

Electronic Thesis and Dissertation Repository

---

3-15-2024 9:30 AM

## Identification and characterization of isoflavone reductase family members in soybean

Negin Azizkhani, *The University of Western Ontario*

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

: Susanne Kohalmi, *The University of Western Ontario*

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

© Negin Azizkhani 2024

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Cell Biology Commons](#), and the [Plant Biology Commons](#)

---

### Recommended Citation

Azizkhani, Negin, "Identification and characterization of isoflavone reductase family members in soybean" (2024). *Electronic Thesis and Dissertation Repository*. 9982.

<https://ir.lib.uwo.ca/etd/9982>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## **Abstract**

Soybean's yield is threatened by *Phytophthora sojae*, a pathogen responsible for stem and root rot disease. Glyceollins, unique antimicrobial agents specific to soybeans in partially preventing *P. sojae* infection, are derived from the isoflavonoid branch of the general phenylpropanoid pathway. One pivotal enzyme exclusively involved in glyceollin synthesis in soybean is the isoflavone reductase (*GmIFR*), which catalyzes the 2'-hydroxydaidzein conversion to 2'-hydroxy-2,3-dihydrodaidzein as a precursor for glyceollin biosynthesis. To comprehensively identify all members of the *GmIFR* gene family within the soybean genome, keyword and blast protein searches were conducted, identifying 98 putative *GmIFRs*. Among these candidates, seven *GmIFR* candidates were selected for further investigation of which six were confirmed to be localized to the cytoplasm. Additionally, *GmIFR* candidates exhibited soluble expressions and were successfully purified. These findings provide a fundamental knowledge of *GmIFR* family members and their functional characterizations in the glyceollin pathway in soybean.

## **Keywords**

Soybean, *Phytophthora sojae*, Glyceollin pathway, Isoflavone reductase, Subcellular localization, Protein purification, Enzyme assay

## Summary for Lay Audience

As the global population is continuously growing, there is a steady rise in the demand for food resources. Nowadays, environmental conservation, greenhouse gas emissions reduction, and preservation of natural resources such as water and land direct people towards consuming plant-based foods and products. Soybean, with its high protein and oil content, is a globally important agricultural valuable source of both human and animal nutrition, human health, and industrial products. However, various stresses, particularly pathogens like *Phytophthora sojae*, severely impact soybean cultivation, leading to substantial annual losses in crop yield. To address this issue, breeders have developed multiple strategies including soil tillage, seed treatments with fungicides, and crop rotation with non host plant-pathogen. However, these approaches prove somewhat less effective due to the pathogen's long-term survival in soil for more than ten years. At the molecular level, glyceollin pathway in soybean is involved in response to various stress including disease, nutritional limitations, drought, and flooding, by creating a defense mechanism involving many enzymes such as isoflavone reductase. So, it is critical to determine all isoflavone reductase members which play key roles in glyceollin defense pathway in response to *P. sojae* infection.

In the current study, by comprehensive identification of all members of the *isoflavone reductase* gene family within the soybean genome, I found seven candidates which are clustered with identified isoflavone reductase in other legumes plants. Analysis of datasets unveiled that *GmIFR* candidates possess three conserved motifs, exhibit transcript expression in infected soybean tissues, and high expression levels in the root. By analyzing functional characterization of *GmIFR* proteins, I discovered that *GmIFRs* in glyceollin pathway are localized in cytoplasm, and they are also expressed in soluble form.

Knowing *GmIFR* family members and its functional characterizations will provide a fundamental knowledge of this family in the glyceollin pathway and help to develop a more profound understanding of soybean's natural resistance against *P. sojae* infection, resulting in increased glyceollin production, and save millions of dollars of annual soybean losses globally.

## **Acknowledgments**

I would like to thank my supervisor Dr. Sangeeta Dhaubhadel for her support, guidance, and experience, providing direction to my research, and helping me write and organize my research into this thesis.

I extend my gratitude to Dr. Susanne Kohalmi, my co-supervisor, for her consistent support and willingness to assist whenever I needed help. I found her so kind to help me to improve my skills and overcome my constraints. I would also like to thank my advisory committee members, Dr. Frédéric Marsolais and Dr. Mark Bernards, for their helpful advice, feedback and guidance throughout my project and meetings.

Thank you to my parents, Sorayya and Hossein, for always supporting me and taking an interest in my studies. I know I couldn't have done it without you. I would particularly like to thank my sweet sister, Negar, for her constant encouragement and help during the challenges of this academic journey. The support from them is incredibly meaningful to me.

Lastly, I would like to thank my lab members (Kufлом, Nishat, Praveen, Liz, Hodan, Deepak, and Muhammad Sufyan) for their helpful advice and helps. Thank you particularly to Kufлом, my lab technician for his unreserved help and valuable experiences. I am also grateful for the support of my friends Vida, Gigi, Nishat, Praveen, and Shabnam.

In addition, thank you to Tim McDowell for his help in doing enzyme assay, and to Dr. Patrick Telmer for protein purification troubleshooting at Agriculture and Agri-Food Canada (AAFC).

## Table of Contents

Abstract.....	ii
Summary for Lay Audience.....	iii
Acknowledgments.....	v
Table of Contents.....	vi
List of Tables.....	ix
List of Figures.....	x
List of Abbreviations.....	xi
Chapter 1.....	1
1 Introduction.....	1
1.1 Soybean.....	1
1.2 Origin and global journey of soybeans.....	2
1.3 Soybean genome.....	3
1.4 <i>Phytophthora sojae</i> causes stem and root rot disease in soybean.....	4
1.5 Different plant's defense mechanisms to provide resistance.....	5
1.6 Strategies for managing <i>P. sojae</i> in soybean cultivation.....	6
1.7 Soybean resistance to <i>P. sojae</i> : complete and partial mechanisms.....	7
1.8 Isoflavonoids.....	9
1.8.1 Unlocking the power of isoflavonoids in soybean growth and defense.....	9
1.8.2 Isoflavonoid is a nutritional source for human health.....	10
1.8.3 Isoflavonoid biosynthetic pathway.....	12
1.9 Glyceollins as a soybean defense mechanism in response to stresses.....	15
1.10Oxidoreductases family.....	16
1.10.1 Structure and function of isoflavone reductase (IFR).....	17
1.11Hypothesis and objective.....	21

Chapter 2.....	22
2 Materials and Methods.....	22
2.1 Plant materials and growth conditions.....	22
2.2 Bacterial strains and growth conditions.....	22
2.3 Identification of IFR candidates.....	23
2.4 Cloning and transformation .....	24
2.4.1 Cloning into the Gateway entry vector.....	25
2.4.2 Cloning into the Gateway destination vector.....	30
2.5 Subcellular localization.....	36
2.5.1 Transient expression of proteins in <i>N. benthamiana</i> leaves .....	36
2.5.2 Confocal microscopy .....	37
2.6 Protein expression.....	38
2.6.1 Bacterial cultivation and protein expression.....	38
2.6.2 Protein extraction and confirmation on SDS polyacrylamide gel .....	38
2.6.3 Western blot analysis .....	39
2.6.4 Protein purification .....	40
2.6.5 Desalting and buffer exchange.....	41
2.6.6 Cleavage of His <sub>6</sub> tag from <i>GmIFR</i> candidates.....	42
2.7 Spectrophotometric enzyme assay.....	42
2.8 HPLC analysis .....	43
Chapter 3.....	44
3 RESULTS .....	44
3.1 Identification of <i>GmIFR</i> candidates in soybean based on <i>in silico</i> analysis.....	44
3.1.1 Phylogenetic analysis clustered ten <i>GmIFRs</i> candidates with identified IFRs in legumes .....	44
3.1.2 <i>GmIFRs</i> contain three conserved motifs.....	46

3.1.3	<i>GmIFR</i> genes family members display tissue-specific gene expression ..	46
3.1.4	The soybean genome has seven <i>GmIFR</i> candidates .....	55
3.2	Six <i>GmIFR</i> candidates localize in cytoplasm .....	55
3.3	Five <i>GmIFR</i> candidates were expressed and purified.....	62
3.4	Spectrophotometric enzyme assay .....	62
Chapter 4	.....	75
4	Discussion .....	75
4.1	Seven <i>GmIFR</i> candidates were identified in soybean.....	77
4.2	<i>GmIFRs</i> were localized in the cytoplasm .....	78
4.3	No product formation for <i>GmIFR</i> candidates compared to 2'-hydroxydaidzein standard .....	79
5	Conclusions and future directions .....	81
References	.....	83
Curriculum Vitae	.....	96



## List of Tables

Table 2.1 List of primers used for full-length cloning, gene expression, and subcellular localization studies.....	26
Table 3.1 The list of identified IFRs in legume plants .....	45
Table 3.2 Characteristics of candidate <i>GmIFRs</i> .....	59

## List of Figures

Figure 1.1 The isoflavonoid biosynthetic pathway in soybean.....	13
Figure 1.2 Oxidation of 2'-hydroxydaidzein by IFR and NADPH cofactor .....	19
Figure 2.1 Gateway donor vector for cloning.....	31
Figure 2.2 Gateway destination vectors for protein expression and subcellular localization.	34
Figure 3.1 Phylogenetic analysis of <i>GmIFRs</i> .....	47
Figure 3.2 Identification of conserved motifs in <i>GmIFR</i> candidates.....	49
Figure 3.3 Tissue-specific expression profile of the <i>GmIFRs</i> gene family .....	53
Figure 3.4 Summary of all identification of <i>GmIFR</i> candidates process using <i>in silico</i> , phylogenetic, tissue-specific expression, and conserved motifs.....	57
Figure 3.5 Subcellular localization of the candidate <i>GmIFRs</i> .....	60
Figure 3.6 Protein expression of soluble fractions of <i>GmIFRs</i> with induction of different concentration of IPTG (0.5 mM and 1 mM).....	64
Figure 3.7 Confirmation of presence of His <sub>6</sub> tagged- <i>GmIFRs</i> and their molecular weight using western blot .....	66
Figure 3.8 Protein purification of <i>GmIFRs</i> using nickel resin and Econo column.....	68
Figure 3.9 Cleavage His <sub>6</sub> tag of <i>GmIFR1A</i> using TEV protease .....	70
Figure 3.10 HPLC chromatogram of candidate <i>GmIFRs</i> enzymatic reactions .....	73

## List of Abbreviations

\*Standard SI units not listed

<i>att</i>	attachment
<i>attB</i>	attachment site of the bacterial genome
<i>attP</i>	attachment site of the plasmid
Bp	base pair
ccdB	coupled cell division protein B
cDNA	complementary DNA
CDS	coding DNA sequences
CFP	cyan fluorescent protein
CHI	chalcone isomerase
CHR	chalcone reductase
CHS	chalcone synthase
4CL	4-coumarate-CoA-ligase
cv.	Cultivar
DNA	deoxyribonucleic acid
3,9DPO	3,9-dihydroxypterocarpan 6a-monooxygenase
ECL	enhanced chemiluminescence
ER	endoplasmic reticulum
G2DT	glycinol 2-dimethylallyltransferase

G4DT	glycinol 4-dimethylallyltransferase
<i>GmIFR</i>	soybean isoflavone reductase
GS	glyceollin synthase
2HID	2-hydroxyisoflavanone dehydratase
HPLC	high-performance liquid chromatography
HPV	human papillomavirus
HRP	horseradish peroxidase
IFR	isoflavone reductase
IFS	isoflavone synthesis
I2'H	isoflavone 2'-hydroxylase
IPTG	isopropylthio- $\beta$ -galactoside
IRL	isoflavone reductase- like
IMAC	immobilized metal affinity chromatography
JA	jasmonic acid
LB	lysogeny broth
LDL	low-density lipoproteins
NADPH	nicotinamide adenine dinucleotide phosphate
NO	nitric oxide
Ori	origin on vector
PAL	phenylalanine ammonia-lyase

PTS	pterocarpan synthase
PVDF	polyvinylidene fluoride
RNA	ribonucleic acid
RNA-Seq	ribonucleic acid-sequencing
ROS	reactive oxygen species
RPM	revolutions per minute
RPS	resistance to <i>Phytophthora sojae</i>
SA	salicylic acid
SDS	sodium dodecyl sulfate
SOC	super optimal broth with catabolite repression
siRNA	small interfering RNA
TEV	tobacco etch virus
TFA	trifluoroacetic acid
UTR	untranslated region
YFP	yellow fluorescent protein

## Chapter 1

### 1 Introduction

#### 1.1 Soybean

Soybean [*Glycine max* (L.) Merrill] is one of the most widely grown grain legumes worldwide (Rong et al., 2020). Global soybean production reached 268.58 million tons in 2012/2013, rising to 388.01 million tons by 2022/2023 (<https://www.statista.com/statistics/263926/soybean-production-in-selected-countries-since-1980/>). Soybean is the third-largest field crop in Canada after wheat and canola and plays a significant role in the country's economy. In 2023, Canada produced 6.54 million tons, with Ontario contributing 4 million tons (<https://soycanada.ca/>).

Soybean is a nutritional source with high-quality protein, beneficial unsaturated fatty acids, and isoflavones (Medic et al., 2014). Soybeans are rich in protein with approximately 36-40% protein content that makes them one of the highest plant-based protein sources available (Li et al., 2019). Soybean's proteins contain all the essential amino acids necessary for human and animal nutrition. Additionally, soybeans are a source of beneficial unsaturated fatty acids, including a balanced combination of omega-3 and omega-6 fatty acids, and various bioactive compounds such as saponins, peptides, and isoflavones (Kumar et al., 2010).

Soybeans are used for a variety of purposes, including human foods, livestock feed, and a diverse industrial application (Messina, 2010a; Messina, 2010b). The protein content of soybean makes them a crucial source in various dietary regimes, especially for

individuals seeking plant-based protein sources. The role of soybean in human food extends beyond its protein, as they are processed into a range of products like tofu, soy milk, and meat substitutes, catering to a variety of dietary preferences and needs (Messina and Messina, 2010). Furthermore, soybean plants serve as a fundamental resource in the production of biodiesel, biodegradable plastics, and numerous other eco-friendly products. This broad spectrum of applications underscores the importance of soybeans across food, agriculture, and industrial sectors (Ali, 2010; Candeia et al., 2009).

## **1.2 Origin and global journey of soybeans**

Soybean initially were domesticated in China approximately 5,000 years ago, primarily valued for their high-protein seeds. Over time, their utility expanded to include oil production (Hymowitz, 1970). By the 16<sup>th</sup> century, soybeans found their way to Japan through Chinese immigrants (Baraibar and Deutsch, 2023). During the 17<sup>th</sup> century, trade routes transported soybeans to Europe, and by the 18<sup>th</sup> century, soybeans reached North America, possibly introduced by European settlers, primarily utilized for animal feed and soil improvement (Du Bois, 2018). During 19<sup>th</sup> century, soybeans became more recognized in North America, setting the stage for research and breeding efforts (Du Bois et al., 2008). In the 20<sup>th</sup> century, the United States and Canada, particularly provinces like Ontario, further contributed to global distribution. In Canada, soybeans were introduced and cultivated in the early 20<sup>th</sup> century (Yoosefzadeh-Najafabadi and Rajcan, 2023). After World War II era, and the growing demand for soybeans and soy-based products, particularly for livestock feed and food processing, encouraged Canadian farmers to expand their soybean cultivation (Baraibar and Deutsch, 2023).

Presently, soybeans form an essential part of Canada's agricultural and agri-food industry, about 35% of the global harvested area, making substantial contributions to the nation's agricultural economy and trade. For decades the province of Ontario, more than half (54.4%) of Canada's soybean acreage, served as a major center for soybean production in Canada, but since the 1970s, production has begun to spread to other provinces including Manitoba, Quebec, Saskatchewan, Alberta, and the Maritimes (<https://soycanada.ca/industry/growing-areas/>).

