi, T., Kaga, A., Hajika, M., Yamamoto, R., Watanabe, K., Aino, M., Matoh, T., Walker, D.R., Biggs, A.R., and Ishimoto, M. (2011). Pathogenic diversity of Phytophthora sojae and breeding strategies to develop Phytophthora-resistant soybeans. Breeding Science **61,** 511-522.

Tahara, S., and Ibrahim, R.K. (1995). Prenylated isoflavonoids-An update. Phytochemistry **38,** 1073-1094.

Takahashi, S., and Koyama, T. (2006). Structure and function of cis-prenyl chain elongating enzymes. Chem. Rec. **6,** 194-205.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular biology and evolution **30,** 2725-2729.

Thomas, R., Fang, X., Ranathunge, K., Anderson, T.R., Peterson, C.A., and Bernards, M.A. (2007). Soybean root suberin: anatomical distribution, chemical composition, and relationship to partial resistance to *Phytophthora sojae*. Plant Physiol. **144,** 299-311.

Tian, L., DellaPenna, D., and Dixon, R.A. (2007). The pds2 mutation is a lesion in the Arabidopsis homogentisate solanesyltransferase gene involved in plastoquinone biosynthesis. Planta **226,** 1067-1073.

Tucker, D.M., Saghai Maroof, M.A., Mideros, S., Skoneczka, J.A., Nabati, D.A., Buss, G.R., Hoeschele, I., Tyler, B.M., St. Martin, S.K., and Dorrance, A.E. (2010). Mapping quantitative trait loci for partial resistance to in a soybean interspecific cross. Crop Sci. **50,** 628-635.

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

Tyler, B.M. (2007). Phytophthora sojae: Root rot pathogen of soybean and model oomycete. Mol. Plant Pathol. **8,** 1-8.

Vega-Sánchez, M.E., Redinbaugh, M.G., Costanzo, S., and Dorrance, A.E. (2005). Spatial and temporal expression analysis of defense-related genes in soybean cultivars with different levels of partial resistance to Phytophthora sojae. Physiol. Mol. Plant Pathol. **66,** 175-182.

Wang, C.S., and Vodkin, L.O. (1994). Extraction of RNA from tissues containing high levels of procyanidins that bind RNA. Plant Mol. Biol. Rep. **12,** 132-145.

Wang, J., Chu, S., Zhu, Y., Cheng, H., and Yu, D. (2015). Positive selection drives neofunctionalization of the UbiA prenyltransferase gene family. Plant Mol. Biol. **87,** 383-394.

Ward, E., Lazarovits, G., Unwin, C., and Buzzell, R. (1979). Hypocotyl reactions and glyceollin in soybeans inoculated with zoospores of *Phytophthora megasperma* var. *sojae*. Phytopathology.

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

Winkel, B.S.J. (2004). Metabolic channeling in plants. Annu. Rev. Plant Biol. **55,** 85-107.

Winkel, B.S.J. (2009). Metabolite Channeling and Multi-enzyme Complexes**,** 195-208.

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. J. Natl. Cancer Inst. **101,** 1141-1155.

Yang, T., Fang, L., Rimando, A.M., Sobolev, V., Mockaitis, K., and Medina-Bolivar, F. (2016). Characterization of stilbenoid-specific prenyltransferase activity in peanut hairy roots. Plant Physiol.

Yazaki, K., Kunihisa, M., Fujisaki, T., and Sato, F. (2002). Geranyl diphosphate:4-hydroxybenzoate geranyltransferase from Lithospermum erythrorhizon. Cloning and characterization of a ket enzyme in shikonin biosynthesis. J. Biol. Chem. **277,** 6240-6246.

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

Yi, S.Y., and Hwang, B.K. (1996). Differential induction and accumulation of β-1,3-glucanase and chitinase isoforms in soybean hypocotyls and leaves after compatible and incompatible infection with Phytophthora megasperma f.sp. glycinea. Physiol. Mol. Plant Pathol. **48,** 179-192.

Young, I.G., Leppik, R.A., Hamilton, J.A., and Gibson, F. (1972). Biochemical and genetic studies on ubiquinone biosynthesis in Escherichia coli K-12:4-hydroxybenzoate octaprenyltransferase. J. Bacteriol. **110,** 18-25.

Young, N.D. (1999). A cautiously optimistic vision for marker-assisted breeding. Mol. Breed. **5,** 505-510.

Zähringer, U., Ebel, J., Mulheirn, L.J., Lyne, R.L., and Grisebach, H. (1979). Induction of phytoalexin synthesis in soybean. Dimethylallylpyrophosphate:trihydroxypterocarpan dimethylallyl transferase from elicitor-induced cotyledons. FEBS Lett. **101,** 90-92.

Zhan, S., and Ho, S.C. (2005). Meta-analysis of the effects of soy protein containing isoflavones on the lipid profile. Am. J. Clin. Nutr. **81,** 397-408.

Zou, C., Lehti-Shiu, M.D., Thomashow, M., and Shiu, S.H. (2009). Evolution of stress-regulated gene expression in duplicate genes of Arabidopsis thalianae. PLoS Genet. **5**.

Curriculum Vitae

Arjun Sukumaran

EDUCATION

2014 – Present	M.Sc candidate Biology	Western University London, Ontario, Canada
2009 – 2013	B.Sc (Honours) Cell and Molecular Biology	University of Toronto Toronto, Ontario. Canada

AWARDS AND SCHOLARSHIPS

2014 – Present	Western Graduate Research Scholarship

WORK EXPERIENCE

2014 – Present	Graduate Research Assistant, Western University
2014 – 2015	Graduate Teaching Assistant Western University

PUBLICATIONS

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

CONFERENCES

Sukumaran, A., and Dhaubhadel, S. (2016). Characterization of isoflavonoid-specific prenyltransferases for stem and root rot resistance in Soybean. Plant Biotech 2016, Queen's University, Kingston, Ontario.

Sukumaran, A., and Dhaubhadel, S. (2015). A systematic characterization of the isoflavonoid-specific prenyltransferase gene family in Soybean. 2015 CSPB Eastern Regional Meeting, University of Toronto, Toronto, Ontario.

Sukumaran, A., and Dhaubhadel, S. (2015). Identification and Characterization of the Isoflavonoid-specific Prenyltransferases gene family in Soybean. Biology Graduates Research Forum (BGRF 2015), London, Ontario.