### **1.3 Soybean genome**

The soybean genome was sequenced and published in 2010 through the Phytozome plant genomics database (Schmutz et al., 2010). The soybean genome from the *G. max* var. Williams 82 variety with a total of 1.1 gigabases has been assembled into 20 chromosomes that encodes 56,044 proteins and contains 88,647 transcripts. Soybean is paleopolyploid in nature, having undergone two whole duplication events approximately 59 and 13 million years ago. These events have resulted in a unique genetic landscape, with 75% of soybean genes having multiple copies (Schmutz et al., 2010). However, duplicated gene pairs have undergone significant diversification, resulting in the silencing and loss of around 25% of soybean gene duplicates, along with the generation of numerous new genes (Shoemaker et al., 2006). The soybean genome has significant genetic diversity, as there are many different varieties and cultivars with unique genetic characteristics, which can influence traits such as yield, disease resistance, nutritional content, etc (Choi et al., 2007; Hyten et al., 2010). Understanding this genetic diversity is crucial for scientists and breeding efforts to



identify genes and characterize their functions, thereby developing improved, resistant soybean varieties (Schmutz et al., 2010).

#### **1.4 *Phytophthora sojae* causes stem and root rot disease in soybean**

*Phytophthora sojae* is a significant and challenging pathogen in soybean fields (Bebber and Gurr, 2015). The soil-borne pathogen oomycete *P. sojae*, commonly known as a water mold pathogen, predominantly targets soybean plants. *P. sojae* was initially identified in the United States in the 1950s and has since spread globally (Thines, 2014). This pathogen is responsible for stem and root rot disease in soybean, resulting in an estimated annual soybean loss of approximately \$1-2 billion on a global scale (Tyler, 2007).

The life cycle of *P. sojae* involves sexual and asexual reproduction, contributing to its ability to adapt and cause disease in soybean crops. The sexual phase of this pathogen starts with the formation of oospores. In the presence of moisture, at temperatures ranging from 16°C to 25°C, and a pH range of approximately 6 to 7, oospores release motile zoospores that can swim through water films in the soil. *P. sojae* infects soybean plants primarily through the roots tissue (Tyler, 2007). Zoospores are attracted to soybean roots by chemical signals such as ethanol, glucosinolates, and flavonoids released by the plant. Zoospores encyst on the root surface and penetrate the plant cells. Then, the pathogen's mycelium grows within the plant, spreading and colonizing root tissues (Kasteel et al., 2023). It secretes enzymes that break down plant cell walls, allowing it to extract nutrients and water from the host plant. After infection, asexual sporangia are produced within infected roots, which then release more zoospores into the soil. These

zoospores can infect adjacent roots or be splashed by rain to nearby plants (Hardham, 2007).

*P. sojae* has become a major threat to soybean farming worldwide (Bebber and Gurr, 2015). It causes significant financial losses for farmers and is found in regions with high humidity and moisture. *P. sojae* infection is very challenging due to its ability to survive in the soil for more than 10 years (Wang et al., 2006b). This pathogen is also genetically diverse and can quickly adapt to new soybean varieties (Dorrance et al., 2003). All these reasons combine to make it a complex issue to manage in soybean cultivation.

## **1.5 Different plant's defense mechanisms to provide resistance**

Plants have a range of physical, hormonal, and metabolic responses to cope with environmental stresses. These responses help the plant adapt to adverse conditions and enhance its chances of survival (Kaur et al., 2022).

Physical responses in soybean against pathogens are part of the plant's defense mechanisms to prevent or limit pathogen invasion and infection. These physical responses include both pre-existing structural barriers and induced responses to prevent pathogen progress: 1) the first line of defense in soybean's physical response to pathogens is the reinforcement of cell walls. Soybean strengthens cell walls with structural components like lignin and suberin. These compounds make the cell walls more resistant to pathogen penetration (Thomas et al., 2007). 2) in the root and shoot system, the exodermis and endodermis are specialized layers that act as a barrier, making it difficult for pathogens to enter plant, and contribute to pathogen resistance (Enstone et al., 2002).

These two cell layers control osmotic pressure and ions, and are contain suberin, an aliphatic polyester of fatty acids, and phenolics (Geldner, 2013).

Upon sensing the pathogen's presence, the plant initiates an intricate network of hormonal signals that activate a range of defense responses (Pieterse et al., 2012). For example, salicylic acid (SA) is a key signaling molecule in the plant defense response against pathogens by producing proteins such as chitinases,  $\beta$ -1,3-glucanases (Durrant and Dong, 2004). Upon detecting the presence of *P. sojae*, soybean initiates hormonal responses, particularly elevating induction of jasmonic acid (JA) and SA (Sugano et al., 2013).

At the molecular level, the defense mainly consists of compounds such as specialized metabolites. Plant specialized metabolites are organic compounds produced by plants that are not directly involved in their growth, development, or reproduction but play vital roles in defense against stresses (Slusarenko et al., 2012). For example, glucosinolates in the *Brassicaceae* family are considered a molecular defense (Chen et al., 2020). In tobacco, in addition to glucosinolates, chitinases, and peroxidases, play a crucial role in preventing pathogen infection and limiting damage (Liu et al., 2003; Loon et al., 2006). Soybean include phytoalexin glyceollins that have antimicrobial properties and help restrict pathogen growth in pathogen infection (Bizuneh, 2021).

## **1.6 Strategies for managing *P. sojae* in soybean cultivation**

To effectively manage *P. sojae*, the pathogen responsible for stem and root rot disease in soybeans, a multifaceted approach is essential. Key elements of this approach include tillage and drainage practices and seed treatments (Dorrance and McClure, 2001;

Sugimoto et al., 2005). In fields with poor drainage, proper tillage enhances soil drainage and raises soil temperature, creating an environment less favorable for the pathogen's growth (Li et al., 2010). Seed treatments with fungicides such as metalaxyl offer a temporary shield, protecting germinating seeds and seedlings until they can develop their own resistance to the disease (Dorrance and McClure, 2001). Crop rotation with non-host plant-pathogen is considered as another agricultural practice to manage stem and root rot disease (Dorrance, 2018). However, all these strategies are restricted due to the pathogen's long-term survival in soil for more than ten years (Sugimoto et al., 2005). So, selecting a cultivar that displays tolerance or resistance to the strains of *P. sojae* found in the field has consistently proven to be the most effective method for preventing crop losses (Dale Young, 1999; Sugimoto et al., 2012).

## **1.7 Soybean resistance to *P. sojae*: complete and partial mechanisms**

Resistance to *P. sojae* in soybean can be two types including complete and partial resistance (Sahoo et al., 2017). Complete resistance in soybeans is associated with specific genes referred to as resistance to *P. sojae* (*RPS*) genes. These genes are specific and offer targeted protection against specific strain of *P. sojae* (Sahoo et al., 2017). *RPS* generally contribute to recognition of *P. sojae* effectors and activation of soybean defense. The first discovered resistance gene, *RPS1a*, was widespread in the United States during the 1960s (Gao et al., 2005). Following that, *RPS1c*, *RPS1k*, *RPS3a*, and *RPS6* were discovered in the subsequent years (Dorrance et al., 2003). To date more than 30 *RPS* genes have been identified such as *RPS1a*, *RPS1b*, *RPS1c* and *RPS1k*, *RPS3a*,

and *RPS6* (de Ronne et al., 2022; Dorrance, 2018). While all these genes have similar functions and participation in soybean defense, each specific *RPS* gene differently recognizes specific pathogen effectors (Dong et al., 2011). However, it's important to note that complete resistance is strain-specific, which means that *RPS* is an effective gene against one strain of pathogen and may or may not protect the plant from a different strain of the pathogen. So, the effectiveness of an *RPS* is limited by new resistant *P. sojae* strains (Yan and Nelson, 2019).

In contrast, partial resistance provides a broader but non-specific protection against only certain *P. sojae* strains. This resistance is more long-term because it does not rely on specific *RPS* genes. Partially resistant soybean plants exhibit a general ability to withstand various strains of *P. sojae* to some extent (Schmitthenner, 1985; Sugimoto et al., 2012). This form of resistance typically involves the activation of defense mechanisms that can decrease infection efficiency, lesion size and oospores production within the tissue. It also delays pathogen colonization, reducing the severity and damage caused by *P. sojae* infection (Dorrance, 2018; Mideros et al., 2007). Partial resistance is particularly valuable in situations where the pathogen population consists of diverse strains, as it provides a more generalized protection (Dorrance et al., 2003). Soybean cultivars with partial resistance are effective against all strains of *P. sojae* (Zenbayashi et al., 2002). Glyceollin phytoalexins are the major defensive compounds produced from the phenylpropanoid pathway in response to pathogen infections and cause partial resistance in soybean (Ng et al., 2011). Studies indicated that silencing of *chalcone reductase* gene in upstream of the isoflavonoid pathway leads to reduced

glyceollin production and the plants' ability resist pathogen attack (Graham et al., 2007; Subramanian et al., 2005).

## **1.8 Isoflavonoids**

### **1.8.1 Unlocking the power of isoflavonoids in soybean growth and defense**

Isoflavonoids serve as pivotal components in soybean's growth, root nodule formation, and defense mechanisms (Dakora and Phillips, 1996). Isoflavonoids influence seed germination and root development in soybean (Graham, 1991). They can affect the balance between the growth of primary and lateral roots, potentially influencing nutrient uptake and overall plant development (Peiretti et al., 2019).

Isoflavonoids, including daidzein, serve as precursors to produce phytoalexins. In soybean, production of phytoalexin glyceollin against *P. sojae* infection serves as a stress tolerance mechanism, ensuring their survival (Graham, 1990). Isoflavonoids play a key role in establishing beneficial symbiotic relationships between soybean plants and specific soil microorganisms. Isoflavonoids act as signaling molecules to attract nitrogen-fixing rhizobia to the plant's roots. These rhizobia form root nodules and convert atmospheric nitrogen into ammonia, a form of nitrogen that the soybean can readily assimilate, thereby enhancing soybean growth and nitrogen uptake (Biswas and Gresshoff, 2014; Rong et al., 2020). It was shown that the catalytic activity of isoflavone synthase and the production of isoflavone aglycones, such as daidzein and genistein,

contribute to the formation of nodules in the symbiotic interaction between *Bradyrhizobium japonicum* and soybeans (Subramanian et al., 2006).

### **1.8.2 Isoflavonoid is a nutritional source for human health**

In addition to polyunsaturated fats, fiber, vitamins, and minerals, soybeans are rich in specialized metabolites such as isoflavonoids and saponins (Saha and Mandal, 2019). Isoflavonoid compounds not only enhance the plant's resilience but also offer advantageous attributes for human health, including antimicrobial, anti-inflammatory, and antioxidant properties (Dakora and Phillips, 1996). They can help lower low-density lipoproteins (LDL) cholesterol levels, reducing the risk of heart disease. Also, studies showed genistein and daidzein inhibit inflammation and cell apoptosis (Abotaleb et al., 2019).

Furthermore, isoflavone aglycons, such as genistein, have antioxidant properties and can inhibit the growth of cancer cells (Taylor et al., 2009). The anticancer activity of isoflavones is further attributed to the regulation of steroid hormone synthesis (Zhang et al., 2017). For example, isoflavones can interact with estrogen receptors in the human body, potentially influencing hormone-related cancers. Research suggests that their binding to estrogen receptors might reduce the risk of hormone-driven cancers (Cederroth and Nef, 2009). As a result, the consumption of soybeans has been associated with a reduced risk of breast, prostate, and other hormone-related cancers (Tice, 2008). Isoflavones in soybean exhibit potent anti-cancer properties by suppressing the expression of tyrosine kinase, an enzyme pivotal in cancer cell growth, physiology, and division, which leads to the inhibition of apoptosis and regulates the cell cycle

(Basson et al., 2021). Notably, a combined treatment of daidzein and genistein has demonstrated effects in effectively inhibiting prostate cancer cell proliferation (Dong et al., 2013). In cases of cervical cancer resulting from human papillomavirus (HPV) infection, genistein's anticancer effect is mediated through apoptosis induction and a reduction in cell viability (Jaudan et al., 2018). In addition, genistein downregulates the expressions of Gli1 protein-related signaling pathways and the cell surface glycoprotein CD44. This inhibition of Gli1 results in control the attenuation of cancer stem cells properties, and could be used as an effective cancer therapy (Yu et al., 2014).

Soy consumption, particularly products like milk and tofu, can help regulate blood sugar levels (Ding et al., 2016). Isoflavonoids content in soy can slow the absorption of sugar, making it a suitable choice for individuals with diabetes I and II (Hu et al., 2021). Moreover, a recent study suggested that genistein could play a key role in diabetic alleviation. Based on this, genistein was able to induce autophagy in mouse renal podocytes, which has been confirmed to be essential for the progression of diabetic alleviation (Wang et al., 2018).

Soybeans are also a good source of dietary fiber, which aids in digestive health (Kim et al., 2021). Flavonoids and their byproducts can support a healthy gut by limiting harmful bacteria growth and promoting beneficial bacteria like Bifidobacterium and Lactobacillus. This helps maintain a balanced immune system, reduce endotoxin production, and improve nutrient absorption (Pei et al., 2020).

Isoflavonoids, with their antioxidant properties, help neutralize harmful free radicals, mitigating oxidative stress and inflammation (Iqbal et al., 2017). For example, genistein



and daidzein help protect against neurodegenerative diseases like Alzheimer's and Parkinson's by reducing nitric oxide (NO) and reactive oxygen species (ROS) production. This suggests a potential benefit of isoflavonoids in supporting brain health through neuroprotection and antioxidative effects (Hussain et al., 2018). Moreover, Ma et al. (2010) suggested genistein could reduce the oxidative stress as a neuroprotective antioxidant and maintain redox balance in cells. Also, another study investigated daidzein acts as a neuroprotectant by anti-inflammatory mechanisms in an in vitro model of microglial activation (Chinta et al., 2013).

### **1.8.3 Isoflavonoid biosynthetic pathway**

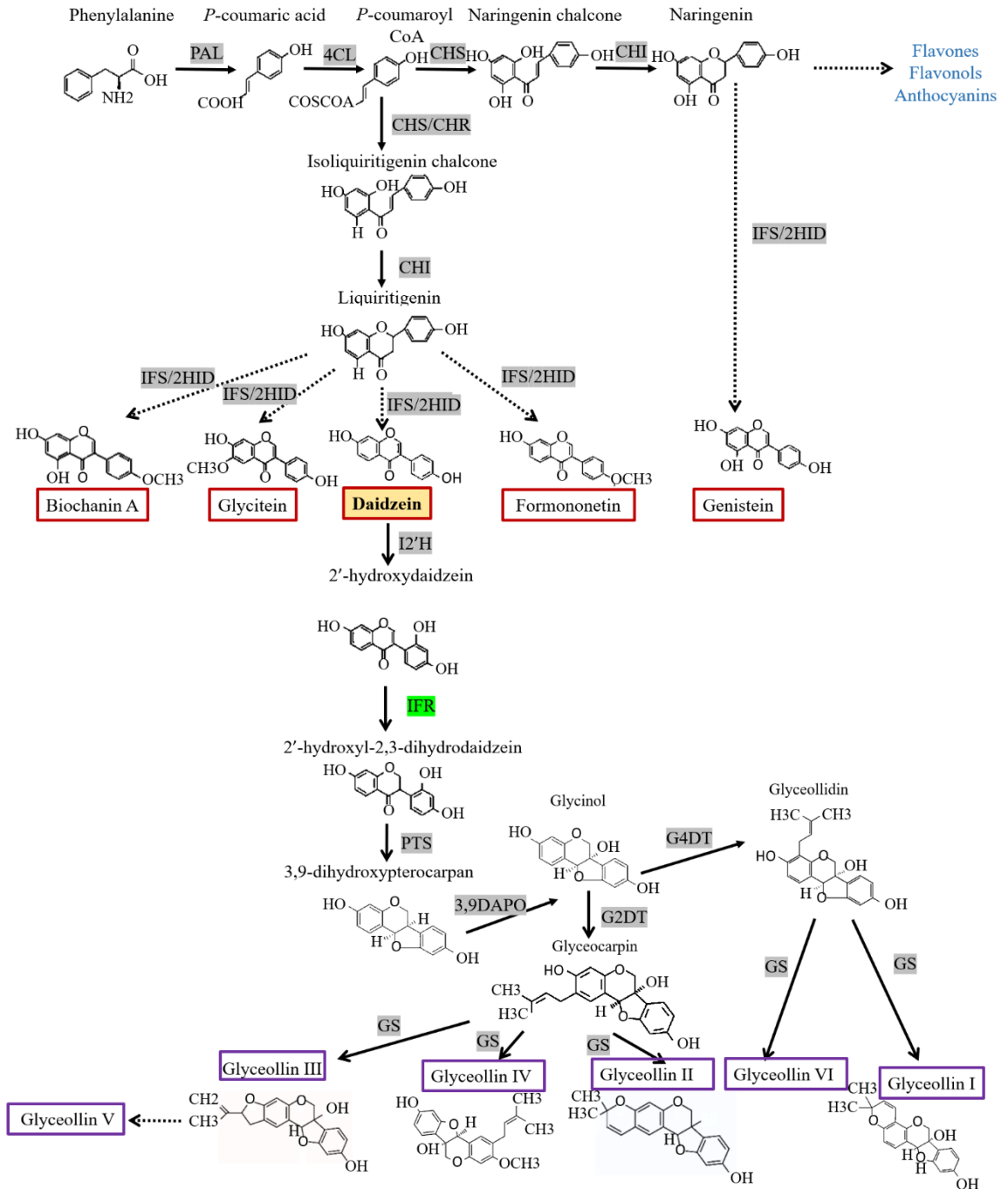
Isoflavonoids are specialized metabolites in plants that are derived from a legume-specific branch of the general phenylpropanoid pathway (Veitch, 2007). While isoflavonoid compounds were primarily associated to be unique to leguminous plants, their presence within certain non-legume plant species such as tobacco, *Arabidopsis thaliana*, and maize was discovered (Broun, 2005; Yu et al., 2000). Other branches of the phenylpropanoid pathway are present in all plant species, leading to the production of compounds like lignin, proanthocyanidins, anthocyanin, and phlobaphenes (Dhaubhadel, 2011).

The initial step in the synthesis of flavonoids and isoflavonoids, known as chalcone production, relies on an enzymatic reaction catalyzed by chalcone synthase (CHS), a specific polyketide synthase found in plants (**Figure 1.1**). Legumes produce two types of chalcones, tetrahydroxy chalcone (naringenin chalcone) and trihydroxy chalcone (isoliquiritigenin chalcone), while non-legume plants only produce tetrahydroxy

### Figure 1.1 The isoflavonoid biosynthetic pathway in soybean

The pathway starts with the phenylalanine which is converted into many metabolites including flavones, flavanols, and anthocyanins (in blue). The combination activity of IFS/2HID produces five isoflavone aglycones including biochanin A, genistein, glycitein, daidzein, and formononetin (highlighted in red boxes). The reduction of daidzein leads to start glyceollin pathway. Then 2'-hydroxydaidzein converts to 2'-hydroxyl-2,3-dihydrodaidzein by isoflavone reductase (highlighted in green). Subsequently, further enzymatic reactions such as PTS, 3,9-DAPO, G2DT, G4DT, and GS result in accumulation six antimicrobial phytoalexins glyceollin including I, II, III, IV, V, and VI (highlighted in purple boxes).

CHI (chalcone isomerase); CHR (chalcone reductase); CHS (chalcone synthase); 4CL (4-coumarate-CoA-ligase); 3,9-DAPO (3,9-dihydroxypterocarpan 6a-monooxygenase); G2DT (glycinol 2-Dimethylallyltransferase); G4DT (glycinol 4- dimethylallyltransferase); GS (glyceollin synthase); 2HID (2-hydroxyisoflavanone dehydratase); IFR (isoflavone reductase); IFS (isoflavanone synthase); I2'H (isoflavone 2'-hydroxylase); PAL (phenylalanine ammonia-lyase); PTS (pterocarpen synthase).



chalcones. In addition to CHS, legumes employ a specific enzyme, chalcone reductase (CHR) to convert isoliquiritigenin chalcone into liquiritigenin. Both naringenin and isoliquiritigenin chalcones are further transformed into their respective flavanones (liquiritigenin) by chalcone isomerase (CHI). From here, the combination activity of isoflavone synthesis (IFS) and 2-hydroxyisoflavone reductase (2HID) produces the five isoflavone aglycones: daidzein, genistein, glycitein, formononetin, and biochanin A. The biosynthesis of isoflavonoid is constitutive and occurs during different growth and development stages of soybean. The isoflavonoid biosynthesis pathway has multiple gene family members with different amino acid sequences and tissue-specific expression patterns during growth, development stages, or in response to stress (Dhaubhadel et al., 2003). For example, the soybean *CHS* (Anguraj Vadivel et al., 2018; Dhaubhadel et al., 2006), *CHR* (Sepiol et al., 2017), *CHI* (Dastmalchi and Dhaubhadel, 2015), and *IFS* (Jung et al., 2000) gene families have 14, 9, 12, and 2 members respectively (**Figure 1.1**).

## **1.9 Glyceollins as a soybean defense mechanism in response to stresses**

Phytoalexins are antimicrobial compounds produced by plants in response to pathogenic attacks. Phytoalexins vary among different legume species. For example, glyceollin in soybeans (Cheng et al., 2015), medicarpin in alfalfa and chickpea (Schlieper et al., 1990; Wang et al., 2006a), pisatin in pea (Sun et al., 1991), and phaseollin in common bean (Rípodas et al., 2013) are produced as antimicrobial phytoalexins in response to stress. The glyceollin pathway is a defense mechanism in response to stresses in soybean that

produces and accumulates glyceollins (I, II, III, IV, V, and VI isomers) in infected tissues (Simons et al., 2011). Upon stresses, specifically *P. sojae* infection, isoflavone aglycone daidzein serves as a substrate to start enzymatic processes that result in the production and accumulation of glyceollin in soybean (Sukumaran et al., 2018). The first step of glyceollin pathway is the hydroxylation of daidzein by isoflavone 2'-hydroxylase to 2'-hydroxydaidzein (Akashi et al., 1998). Following this, isoflavone reductase (IFR) catalyses 2'-hydroxydaidzein to 2'-hydroxy-2,3-dihydrodaidzein and subsequent cyclization, a step catalyzed by pterocarpan synthase (PTS) (Fischer et al., 1990). The resulting compound is then transformed into glycinol with the involvement of 3,9-dihydroxypterocarpan 6a-monooxygenase (3,9DPO) (Schopfer et al., 1998). Then, the enzyme glycinol 4-dimethylallyltransferase (G4DT) mediates prenylation, resulting in the formation of 4-glyceollidin. Conversely, glycinol 2-dimethylallyltransferase (G2DT) performs prenylation on the C-2 position of glycinol, leading to formation of glyceocarpin. The last step in the glyceollin pathway is regulated by glyceollin synthase (GS), which catalyzes the cyclization of 4-glyceollidin to produce six isomerases of glyceollin (Akashi et al., 2008; Welle and Grisebach, 1988).

## **1.10 Oxidoreductases family**

Oxidoreductases are a diverse group of enzymes found in microbes, plants, and animals (Younus, 2019). These enzymes play a crucial role in catalyzing electron exchange between donor and acceptor molecules, facilitating processes such as electron transfer, proton or hydrogen extraction, hydride transfer, oxygen insertion, and other essential reactions (Husain, 2017; Toone, 2010). Oxidoreductases include various types: oxidases,

dehydrogenases, hydroxylases, oxygenases, peroxidases, and reductases (Özgen et al., 2019).

Oxidoreductases are essential in both aerobic and anaerobic metabolism and can work with a broad range of substances, including organic compounds like alcohols, amines, ketones, and inorganic substances such as small anions and metals (Özgen et al., 2019). Among oxidoreductases, NAD(P)H-dependent dehydrogenases are particularly important for asymmetric or chiral synthesis. These enzymes can reversibly convert prochiral aldehydes and ketones into chiral alcohols, hydroxy acids, and amino acids. Hydroxy acid dehydrogenases can synthesize chiral hydroxy acids from various 2-oxo acids, including aliphatic, linear, branched, or aromatic types, which are valuable building blocks (Santaniello et al., 1992).

### **1.10.1 Structure and function of isoflavone reductase (IFR)**

IFRs are classified as a subfamily of the oxidoreductase family because they catalyze reactions involving the reduction of isoflavones (Huang et al., 2014). These enzymes are relied on NADPH for their functions, particularly in the synthesis of phytoalexins in leguminous plants (Fischer et al., 1990; Graham et al., 1990). Structurally, IFRs contain three conserved motifs, the NADPH-binding motif at the N-terminus and two substrate-binding motifs at the C-terminus (Bhinija et al., 2022; Wang et al., 2006a). IFRs are unique to the plant kingdom and play vital roles in how plants react to different biotic or abiotic stresses (Bhinija et al., 2022; Gang et al., 1999; Guo et al., 1994; Paiva et al., 1991). IFR was initially recognized as a critical enzyme in the final stages of the medicarpin biosynthetic pathway (Daniel and Barz, 1990). In alfalfa, IFR plays a pivotal

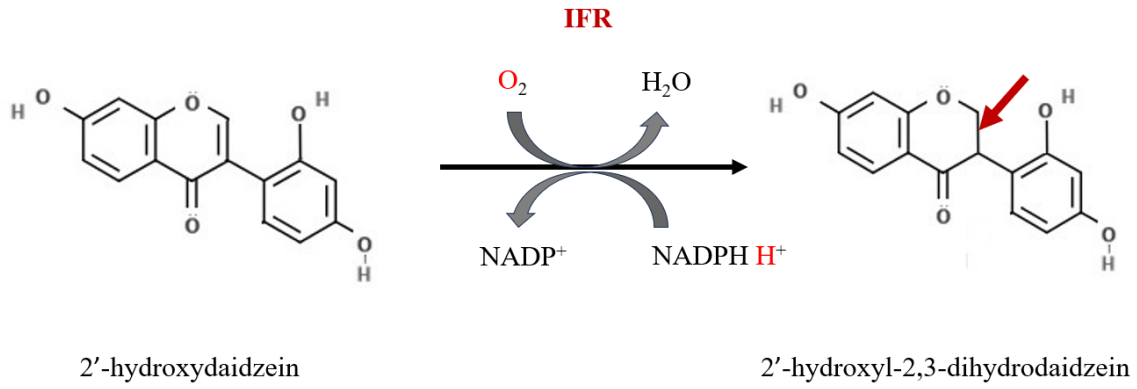
role in transforming 2'-hydroxyformononetin into vestitone, which acts as a precursor to produce medicarpin phytoalexin (Guo et al., 1994; Paiva et al., 1991). IFRs have also been discovered and their complementary DNAs (cDNAs) cloned from various legume species, including peas (*Pisum sativum*; *PsIFR*; Paiva et al. (1991)), chickpeas (*Cicer arietinum*; *CaIFR*; Rípodas et al. (2013)), common bean (*Phaseolus vulgaris*; *PvIFR*; Sun et al. (1991)), and soybeans (*Glycine max*; *GmIFR*; Cheng et al. (2015)) with a specialized role in the biosynthesis of isoflavonoid phytoalexins. In soybeans, IFR can utilize 2'-hydroxy forms of isoflavone aglycons, including 2'-hydroxydaidzein, 2'-hydroxyformononetin, and 2'-hydroxygenistein, as substrates. In peas and chickpeas, IFRs have shown activity with 7,2' dihydroxy-4',5'-methylenedioxyisoflavone and 2'-hydroxyformononetin, respectively.

In soybeans, a single monomeric cytosolic IFR enzyme has been identified and characterized. This enzyme facilitates the conversion of 2'-hydroxyformononetin to an unknown product. It was observed that the transcript level of *GmIFR* in soybeans increases following infection by *P. sojae* and in response to wounding stress (Cheng et al., 2015). Moreover, soybean's cytosolic IFR can convert 2'-hydroxydaidzein into 2'-hydroxydihydrodaidzein when NADPH is present (Fischer et al., 1990) (**Figure 1.2**). In addition to legumes, IFRs were called IFR-like (IRL) proteins, were isolated from various non legume plants including tobacco (*Nicotiana tabacum*; (Shoji et al., 2002), rice (*Oryza sativa*; (Kim et al., 2003), *A. thaliana* (Babiychuk et al., 1995), Ginkgo (*Ginkgo biloba*; (Hua et al., 2013), etc. IRLs have sequence similarity to legume IFRs and have been implicated in response to biotic or abiotic stresses (Babiychuk et al., 1995; Kim et al., 2003; Shoji et al., 2002). For instance, *OsIRL* in rice acts as an

**Figure 1.2 Oxidation of 2'-hydroxydaidzein by IFR and NADPH cofactor**

IFR oxidizes substrate by transferring hydrogen to acceptor molecules, NADP, and reduces one oxygen of 2'-hydroxydaidzein to produce 2,3-dihydrodaidzein with a hydroxyl group at position 2' (Fischer et al., 1990). Red arrow shows a reduced double bond at the 2,3 positions.





antioxidant in reaction to ROS in both suspension-cultured cells and during root development. It's also stimulated after *Pyricularia grisea* inoculation (Kim et al., 2010a; Kim et al., 2010b). Additionally, two tobacco genes, *TP7* and *A622*, which code for IFR-like proteins, play roles in defense mechanisms like lignan biosynthesis and in the metabolism of nicotine or related alkaloids (Shoji et al., 2002).

## 1.11 Hypothesis and objective

Given soybean's history of whole genome duplication and the presence of multigene families (such as *CHS*, *CHR*, *CHI*, *IFS*, and *PT*) in the isoflavonoid biosynthetic pathway, I hypothesize that soybean has multiple *IFR* genes and that some of these play a key role in response to *P. sojae* in root tissue.

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

To fulfill this objective, I will look to

- a) Identify members of the *GmIFR* gene family in soybean.
- b) Determine the subcellular localization of all *GmIFR* candidates.
- c) Express and purify *GmIFR* candidates.
- d) Characterize functional *GmIFR* enzyme activity.

## Chapter 2

### 2 Materials and Methods

#### 2.1 Plant materials and growth conditions

*Nicotiana benthamiana* (tobacco) seeds were scattered on wet PRO MIX® BX MYCORRHIZAETM soil (Rivière du Loup, Canada) in a small tray and incubated in a controlled environment with a light cycle of 16 hours at 25°C followed by an 8-hour darkness at 20°C. The humidity level in the growth chamber was maintained at 50%, and a constant light intensity of 70  $\mu\text{mol photons m}^{-2}\text{s}^{-1}$ . The tray was initially covered with a clear plastic dome for two days creating a high-humidity environment to promote successful seed germination. After a week, individual seedlings were transferred into sterilized pots filled with soil and watered once a week using a nutrient mixture of nitrogen, phosphorous, and potassium (20-20-20).

#### 2.2 Bacterial strains and growth conditions

*Escherichia coli* DH5 $\alpha$  and TOP10 (Invitrogen) were used for cloning to maintain and propagate entry and/or destination vectors. *Agrobacterium tumefaciens* GV3101 was used for transient expression in tobacco leaves for subcellular localization. *E. coli* Rosetta-gami (DE3) was used for protein expression.

Strains of *E. coli* including Rosetta-gami (DE3) and *A. tumefaciens* strain GV3101 were grown in lysogeny broth (LB) (Lennox, 1955) or super optimal broth with catabolite

repression (SOC) media supplemented with appropriate antibiotics at 37°C and 28°C, respectively, in a shaker incubator at 225 RPM. Finally, the stocks of bacteria transformed with desired plasmids were prepared in 50% glycerol, and stored at –80°C.

### 2.3 Identification of IFR candidates

To identify all the putative soybean *IFRs*, keyword searches using the words ‘isoflavone reductase’ and ‘oxidoreductase’ were conducted in the annotated *G. max* Wm82.a4. v1 genome in the Phytozome 13 database (<https://phytozome-next.jgi.doe.gov/>). In addition, Glyma.01G172600.1, the first identified IFR in *G. max* (Cheng et al., 2015), and CAA41106.1, IFR in *M. sativa* (Wang et al., 2006a), were used as queries for a protein BLAST (BLASTp) search. To ensure the comprehensiveness of the keyword search, each *GmIFR* was employed as a query in additional BLASTp (<https://phytozome-next.jgi.doe.gov/>).

To create phylogenetic tree, protein sequence of all putative *GmIFR* candidates were aligned with IFR amino acid sequences from other legume plants using ClustalW. Then, phylogenetic tree was built using maximum likelihood method and 1000 bootstrap replicates in MEGAX (Kumar et al., 2018).

To determine whether candidate *GmIFRs* contain conserved sequence motifs, all *GmIFRs* were screened for the splice variant (<https://phytozome-next.jgi.doe.gov/>). Then the amino acid sequence of motifs was identified within the protein sequences of each candidate. This identification was carried out through both manual analysis and the utilization of the MEME tool (<https://meme-suite.org/meme/tools/meme>)

The publicly available ribonucleic acid-sequencing (RNA-Seq) data was used to identify the *GmIFR* candidates with transcript expression in root tissue on soybase (<https://www.soybase.org/soyseq/>), and bioproject IDs including PRJNA324419 (Li et al., 2016), PRJNA544432 (Jahan et al., 2020), PRJNA574764, and PRJNA318321 (Jing et al., 2016). These datasets contained samples taken at different times after soybean plants were infected with *P. sojae*. The samples were from soybean varieties that either could resist the infection or were susceptible to it. One of the datasets, PRJNA574764, was particularly about partial resistance. All four bioprojects in this analysis were included to identify the *GmIFR* candidates with transcript expression in root tissue involved in resistance against *P. sojae*. To normalize transcript levels, reads per kilobase of transcript per million reads mapped (RPKM) were used.

*In silico* analysis was performed to predict the subcellular localization and molecular weight of *GmIFR* candidates. The WoLF PSORT was used to determine where *GmIFRs* are predicted to localize within cell (Horton, 2006). The molecular weight of *GmIFRs* was calculated using ExPASy (Wilkins et al., 1999).

## **2.4 Cloning and transformation**

Gateway cloning strategy-based recombination system (Thermo Fisher Scientific, USA) was used for cloning all *GmIFR* candidates to analyze functional characterization including subcellular localization and protein expression studies.

### 2.4.1 Cloning into the Gateway entry vector

In the Gateway-compatible system, all the primers for cloning of *GmIFR* candidates were designed with the attachment site of the bacterial genome (*attB1*) adaptor sequence (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT-3') in the forward primers, and the *attB2* adaptor sequence (5'-GGGGACCACTTTGTACAAGAAAGCTGGGT-3') in the reverse primers. Gene-specific primers (**Table 2.1**) were used in polymerase chain reactions (PCR) using cDNA synthesized from ribonucleic acid (RNA) isolated from soybean cv. Williams 82 hypocotyl infected with *P. sojae*. The synthesis of cDNA

**Table 2.1 List of primers used for full-length cloning, gene expression, and subcellular localization studies**

Gene (given name)	Primer Name	Sequence (5' to 3')	Amplicon Size (bp)
<b>Protein Expression</b>			
<i>GmIFR1A</i>	<i>GmIFR1AF</i>	GGGACAAAGTTTGTACAAAAAAGCAGGCTTCATGGCTCCGAAA GATAGAAATCC	957
	<i>GmIFR1AR</i>	GGGACCACTTTGTACAAAGAAAGCTGGGTCTCAGACAAACTGA TCCAAATATTCA	
<i>GmIFR1A</i>	<i>GmIFR11AF</i>	GGGACAAAGTTTGTACAAAAAAGCAGGCTTCATGTTGGACTG CCATTCAAAATGAT	1131
	<i>GmIFR11AR</i>	GGGACCACTTTGTACAAAGAAAGCTGGGTCTCAGACAAACGCA TTCAAATAGTTGTCG	
<i>GmIFR1B</i>	<i>GmIFR1BF</i>	GGGACAAAGTTTGTACAAAAAAGCAGGCTTCATGGCAGCGAA GAGCAAAATCCTAG	924
	<i>GmIFR1BR</i>	GGGACCACTTTGTACAAAGAAAGCTGGGTCTCAGACAAACGCA TTCAAATAGTTGTCG	
<i>GmIFR1D</i>	<i>GmIFR1DF</i>	GGGACAAAGTTTGTACAAAAAAGCAGGCTTCATGTGGAATTAT CAGTTTTTAGGCA	1155
	<i>GmIFR1DR</i>	GGGACCACTTTGTACAAAGAAAGCTGGGTCTCAGACAAACGCA TTCAAATAGTTGTCG	

Table 2.1 Continue

Gene (given name)	Primer Name	Sequence (5' to 3')	Protein Expression	Amplicon Size (bp)
<i>GmIFR11D</i>	<i>GmIFR11DF</i>	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCTGGGAAAGAT AGAATCCTG		957
	<i>GmIFR11DR</i>	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGACAAACTGATTC AAATATTCATCG		
<i>GmIFR1E</i>	<i>GmIFR1EF</i>	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCTGGGAAAGAT AGAATCC		957
	<i>GmIFR1ER</i>	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGACAAACTGATTC AAATATTCATC		
<i>GmIFR1IE</i>	<i>GmIFR1IEF</i>	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCTGGGAAAGAT AGAATCC		957
	<i>GmIFR1IER</i>	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGACAAACTGATTC AAATATTC		



Table 2.1 Continue

Gene (given name)	Primer Name	Sequence (5' to 3')	Amplicon Size (bp)
<b>Subcellular Localization</b>			
<i>GmIFR1A</i>	<i>GmIFR1AF</i>	GGGGACAAAGTTTGTACAAAAAGCAGGCTTTCATGGCTCCGAAA GATAGAAATCC	954
	<i>GmIFR1AR</i>	GGGGACCACTTTGTACAAGAAAGCTGGGTGACAAACTGATCCA AATATTCA	
<i>GmIFR11A</i>	<i>GmIFR11AF</i>	GGGGACAAAGTTTGTACAAAAAGCAGGCTTTCATGTTGGGACTGC CATTCAAAATGAT	1128
	<i>GmIFR11AR</i>	GGGGACCACTTTTGTACAAGAAAGCTGGGTGACAAACGCATTCA AATAGTTGTCG	
<i>GmIFR1B</i>	<i>GmIFR1BF</i>	GGGGACAAAGTTTGTACAAAAAGCAGGCTTTCATGGCAGCGAAG AGCAAAATCCTAG	921
	<i>GmIFR1BR</i>	GGGGACCACTTTTGTACAAGAAAGCTGGGTC GACAAACGCATTCAAATAGTTGTCG	
<i>GmIFR1D</i>	<i>GmIFR1DF</i>	GGGGACAAAGTTTGTACAAAAAGCAGGCTTTCATGTGGAATTATC AGTTTTTAGGCA	1152
	<i>GmIFR1DR</i>	GGGGACCACTTTTGTACAAGAAAGCTGGGTGACAAACGCATTCA AATAGTTGTCG	

Table 2.1 Continue

Gene (given name)	Primer Name	Sequence (5' to 3')	Amplicon Size (bp)
<b>Subcellular Localization</b>			
<i>GmIFR11D</i>	<i>GmIFR11DF</i>	GGGGACAAAGTTTGTACAAAAAAGCAGGCTTCATGGCTGGGAAAGA TAGAATCCTG	954
	<i>GmIFR11DR</i>	GGGACCACCTTTGTACAAGAAAGCTGGGTCGACAAAAGCTGATTCAA ATATTCATCG	
<i>GmIFR11E</i>	<i>GmIFR11EF</i>	GGGGACAAAGTTTGTACAAAAAAGCAGGCTTCATGGCTGGGAAAGA TAGAATCC	954
	<i>GmIFR11ER</i>	GGGACCACCTTTGTACAAGAAAGCTGGGTCGACAAAAGCTGATTCAA ATATTCATC	
<i>GmIFR11E</i>	<i>GmIFR11EF</i>	GGGGACAAAGTTTGTACAAAAAAGCAGGCTTCATGGCTGGGAAAGA TAGAATCC	954
	<i>GmIFR11ER</i>	GGGACCACCTTTGTACAAGAAAGCTGGGTCGACAAAAGCTGATTCAA ATATTC	

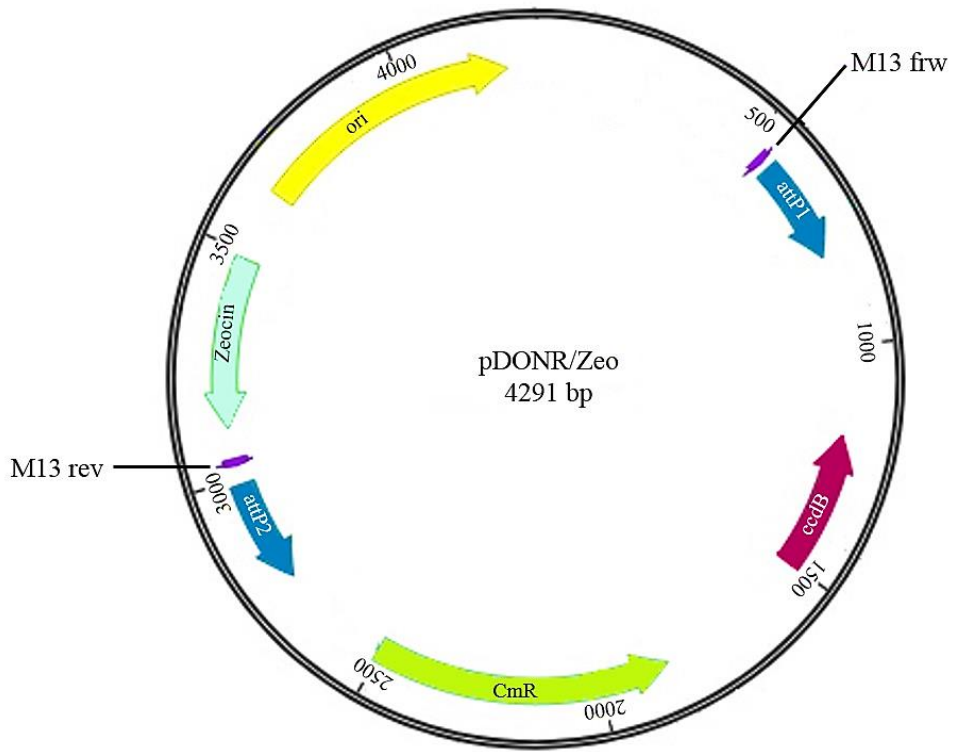
from 1  $\mu\text{g}$  of total RNA and 1  $\mu\text{L}$  of reverse transcriptase (200 U/ $\mu\text{L}$ ) was achieved using the thermoscript reverse transcription-polymerase chain reaction (RT-PCR) System by Life Technologies in a 20  $\mu\text{L}$  reaction. In the subsequent RT-PCR reactions, 2  $\mu\text{L}$  of the undiluted RT reaction served as a template. PCR products were loaded on a 1% agarose gel, and RedSafe stain (iNtROn Biotechnology) was applied for visualization. Gel images were captured using a Bio-Rad Gel Doc system. Following gel electrophoresis, PCR products were purified via gel extraction using the EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc.). PCR products, amplified with Gateway primers were recombined into pDONR/Zeo (Invitrogen) using Gateway BP Clonase® II Enzyme Mix (Invitrogen) (**Figure 2.1**). The recombinant plasmids were then transformed into *E. coli* DH5 $\alpha$  and plated on LB agar supplemented containing zeocin antibiotic (50  $\mu\text{g}/\text{mL}$ ). Positive transformants were screened by colony PCR using gene-specific Gateway primers, selected, and cultured overnight at 37°C in LB medium containing zeocin. Plasmid DNA was then extracted using EZ-10 Spin Column Plasmid DNA Kit (Bio Basic Inc.). Extracted plasmid DNA was quantified using NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific). Finally, the insertion of the correct DNA was confirmed by DNA sequencing, carried out at Eurofins Genomics LLC, USA. For sequence analysis, the DNASTAR® Lasergene software, MEGAX, and ClustalW were used.

#### **2.4.2 Cloning into the Gateway destination vector**

For subcellular localization and protein expression, the pDONRZeo-*GmIFR* were recombined into pEarleyGate101 (pEG101) (Earley et al., 2006) and pET160 DEST

**Figure 2.1 Gateway donor vector for cloning**

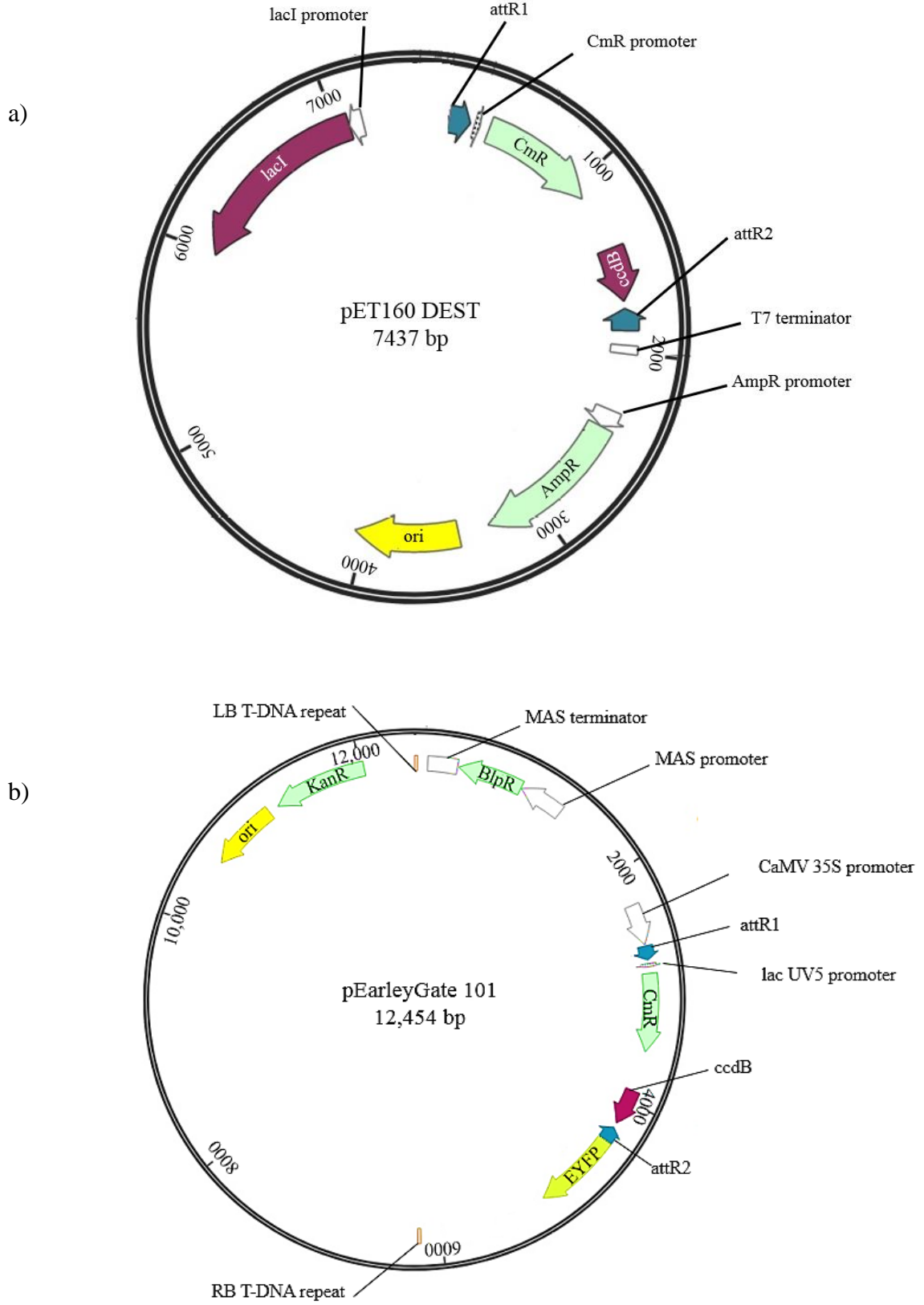
The pDONRzeo vector map with coupled cell division protein B (*ccdB*), *attP*: attachment site on the plasmid genome (*attP1*) and *attP2* sites and a Zeocin™ resistance marker. *cmR*: chloramphenicol resistance, *ccdB*: coupled cell division protein B, *attB*: attachment site on the bacterial genome, *attP*: attachment site on the Plasmid, Ori: origin, bp: base pair.



(Francis and Page, 2010) respectively using LR clonase (Invitrogen). pEarleyGate101 was used to produce translational fusions with C-terminal yellow fluorescence fusion protein (*GmIFR*- YFP) for subcellular localization (**Figure 2.2a**). pET160DEST was used to add a His<sub>6</sub>-tag and a tobacco etch virus (TEV) cleavage site to the N-terminus of the candidate proteins for expression and purification (**Figure 2.2b**). The recombinant plasmids were propagated and maintained in *E. coli* DH5 $\alpha$  for subcellular localization and TOP10 for protein expression. Transformed *E. coli* colonies containing pEG101-*GmIFR*s and pET160 DEST-*GmIFR* plasmids were screened by colony PCR using gene-specific primers (**Table 2.1**). LB media with kanamycin (50  $\mu$ g/mL) and ampicillin (100  $\mu$ g/mL) was used separately for growing positive colonies containing pEG101-*GmIFR*s, and pET160DEST-*GmIFR*, respectively, at 37°C for overnight. Then, plasmids were extracted and purified using the EZ-10 Spin Column Plasmid DNA Kit (Bio Basic Canada Inc). pEG101-*GmIFR*s plasmids transformed into *A. tumefaciens* GV3101 electro-competent cells, and pET160DEST-*GmIFR* into *E. coli* Rosetta-gami (DE3) electro-competent cells using electroporation (Traore and Zhao, 2011). Electroporation was carried out in a Gene Pulser® Cuvette (BioRad Laboratories) with a 0.1 cm electrode gap using MicroPulser™ (BioRad Laboratories) for the transformation of recombinant plasmids into bacterial strains. The voltage of 1.80 kV and 2.18 kV were used for *E. coli* strains and *Agrobacterium* sp., respectively. The transformed *A. tumefaciens* were grown on LB agar supplemented with rifampicin (10  $\mu$ g/mL), gentamycin (50  $\mu$ g/mL), and kanamycin (50  $\mu$ g/mL) to select positive colonies contain pEG101-*GmIFR*s. The transformed *E. coli* Rosetta-gami (DE3) carrying

**Figure 2.2 Gateway destination vectors for protein expression and subcellular localization**

a) The pET160 DEST for protein expression b) The pEarleyGate 101 vector map for subcellular localization. Vector maps were adapted from <https://www.snapgene.com/>. EYFP: enhanced yellow fluorescent protein, cmR: chloramphenicol resistance, AmpR: ampicillin resistance, KanR: kanamycin resistance, lacI: lactose operon repressor, ccdB: coupled cell division protein B, attR: attachment site on the right, Ori: origin, bp: base pair. MAS: multiple cloning site, LB T-DNA: left border sequences flanked the transfer DNA, RB T-DNA: right border sequences flanked the transfer DNA, CaMV: Cauliflower Mosaic Virus.





pET160DEST *GmIFR* was grown on LB agar supplemented with ampicillin (100 µg/mL), tetracycline (10 µg/mL), kanamycin (50 µg/mL), and chloramphenicol (34 µg/mL) at 37°C overnight to select positive colonies contain pET160DEST-*GmIFR*. Finally, bacterial transformants were screened by colony PCR using gene-specific primers (Table 2.1).

## 2.5 Subcellular localization

### 2.5.1 Transient expression of proteins in *N. benthamiana* leaves

To determine the subcellular localization of *GmIFRs*, the infiltration of *N. benthamiana* leaves with *A. tumefaciens* GV3101 carrying pEG101-*GmIFR*, was employed as outlined by Sparkes et al., 2006. Initially, a single colony of *A. tumefaciens* GV3101 grew in an infiltration medium at 28°C. This medium consisted of LB containing 10 mM 2-N-morpholino-ethanesulfonic acid (MES) pH of 5.6, supplemented with kanamycin (50 µg/mL), rifampicin (10 µg/mL), and gentamycin (50 µg/mL) for selecting *A. tumefaciens* GV3101 carrying pEG101-*GmIFR*. The culture was then incubated at 28°C until the optical density at 600 nm (OD<sub>600</sub>) reached a range of 0.5-0.8.

Bacterial cells were pelleted by centrifuging at 3000 RPM for 30 min in a microcentrifuge and then re-suspended in Gamborg's medium (3.2 g/L Gamborg's B5 and vitamins, 20 g/L sucrose, 10 mM MES at pH 5.6, and 200 µM acetosyringone) until the final OD<sub>600</sub> reached 1.0. To activate the virulence genes, the *Agrobacteria* harboring the plasmids of interest were incubated at room temperature for 1 hour with gentle agitation. The leaves of 4–6-week-old *N. benthamiana* plants were infiltrated with a

bacterial culture. This was achieved by gently placing the syringe barrel against the underside of the leaf and applying moderate pressure (Sparkes et al., 2006).

To confirm the subcellular localization, the *Agrobacterium* culture containing pEG101-*GmIFR* was co-infiltrated (1:1 ratio) with an endoplasmic reticulum (ER) organelle marker that was translationally fused to cyan fluorescent protein (CFP).

Following the infiltration, the plants were returned to the growth room at normal growth condition as described in Section 2.1.1 for 48 hours.

### **2.5.2 Confocal microscopy**

Following 48 hour incubation, the translational fusion of each *GmIFR* candidates with YFP in the epidermal cell layers of the infiltrated *N. benthamiana* leaves were visualized using an Olympus FV1000 confocal microscope (Olympus Corporation, Japan) with a 63× water immersion objective lens at 514 nm excitation wavelength (Keller, 1995; Sparkes et al., 2006). 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 used.

## **2.6 Protein expression**

### **2.6.1 Bacterial cultivation and protein expression**

Protein expression was started by preparing liquid LB media with appropriate antibiotics including ampicillin (100  $\mu\text{g}/\text{mL}$ ), tetracycline (10  $\mu\text{g}/\text{mL}$ ), kanamycin (50  $\mu\text{g}/\text{mL}$ ), and chloramphenicol (34  $\mu\text{g}/\text{mL}$ ) to select positive *E. coli* Rosetta-gami (DE3) colonies carrying pET160 DEST-*GmIFRs*. First, each transformed *E. coli* Rosetta-gami (DE3) carrying cytoplasm-localized pET160 DEST-*GmIFRs* was grown in a small initial culture (10 mL) at 37°C for overnight. These overnight cultures were transferred and grown in 300 mL liquid LB media with the antibiotics as described above at 37°C for 2-3 hours to reach an OD<sub>600</sub> of 0.8. Then, the culture was equally divided into three 100 mL into 500 ml flasks. The protein expression of three cultures was induced by different concentrations of isopropylthio- $\beta$ -galactoside (IPTG) including 0 mM (as control), 0.5 mM, and 1 mM at 30°C for 16 hours with 225 RPM shaking.

### **2.6.2 Protein extraction and confirmation on SDS polyacrylamide gel**

After 16 hours of protein expression, cultures for induced and un-induced protein samples were centrifuged at 3000 RPM for 15 min to get pellets. Pellets were stored at 20°C for at least 2 hours, resuspended, and extracted using lysis buffer (50 mM potassium phosphate, pH 7.4, 400 mM NaCl, 100 mM KCl, 10% glycerol, 0.5% Triton X-100, 10 mM imidazole, and 100  $\mu\text{g}/\text{mL}$  lysozyme). Lysates were sonicated for 5 min (10 sec on and off) at an output level of 40% using a sonicate machine, and

were centrifuged for 20 min at 3000 RPM and 4°C. Then, supernatants (soluble) and inclusion bodies (insoluble) of induced and un-induced samples for each *GmIFR* candidate were taken and stored at 4°C. To prepare the inclusion bodies, the pellets were fully resuspended in a 20 mM Tris-HCl (pH 7.4) buffer. Afterwards, they were centrifuged at 3000 RPM for 10 min at 4°C, and the supernatant was carefully removed. The pellets were resuspended again in 10 mL of 20 mM Tris-HCl pH 7.4 containing 1% SDS to solubilize inclusion bodies. Then the selections were incubated at room temperature for 1 hour followed by centrifugation at 3000 RPM for 10 min at room temperature. Finally, the protein concentration and yield were determined using the Bradford assay (Bradford, 1976).

To check that *GmIFRs* are present in soluble form and to confirm the size of proteins based on molecular weight, both un-induced and induced samples (20 µg) were separated on 10% sodium dodecyl sulfate (SDS) polyacrylamide electrophoretic gel.

### **2.6.3 Western blot analysis**

Total soluble proteins (15 µg) were loaded on a 10% SDS- polyacrylamide electrophoretic gel. The proteins from the gel were then transferred onto an Immun-Blot™ polyvinylidene fluoride (PVDF) membrane (Bio-Rad) using a Trans-Blot SemiDry Electrophoretic Transfer Cell (Bio- Rad) at 20 V for 20 min. The membrane was washed with tris-buffered saline (TBS)+0.1% Tween 20 three times for 15 min followed by blocking in TBS+1% bovine serum albumin (BSA) and 0.1% Tween 20 at 4°C overnight. After blocking, the membrane was washed three times with washing buffer (TBS+0.1% Tween 20). His<sub>6</sub> tagged-*GmIFR* proteins were detected using a

Monoclonal Anti His mouse primary antibody (Sigma-Aldrich) and 1.5 mL of conjugated horseradish peroxidase (HRP) Goat Anti-Mouse IgG (1:5000 dilution) secondary antibody (EMD Millipore). The immune complexes, which had bound to their targets, were detected using enhanced chemiluminescence (ECL) Prime Western Blot detection reagents (GE Health Care Life Sciences) and then exposed using the MicroChemi system by DNR Bio-Imaging (Bio-Rad).

#### **2.6.4 Protein purification**

For purification of *GmIFR* proteins, after confirming the appropriate concentration of IPTG to achieve the highest protein expression, and the proteins were expressed in a soluble form, the protein expression in a larger volume of 400 mL LB culture was repeated as described in the section 2.6.2. The soluble *GmIFR* protein samples were prepared by centrifuging at 10 min at 3000 RPM and 4°C, and filtering through a 0.45 µm filter to avoid purification column clogging and remove cell debris or other unwanted material. Since the His<sub>6</sub>-tag interacts specifically with nickel resin were selectively captured and isolated the target proteins from complex mixtures, by using nickel glass Econo-chromatography columns (2.5 × 20 cm; Bio Rad) as an immobilized metal affinity chromatography (IMAC).

First, the columns were washed with 5 column volume (CV) of distilled water, and then were calibrated with 5 CV of binding buffer (20 mM sodium phosphate, 0.5 M NaCl, and 20 mM imidazole). Two mL of 50% slurry of nickel sepharose (Cytiva) with a binding capacity/mL chromatography medium around 40 mg His<sub>6</sub>-tag proteins was made with binding buffer and added into the calibrated column. Filtered protein samples were added

into nickel resin in the column and incubated on a racker for 1 hour at 4°C in the cold room. After incubation, the prepared protein samples with nickel and binding buffer were allowed to reach the column bead using a holder. Then, the flow through of each *GmIFR* candidates were collected in new test tubes on ice. The columns were washed with 5 CV of two washing buffers (mix of 20 mM sodium phosphate, 0.5 M NaCl, and 35 mM imidazole; and filtered mix of 20 mM sodium phosphate, 0.5 M NaCl, and 55 mM imidazole) contain different imidazole concentrations, and the washed samples were collected. His<sub>6</sub> tagged-*GmIFR*s were eluted 4 times with 5 mL elution buffer at a high concentration of imidazole (filtered mix of 20 mM sodium phosphate, 0.5 M NaCl, and 350 mM imidazole) which can compete with histidine for binding to nickel and pure protein from the total protein extracted from *E. coli* Rosetta-gami (DE3). The eluates were checked on SDS-polyacrylamide gel beside induced, uninduced, wash, and flow through. The concentration of confirmed eluates was determined using the Bradford assay.

### **2.6.5 Desalting and buffer exchange**

PD-10 desalting columns packed with Sephadex G-25 resin using gravity flow protocol from Cytiva was applied to remove salts, imidazole, and replace the buffer from purified *GmIFR* samples. The PD-10 columns were washed twice with 3 mL of distilled water and were equilibrated four times with 3 mL of 0.1 M Tris-HCl buffer (pH 8.5). The eluted *GmIFR* samples (maximum 3 mL) were applied an equilibrated PD-10 column, and buffer was replaced with 3 mL of 0.1 M Tris-HCl buffer (pH 8.5). The desalted and buffer exchanged eluates were collected in 1.5 mL tubes, checked on

SDS-polyacrylamide gel, and measured the concentration using Bradford assay. At the end, the purified *GmIFR* samples were stored at -80°C.

### **2.6.6 Cleavage of His<sub>6</sub> tag from *GmIFR* candidates**

For further purification of *GmIFRs*, tobacco etch virus (TEV) protease was used to cleave the His<sub>6</sub>-tags from the *GmIFR* candidates, as described by Block et al. (2009). To initiate the cleavage reaction, a total reaction volume of 50 µL, which consisted of 15 µg of purified *GmIFR1A* with the His<sub>6</sub>-tag, 5 µL of a 10X TEV protease reaction buffer, and 1 µL of TEV protease was prepared. The reaction mixture was then incubated at 30°C for 1 hour as designed by the New England Bio Lab protocol. After the incubation, an Amicon microcentrifugal filter with a 10 kilo-Dalton (kDa) MWCO (Sigma-Aldrich) was used to separate the TEV protease and the cleaved His<sub>6</sub>-tag from *GmIFR*.

### **2.7 Spectrophotometric enzyme assay**

Analysis of *GmIFR* enzyme activity was performed using Tris-HCl buffer (pH 8.5), Nicotinamide adenine dinucleotide phosphate (NADPH) cofactor, and 2'-hydroxydaidzein substrate. Five µg of purified *GmIFR* samples were incubated in 500 µL total reaction volume with 70 mM Tris-HCl buffer (pH 8.5), 1 mM NADPH, 0.1 mM 2'-hydroxydaidzein substrate at 30°C for 1 hour. After incubation, the reactions were stopped by adding two times 500 µL of ethyl acetate and centrifuging at 13000 RPM for 3 min at room temperature. Then the supernatants were combined and transferred to new 1.5 µL centrifuge tubes. The ethyl acetate was evaporated using nitrogen gas (N<sub>2</sub>) for 20 min. The dried samples were re-suspended in 100 µL of

100 % methanol and sonicated for 15 min using a sonication machine (Cole Palmer 8893 Ultrasonic cleaner) to dissolve pellets. Samples were centrifuged for 5 min at 13000 RPM and were filtered through a 0.2  $\mu\text{m}$  filter (Millipore, USA) into a 2 mL amber glass HPLC vial.

## 2.8 HPLC analysis

HPLC analysis was performed to evaluate *Gm*IFR activity using an Agilent 1260 series HPLC system using an established method described by Anguraj Vadivel et al. (2018). The analysis employed a 10  $\mu\text{L}$  injection in a mobile phase of 0.8 mL/min onto a Phenomenex Kinetex XB-C18 column (4.6 mm, 100 mm, 2.6  $\mu\text{m}$ ) with a controlled temperature of 35°C. A mobile-phase gradient was applied consisted of two solvents including water with 0.1% trifluoroacetic acid (TFA) as solvent A, and acetonitrile with 0.1% TFA as solvent B in the range of 12–100%. The flow rate was maintained at 1 mL min<sup>-1</sup> with detection at 257 nm. A 5  $\mu\text{L}$  sample was injected, and the column temperature was controlled at 35°C. The gradient elution program was configured as follows: mobile phase B was maintained at 12% for 1 min, after which it was gradually raised to 35% over 16 min. Subsequently, mobile phase B, was swiftly increased to 80% for 1 min, and within the subsequent 1.5 min, it reached 100% B. This composition was sustained for 2 min before reverting to the initial condition of 12% B over a duration of 1 min. The system then remained at 12% B for an additional 3.5 min, resulting in a total runtime of 24 min. The UV spectrum and retention time of the substrate and product were compared to synthesized 2'-hydroxydaidzein substrate in Dhaubhadel laboratory.



## Chapter 3

### 3 RESULTS

#### 3.1 Identification of *Gm*IFR candidates in soybean based on *in silico* analysis

Identification of putative soybean *Gm*IFRs was performed through a keyword search using the words “isoflavone reductase” and “oxidoreductase” within the annotated *G. max* Wm82.a4.v1 genome, in the Phytozome 13 database (<https://phytozome-next.jgi.doe.gov/>). This search yielded 83 hits. Subsequently, using Glyma.01G172600.1 and CAA41106.1 as query sequences in a BLASTp search, 34 and 26 putative *Gm*IFRs were found, respectively, using a cut off from 40% amino acid identity, E-value of 1e-10, and bit-score greater than 60 as threshold. A total of 97 putative *Gm*IFRs were obtained by merging the outputs from all three searches and removing any recurring sequences.

##### 3.1.1 Phylogenetic analysis clustered ten *Gm*IFRs candidates with identified IFRs in legumes

To identify *Gm*IFRs, a phylogenetic tree was created using the amino acid sequences of all 97 putative *Gm*IFRs, along with previously characterized IFRs from other legume plants (**Table 3.1**). Out of 97 putative *Gm*IFRs, only 10 protein candidates clustered with previously characterized IFRs from other leguminous plants. The putative

Table 3.1 The list of identified IFRs in legume plants

Gen ID	Given name	Plant species	Peptide length (aa)	Reference
<i>Glyma.01G172600.1</i>	<i>GmIFR</i>	<i>Glycine max</i> (Soybean)	319	(Cheng et al., 2015)
<i>CAA41106.1</i>	<i>MsIFR</i>	<i>Medicago sativa</i> (Alfalfa)	307	(Wang et al., 2006a)
<i>Phvul.002G0333.1</i>	<i>PvIFR</i>	<i>Phaseolus vulgaris</i> (Common bean)	308	(Rípodas et al., 2013)
<i>AAB31368.1</i>	<i>PsIFR</i>	<i>Pisum sativum</i> (Pea)	318	(Sun et al., 1991)
<i>Lj2g0027531</i>	<i>LjIFR</i>	<i>Lotus japonicus</i>	308	(García-Calderón et al., 2020)
<i>Ca.02018</i>	<i>CaIFR</i>	<i>Cicer arietinum</i> (Chickpea)	318	(Schlieper et al., 1990)
<i>Medtr5g020740.1</i>	<i>MtIFR</i>	<i>Medicago truncatula</i>	309	(Liu et al., 2003)

*GmIFR* candidates were named in the following manner: *GmIFR1A*, *GmIFR11A*, *GmIFR1B*, *GmIFR11B*, *GmIFR1D*, *GmIFR11D*, *GmIFR1E*, *GmIFR11E*, *GmIFR1C*, and *GmIFR11C*. These names (given names) are based on the order of chromosome number and unique gene identifier number, which serves to distinguish each gene within the genome (**Figure 3.1**).

### **3.1.2 *GmIFRs* contain three conserved motifs**

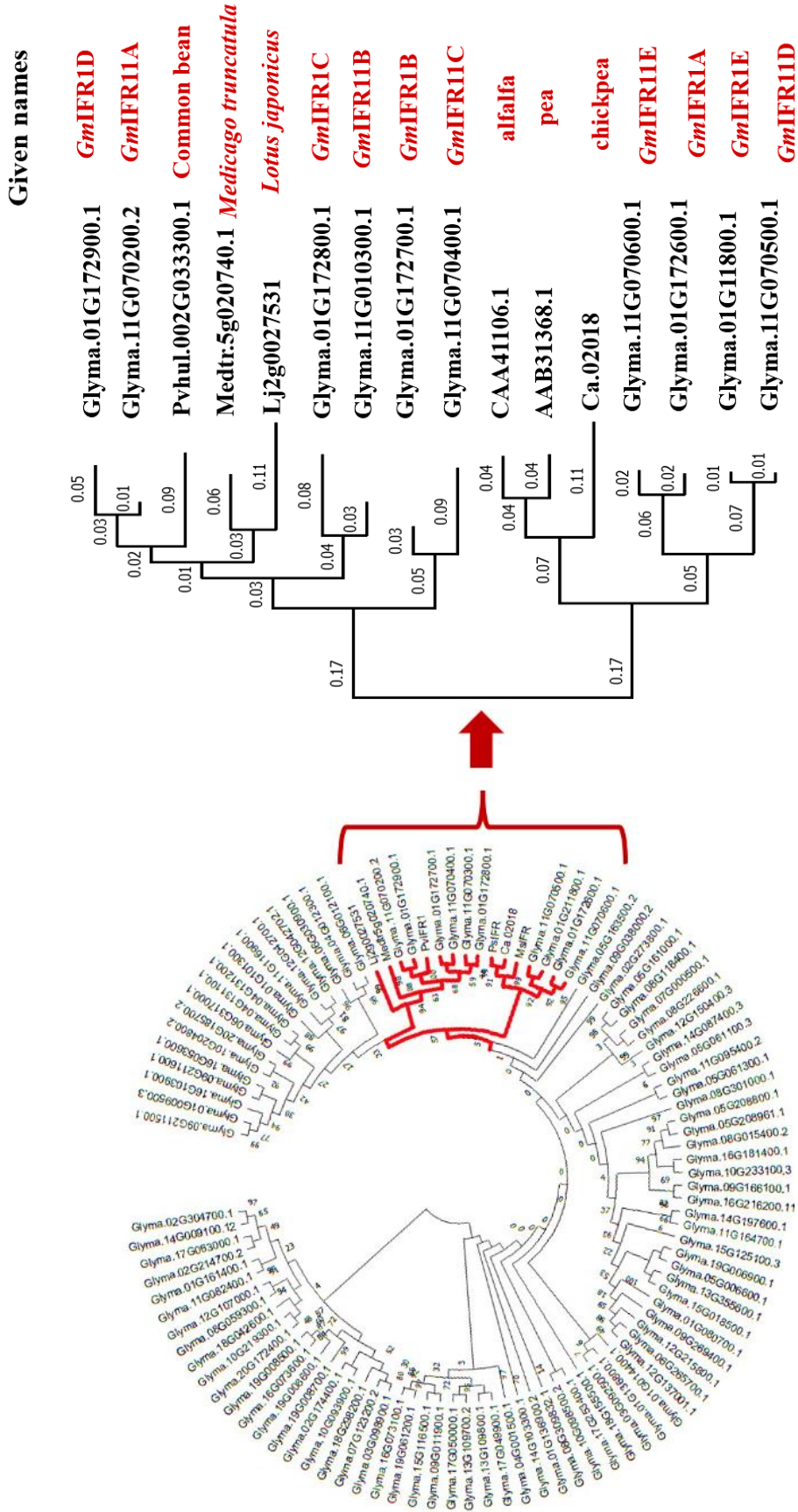
IFRs contain three conserved motifs including the NADPH-binding motif at the N-terminus and two substrate-binding motifs at C-terminus of proteins sequence (Bhinija et al., 2022; Wang et al., 2006a) (**Figure 3.2a**). Based on this, a screening process was conducted to examine all 10 putative *GmIFRs*. All *GmIFR* amino acid sequences were aligned and the presence of NADPH-binding (GXXGXXG) at the N-terminus and isoflavone substrate binding (DPXXLNK and EASXXYPXV) motifs at the C-terminus of *GmIFR* proteins sequence were identified. As depicted in Figure 3.2b, all *GmIFR* candidates exhibit the presence of three conserved motifs, except for *GmIFR11C*, which has short peptide sequence and lacks the substrate-binding motif (EASXXYPXV) at the C-terminus (**Figure 3.2b**).

### **3.1.3 *GmIFR* genes family members display tissue-specific gene expression**

To determine the tissue-specific expression patterns of all 10 putative *GmIFRs*, publicly available RNA-Seq data from soybase (Libault et al., 2010; Wang et al., 2010) of

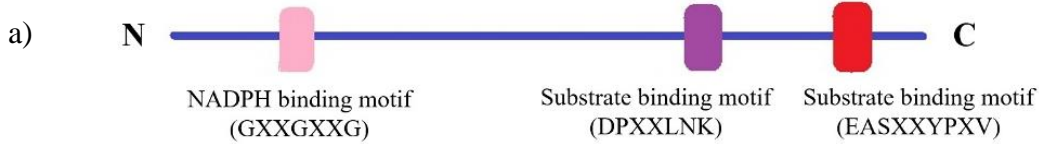
### **Figure 3.1 Phylogenetic analysis of *GmIFRs***

The protein sequences of the *GmIFRs* and IFRs of other legumes were aligned using ClustalW, followed by the construction of a Maximum likelihood tree with 1000 bootstrap replications using MEGA-X software. The tree indicates the relationship between candidates and their genetic distance. *GmIFR* candidates are shown with their given names in smaller tree, and the numbers on tree are the branch lengths that represent genetic distance between candidates and IFRs in other legume plants.



**Figure 3.2 Identification of conserved motifs in *GmIFR* candidates**

**a)** Schematic diagram of *GmIFRs* conserved motifs. Boxes show conserved motifs on *GmIFRs* protein sequence. **b)** Candidate *GmIFRs* were aligned using ClustalW with characterized *MsIFR*. *GmIFRs* contain three conserved motifs including the NADPH-binding motif at the N-terminus and two substrate-binding motifs at C-terminus which are highlighted in blue, pink, and orange boxes, respectively.



b)

**NADPH binding motif (GXXGXXG)**

	90	100	110	120	130	140	150	160	
<i>GmIFR1A</i>	KDRLLVIGPTCAIGRHIVASVKNAGNPTFLVROTPASVNKPRLVAAANPEETREELTQSFQNSGVTLIQGDLNDHESLVN	83							
<i>GmIFR11E</i>	KDRLLVIGPTCAIGRHIVASVKNAGNPTFLVROTPASVNKPRLVAAANPEETREELTQSFQNSGVTLIQGDLNDHESLVN	83							
<i>GmIFR11D</i>	KDRLLVIGPTCAIGRHIVASVKNAGNPTFLVROTPASVNKPRLVAAANPEETREELTQSFQNSGVTLIQGDLNDHESLVN	83							
<i>GmIFR11A.1</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	140							
<i>GmIFR11A.2</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	72							
<i>GmIFR11B</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	72							
<i>GmIFR1D.1</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	148							
<i>GmIFR1D.2</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	72							
<i>GmIFR1B</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	72							
<i>GmIFR1C</i>	RDRLLVIGPTCAIGRHIVASVKNAGNPTFLVROTPASVNRVNLVKAANPEETREELTQSFQNSGVTLIQGDLNDHESLVN	83							
<i>GmIFR1E</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	67							
<i>GmIFR11C</i>	RSKILLVIGGTCYIGKPIVWASABAGNPTFLVRES-----VLSHPE-RSKLLIESFKNSGVTLIQGDLNDHESLVK	72							
<i>MsIFR</i>	ENKILLVIGPTCAIGRHIVASVKNAGNPTFLVROTPASVNRVNLVKAANPEETREELTQSFQNSGVTLIQGDLNDHESLVK	72							

	170	180	190	200	210	220	230	240	
<i>GmIFR1A</i>	AIKQVDVVI CSFGRLLDEQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	163							
<i>GmIFR11E</i>	AIKQVDVVI CSFGRLLDEQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	163							
<i>GmIFR11D</i>	AIKQVDVVI CSFGRLLDEQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	163							
<i>GmIFR11A.1</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	220							
<i>GmIFR11A.2</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	152							
<i>GmIFR11B</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	152							
<i>GmIFR1D.1</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	228							
<i>GmIFR1D.2</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	152							
<i>GmIFR1B</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	152							
<i>GmIFR1C</i>	AIKQVDVVI SALGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	163							
<i>GmIFR1E</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	163							
<i>GmIFR11C</i>	AIKQVDVVI STLGGQCHDQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	105							
<i>MsIFR</i>	AIKQVDVVI CAAGRLLDEQVKI IAAIKKEAGNPKRF PSEFGLDVRHDAVEPVRVREVFBEKAKIRRIEAEGIPYTYLCC	152							

**Substrate binding motif (DPXXLNK)**

	250	260	270	280	290	300	310	320	
<i>GmIFR1A</i>	HAEFGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	243							
<i>GmIFR11E</i>	HAEFGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	243							
<i>GmIFR11D</i>	HAEFGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	243							
<i>GmIFR11A.1</i>	NBAAGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	300							
<i>GmIFR11A.2</i>	NBAAGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	232							
<i>GmIFR11B</i>	NBAAGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	232							
<i>GmIFR1D.1</i>	NBAAGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	308							
<i>GmIFR1D.2</i>	NBAAGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	232							
<i>GmIFR1B</i>	NBAAGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	232							
<i>GmIFR1C</i>	NBAAGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	243							
<i>GmIFR1E</i>	HAEFGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	185							
<i>GmIFR11C</i>	HAEFGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	232							
<i>MsIFR</i>	HAEFGYFIRNACLDITVPPRDKVI LLDGDNVKGAVYVDEAVGATIRKAVDPRTLNKLHYRPPANILITENELVSLWEN	232							

**Substrate binding motif (EASXXYPXV)**

	330	340	350	360	370	380	390	400	
<i>GmIFR1A</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	318							
<i>GmIFR11E</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	318							
<i>GmIFR11D</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	318							
<i>GmIFR11A.1</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	376							
<i>GmIFR11A.2</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	308							
<i>GmIFR11B</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	307							
<i>GmIFR1D.1</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	384							
<i>GmIFR1D.2</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	308							
<i>GmIFR1B</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	307							
<i>GmIFR1C</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	307							
<i>GmIFR1E</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	318							
<i>GmIFR11C</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	249							
<i>MsIFR</i>	KIKGFLERKIVSEEVVAKQKQESFPANNYLALYHSQQIKGDAV-YEIDAKDI EASEAYPDVNTTVSDYLNQV	307							

a variety of soybean tissues and development stages including of seed development, pod, root, young leaves, flowers, and nodules in normal soybean's growth condition was used to create a heatmap (**Figure 3.3a**). Additionally, to enhance the understanding of the impact of *P. sojae* infection on the expression patterns of *GmIFR* candidates, PRJNA324419 (Li et al., 2016), PRJNA544432 (Jahan et al., 2020), PRJNA574764 (de Ronne et al., 2020), and PRJNA318321 (Jing et al., 2016) datasets were used. The datasets had samples taken at different times and tissues after *P. sojae* infection in soybean plants, representing varieties with resistance or susceptibility such as Misty. One dataset, PRJNA574764, looked specifically at partial resistance. All four bioprojects in this study aimed to find *GmIFR* candidates showing transcript expression in root tissue related to resistance against *P. sojae* either *RPS*-mediated or quantitative. So, a heatmap were obtained from total RNA samples collected from various tissues, including seedlings, hypocotyls, hairy roots, and roots from soybean cultivars such as Williams 82, Misty, and PI449459, both before and after *P. sojae* infection. Since soybean root rot caused by *P. sojae* is a soil borne disease, only the *GmIFR* genes expressed in roots were selected for further study. Out of 10 putative *GmIFRs*, *GmIFR1A*, *GmIFR11A*, *GmIFR1B*, *GmIFR1E*, *GmIFR11E*, *GmIFR1D*, and *GmIFR11D* showed higher transcript expression level in roots as compared to other tissues (**Figure 3.3a**). In Figure 3.3a, *GmIFR11B* and *GmIFR1C* display elevated transcript expression levels only in seeds, while *GmIFR11C* does not manifest any expression in the examined tissues.

Expanding my analysis to infected tissues, including seedlings, hypocotyl, hairy roots, and roots following *P. sojae* infection, *GmIFR1A*, *GmIFR11A*, *GmIFR1B*, *GmIFR1E*, *GmIFR11E*, *GmIFR1D*, and *GmIFR11D* exhibit increased transcript expression levels

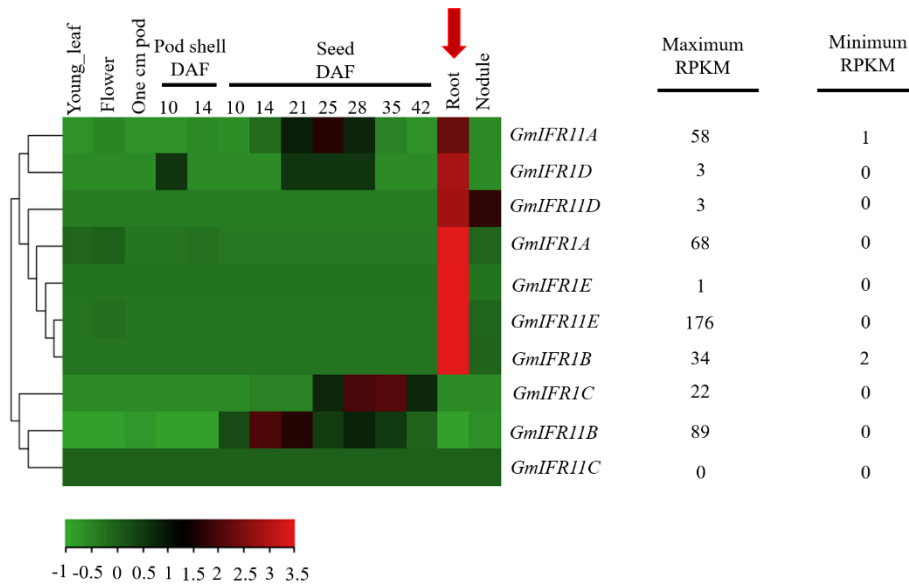


in root tissues following 7 and 14 hours *P. sojae* infection. However, resistance soybean cultivars such as PI449459 and Williams, *RPS*-mediated or quantitative, showed high transcript expression levels of *GmIFRs* in root tissues even before *P. sojae* infection. Interestingly, the transcript expression levels of *GmIFR11B*, *GmIFR1C*, and *GmIFR11C* remain consistently low, both before and after pathogen attack (**Figure 3.3b**). Consequently, only *GmIFR1A*, *GmIFR11A*, *GmIFR1B*, *GmIFR1E*,

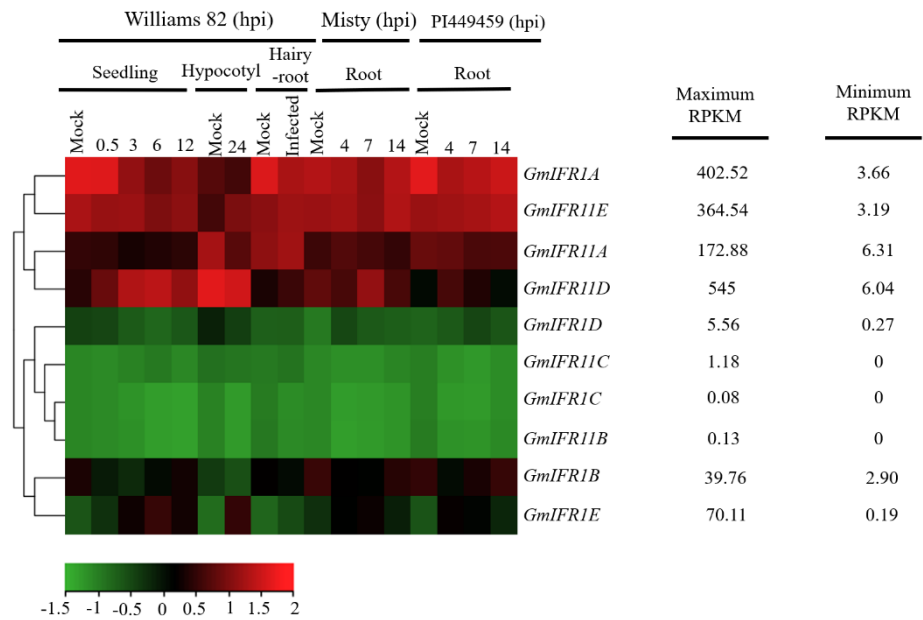
**Figure 3.3 Tissue-specific expression profile of the *GmIFRs* gene family**

Transcript expression of candidates *GmIFR* in different infected and un-infected tissues across various soybean cultivars, as obtained from **a)** Soybase, and **b)** PRJNA324419, PRJNA574764, and PRJNA318321 public RNA-Seq databases with RPKM values, stands for reads per kilobase million to quantify gene expression levels. DAF: day after formation, Mock: un-pathogen infected tissues, hpi: hours after pathogen infection. The color scale indicates normalized expression value, red color shows the highest and green indicating low transcript expression levels on both heatmaps. Maximum and minimum RPKM values for each gene are shown to the right of the diagram.

a)



b)



*GmIFR11E*, *GmIFR1D*, and *GmIFR11D*, which consistently exhibit transcript expression in root tissue before and after *P. sojae* infection, were selected for further analysis.

### 3.1.4 The soybean genome has seven *GmIFR* candidates

After an extensive *in silico* analysis involving various techniques such as keyword search, BLASTp; and multiple alignment, phylogenetic tree construction, identification of conserved motifs, and tissue-specific expression analysis, seven *GmIFRs*, excluding *GmIFR11B*, *GmIFR1C*, and *GmIFR11C* with no transcript expression in root tissue, have been selected as the putative candidates in soybean (**Figure 3.4; Table 3.2**).

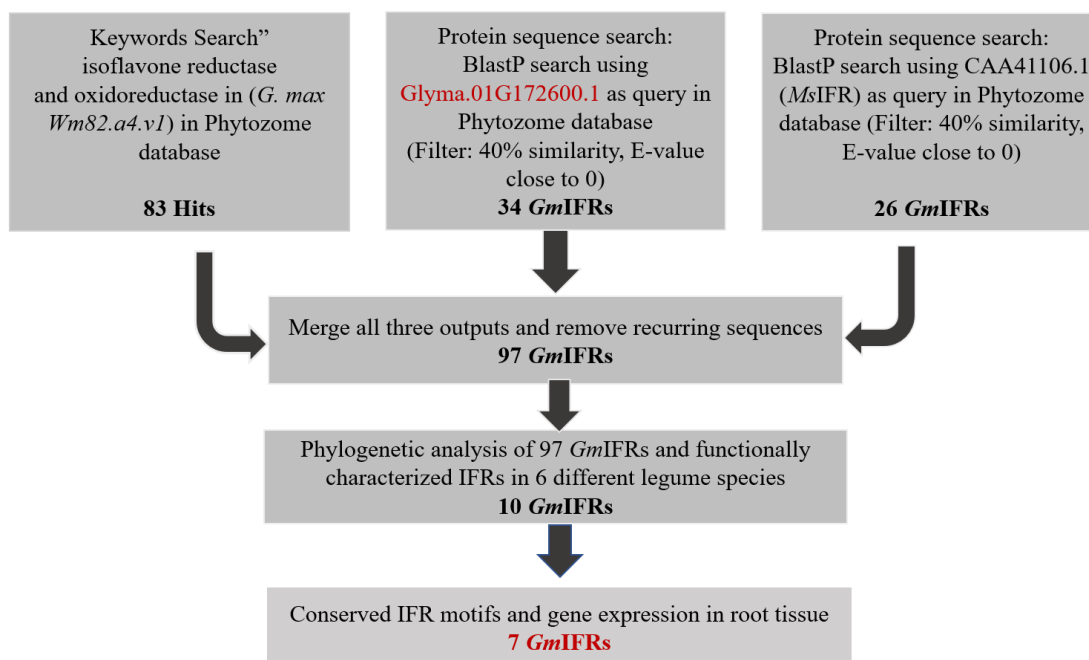
### 3.2 Six *GmIFR* candidates localize in cytoplasm

Previously, *GmIFR1A*, which was the first identified and characterized IFR in soybean, was shown to localize to cytoplasm (Cheng et al., 2015). So, *GmIFR1A* was used as a reference for subcellular localization in my study. To determine the subcellular localization of the putative *GmIFR* candidates, each candidate was translationally fused with YFP. The *GmIFR*-YFP fusion protein was transiently expressed in leaf cells of *N. benthamiana* and confocal microscopy was used to visualize the YFP fluorescence. The results revealed *GmIFR1A*, *GmIFR1B*, *GmIFR11D*, *GmIFR1E*, and *GmIFR11E* localize to the cytoplasm and nucleus, and *GmIFR11A* localizes in cytoplasm. Only *GmIFR1D* was localized to the endoplasmic reticulum (ER) (**Figure 3.5a**) and this location was confirmed through co-localization of *GmIFR1D*-YFP with ER organelle marker-CFP (**Figure 3.5b**). Given that both the glyceollin pathway and IFRs

are localized in the cytoplasm, only cytoplasmic candidate *GmIFRs* are considered further, and *GmIFR1D* was not included in the protein expression study.

**Figure 3.4 Summary of all identification of *GmIFR* candidates process using *in silico*, phylogenetic, tissue-specific expression, and conserved motifs**

Identification of *GmIFR* candidates using keyword searches including 'isoflavone reductase' and 'oxidoreductase' in the annotated *G. max* Wm82.a4. v1 genome in the Phytozome 13 database shows 83 hits. In addition, protein BLAST search using Glyma.01G172600.1, the first identified IFR in *G. max*, and CAA41106.1, IFR in *M. sativa*, as queries shows 34 and 26 candidates, respectively. After an extensive *in silico* analysis, phylogenetic tree construction, identification of conserved motifs, and tissue-specific expression analysis introduced seven *GmIFRs* as the putative candidates in soybean.



**Table 3.2 Characteristics of candidate *GmIFRs***

<b>Gene ID</b>	<b>Given name</b>	<b>Predicted protein mol. weight (KDa)</b>	<b>CDS sequence (bp)</b>	<b>Peptide sequence (aa)</b>
<i>Glyma.01G172600</i>	* <i>GmIFR1A</i>	35.58	957	319
<i>Glyma.11G070200</i>	<i>GmIFR11A</i>	41.85	1131	377
<i>Glyma.01G172700</i>	<i>GmIFR1B</i>	33.77	924	308
<i>Glyma.01G172900</i>	<i>GmIFR1D</i>	42.98	1155	385
<i>Glyma.11G070500</i>	<i>GmIFR11D</i>	35.53	957	319
<i>Glyma.01G211800</i>	<i>GmIFR1E</i>	35.51	957	319
<i>Glyma.11G070600</i>	<i>GmIFR11E</i>	35.58	957	319

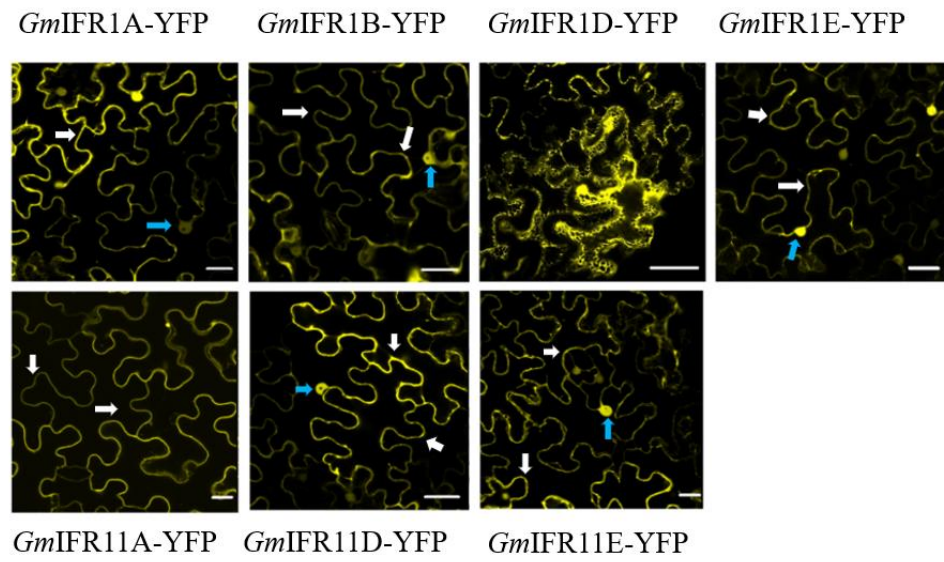
\* *GmIFR1A* is the first identified and characterized *IFR* in soybean (Cheng et al., 2015).



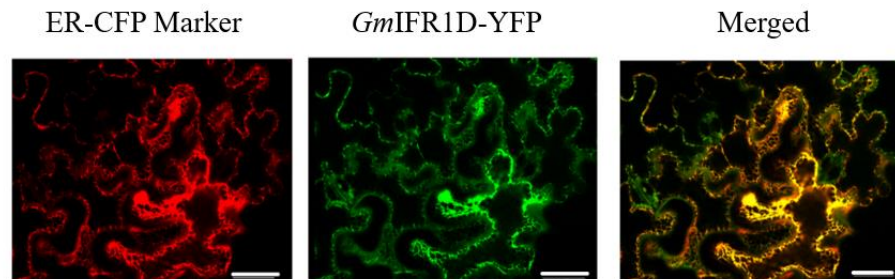
**Figure 3.5 Subcellular localization of the candidate *GmIFRs***

**a)** Translational fusions of *GmIFR*-YFP were transiently expressed in the leaf of *N. benthamiana* and visualized with confocal microscopy. White arrows show cytoplasm and blue arrows show nucleus. The scale bar represents 40  $\mu\text{m}$ . **b)** Confirmation of localization of *GmIFR1D* to the ER was performed through co-localization of *GmIFR1D*-YFP fusion with an ER marker fused to cyan fluorescence protein (CFP). The merged signal was obtained by sequential scanning of the two channels. The scale bar represents 40  $\mu\text{m}$ .

a)



b)



### 3.3 Five *GmIFR* candidates were expressed and purified

For functional characterization of *GmIFR* candidates, five candidates were cloned successfully into the pET160 DEST vector and transformed into *E. coli* Rosetta-gami (DE3) at Section 2.4.2. Despite several attempts, cloning *GmIFR1B* into the entry vector (pDONRz) was not successful. Hence, only *GmIFR1A*, *GmIFR11A*, *GmIFR1E*, *GmIFR11E*, and *GmIFR11D* were expressed using two different concentrations of IPTG induction (0.5 mM and 1 mM). Subsequently, the total soluble proteins were extracted, and it was observed that all *GmIFR* candidates had more expression when induced with 1 mM IPTG (**Figure 3.6**). To further confirm the presence of *GmIFRs* with His<sub>6</sub>-tag and their molecular weight, western blot analysis was performed. The results showed presence of His<sub>6</sub> tagged-*GmIFRs* in total soluble protein extraction from *E. coli* Rosetta-gami (DE3) through a signal for only His<sub>6</sub> tagged-*GmIFRs* (**Figure 3.7**). Then all five samples were purified, desalted, and presence of proteins were confirmed based on molecular weight on SDS-polyacrylamide gel (**Figure 3.8**).

Given the potential interference of histidines with protein function, the removal of the His<sub>6</sub>-tag was performed for only *GmIFR1A* using TEV protease. This pivotal step improved the purity and integrity of *GmIFR* protein samples (**Figure 3.9**).

### 3.4 Spectrophotometric enzyme assay

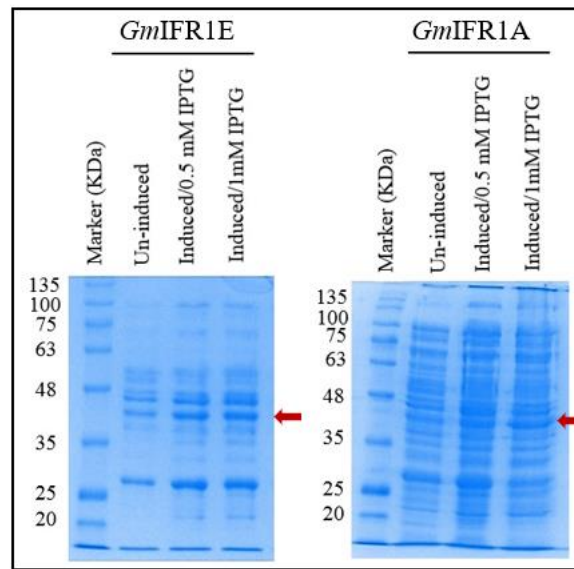
For functional characterization of *GmIFR* candidates, the purified His<sub>6</sub> tagged-*GmIFR* proteins were used to check their ability to convert 2'-hydroxydaidzein to

2'-hydroxy-2, 3-dihydrodaidzein in presence of the NADPH cofactor. The enzyme assay parameters

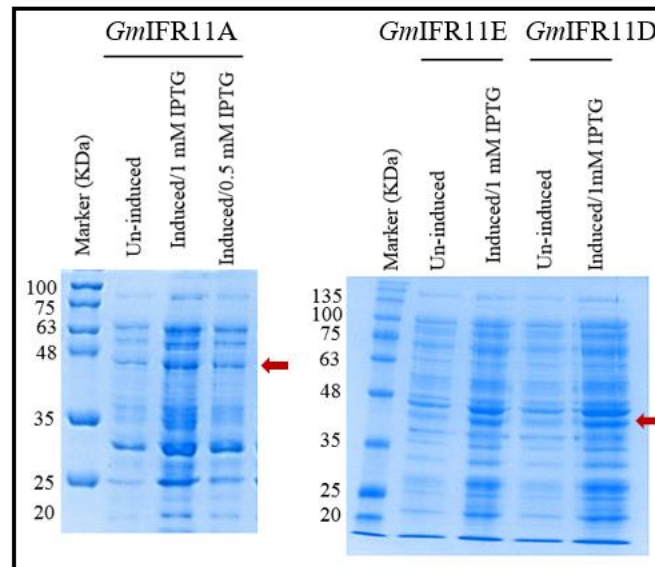
**Figure 3.6 Protein expression of soluble fractions of *GmIFRs* with induction of different concentration of IPTG (0.5 mM and 1 mM)**

**a and b)** *GmIFRs* were induced with 0.5 and 1 mM IPTG at 30°C for 16 hours. Soluble proteins were extracted from *E. coli* Rosetta-gami (DE3) pET160DEST-*GmIFRs* and were separated on 10% SDS-polyacrylamide gel. **a)** The protein expression of *GmIFR1A* and *GmIFR1E*. **b)** The protein expression of *GmIFR11D*, *GmIFR11E*, and *GmIFR11A*. **a and b)** *GmIFR1A* and *GmIFR1E* *GmIFR11D*. The calculated molecular weight of *GmIFR1A* and *GmIFR1E* *GmIFR11D*, and *GmIFR11E* is ~38 KDa; and *GmIFR11A* is ~44.8 KDa.

a)

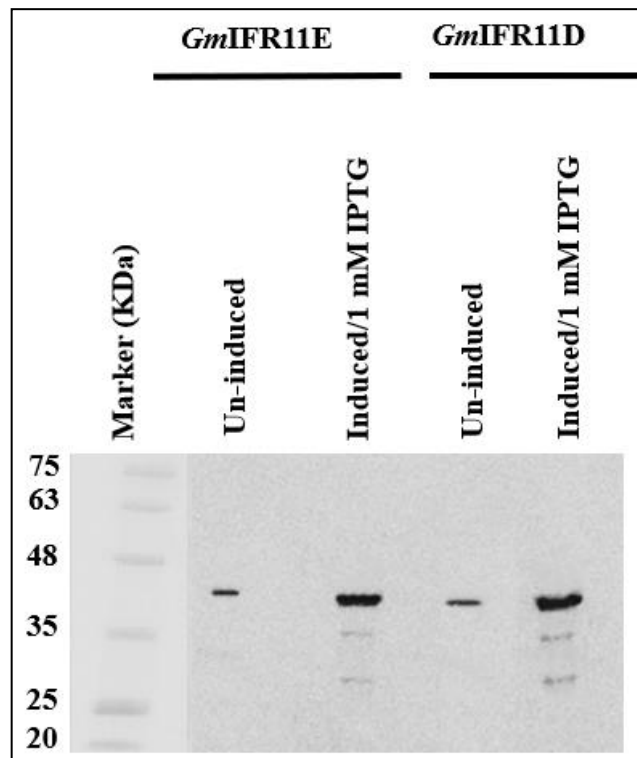
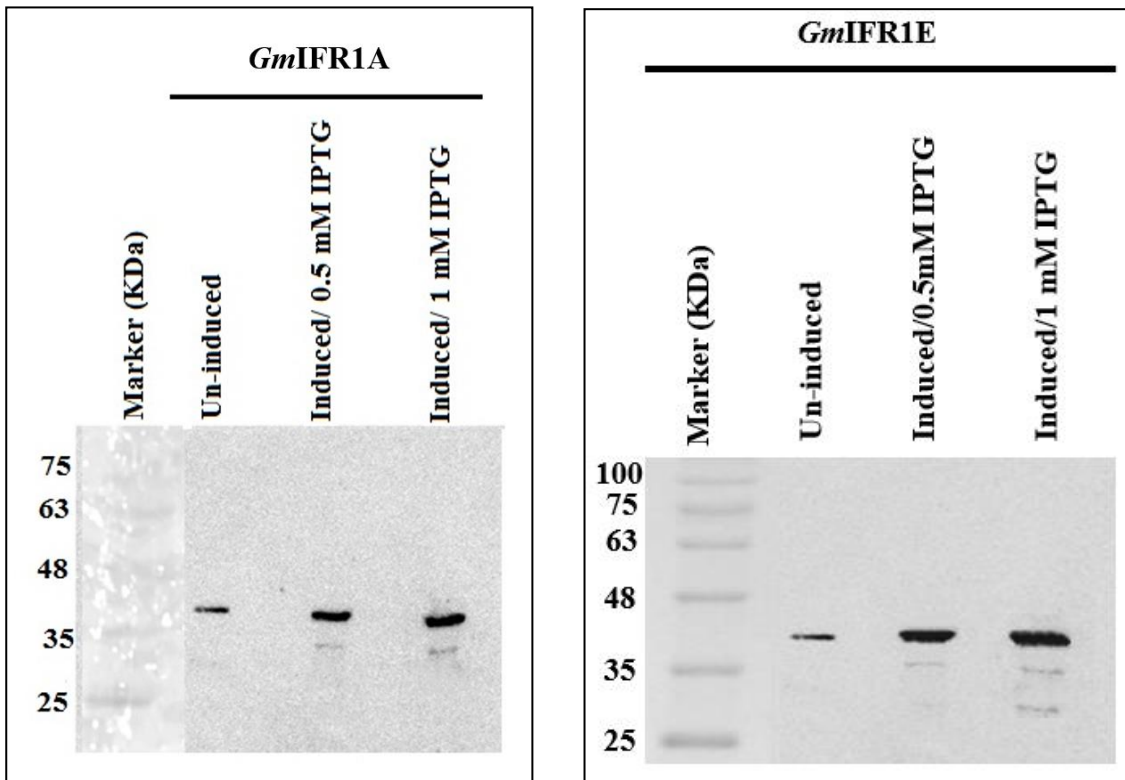


b)



**Figure 3.7 Confirmation of presence of His<sub>6</sub> tagged-*GmIFRs* and their molecular weight using western blot**

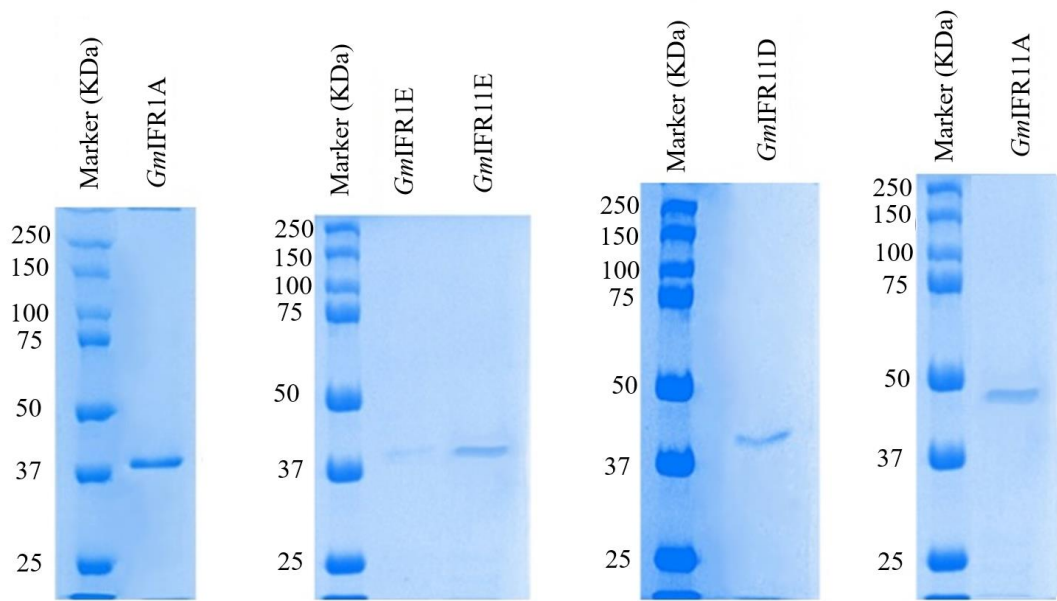
His<sub>6</sub> tagged-*GmIFR1A*, His<sub>6</sub> tagged-*GmIFR1E*, His<sub>6</sub> tagged-*GmIFR11D*, and His<sub>6</sub> tagged-*GmIFR11E* samples including un-induced and induced were separated on 10% SDS-polyacrylamide gel. Presence of His<sub>6</sub> tagged-*GmIFR1A*, His<sub>6</sub> tagged-*GmIFR1E*, His<sub>6</sub> tagged-*GmIFR11D*, and His<sub>6</sub> tagged-*GmIFR11E*, and molecular weight (~38.5 KDa) were confirmed by western blotting. This process involved transferring proteins onto a PVDF membrane using a Trans-Blot SemiDry Electrophoretic Transfer Cell, followed by detection using a Monoclonal Anti-His mouse primary antibody and a conjugated horseradish peroxidase (HRP) Goat Anti-Mouse IgG secondary antibody.





**Figure 3.8 Protein purification of *GmIFRs* using nickel resin and Econo column**

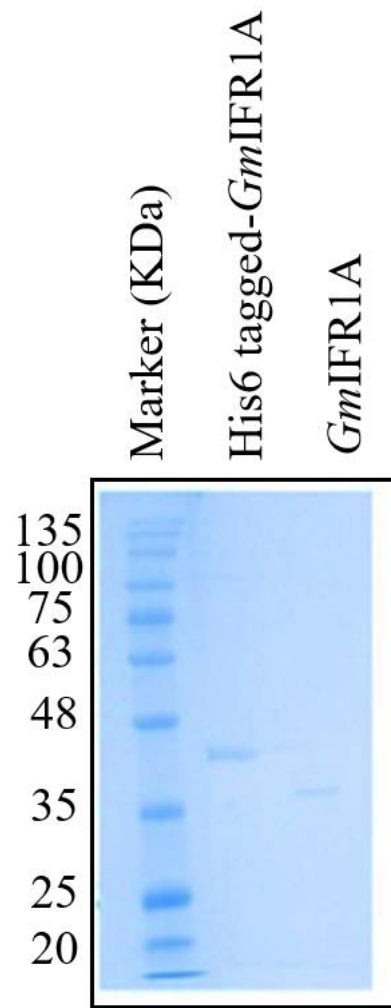
*GmIFR1A*, *GmIFR1E*, *GmIFR11E*, *GmIFR11D*, and *GmIFR11A* were purified, desalted, and separated on 10% SDS-polyacrylamide gel. The calculated molecular weight of *GmIFR1A* and *GmIFR1E*, *GmIFR11D*, and *GmIFR11E* is ~38 KDa; and *GmIFR11A* is ~44.8 KDa.



**Figure 3.9 Cleavage His<sub>6</sub> tag of *GmIFR1A* using TEV protease**

Tobacco etch virus (TEV) protease was used to cleave the His<sub>6</sub>-tags from the *GmIFR* candidates. The reaction mixture was then incubated at 30°C for 1 hour and was filtered by an Amicon microcentrifugal filter with a 10 KDa to separate the TEV protease and the cleaved His<sub>6</sub>-tag from *GmIFR*.

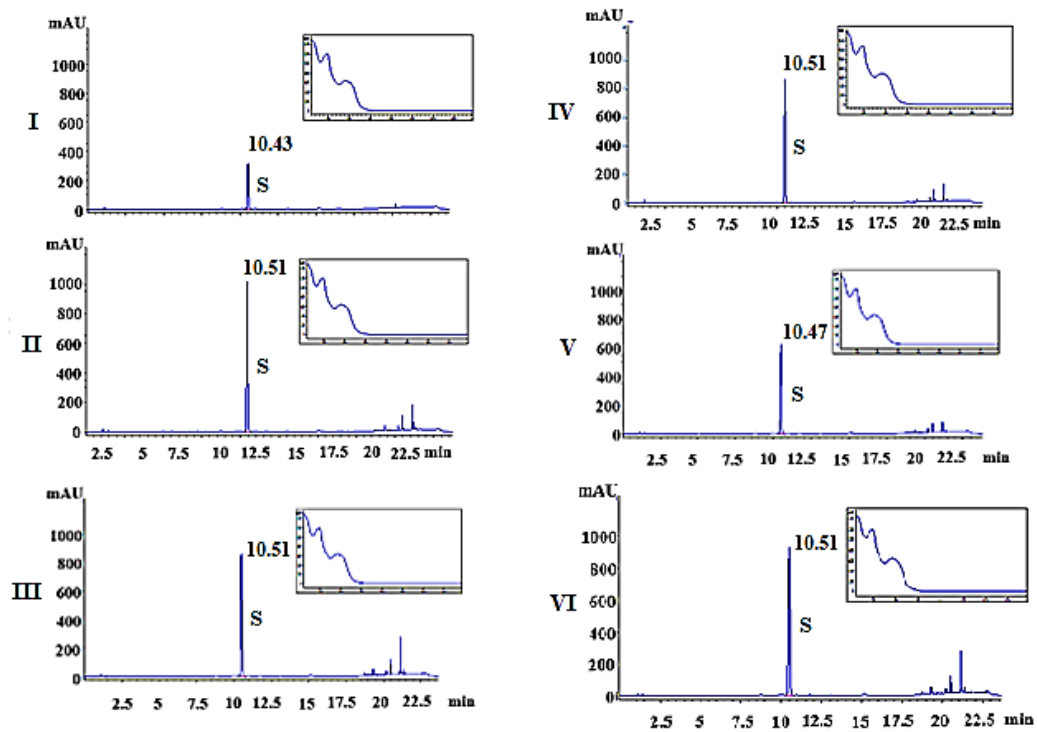
*GmIFR1A* before and after cleaving His<sub>6</sub> tag were loaded on 10% SDS-polyacrylamide gel and the size of each sample was ~38 KDa and ~35 KDa, respectively.



remained consistent across all five purified candidates including *GmIFR1A*, *GmIFR11A*, *GmIFR1E*, *GmIFR11E*, and *GmIFR11D* with His<sub>6</sub>-tag. The reaction with only 2'-hydroxydaidzein substrate was used as a standard. Analysis of the chromatogram revealed a peak at 10.51 minutes retention time for the standard. A product peak before the substrate peak at 10.51 minutes retention time was expected for the functional enzymes. However, no enzyme activity was observed in the reaction of *GmIFR* candidates with NADPH and substrate (**Figure 3.10**). This result showed that enzyme activity under the conditions tested in Section 2.7 was not successful, and more experiments are required to verify enzymatic activity to properly characterize *GmIFRs*.

**Figure 3.10 HPLC chromatogram of candidate *GmIFRs* enzymatic reactions**

His<sub>6</sub> tagged-*GmIFRs* were incubated with 2'-hydroxydaidzein substrate and NADPH cofactor at 30°C for 1 hour. Ethyl acetate extracts were analyzed by high-performance liquid chromatography (HPLC). HPLC chromatograms determined no enzyme activity in the reaction of *GmIFR* candidates with NADPH cofactor and 2'-hydroxydaidzein substrate. S: substrate and mAU: milli-absorbance units. **I)** Reaction using only 2'-hydroxydaidzein substrate as standard with a retention time of 10.43 minutes. **II-VI)** Reactions contain His<sub>6</sub> tagged-*GmIFR1A*, His<sub>6</sub> tagged-*GmIFR1E*, His<sub>6</sub> tagged-*GmIFR11A*, His<sub>6</sub> tagged-*GmIFR11E*, and His<sub>6</sub> tagged-*GmIFR11D* with substrate and NADPH, respectively.



## Chapter 4

### 4 Discussion

An essential part of the soybean defense response is glyceollin production. The production of phytoalexin glyceollins upon exposure of soybean to environmental challenges has been long known (Graham et al., 2007). Upon stresses such as *P. sojae* attack, daidzein (an isoflavone aglycone) acts as a substrate for further enzymatic activity in the production of the glyceollin. The first step towards glyceollin production is the hydroxylation of daidzein by isoflavone 2'-hydroxylase (Akashi et al., 1998). This product is then subsequently reduced by isoflavone reductase (IFR (Fischer et al., 1990)). The enzyme isoflavone reductase of soybean (*GmIFR*) is a member of a large oxidoreductase family that catalyzes the conversion of 2'-hydroxydaidzein to 2'-hydroxy-2,3-dihydrodaidzein in the glyceollin biosynthesis. IFRs are NADPH-dependent oxidoreductase enzymes that are involved in phytoalexins synthesis in legumes (Fischer et al., 1990; Graham et al., 1990). IFRs contain three conserved motifs including the NADPH-binding motif at the N-terminus and two substrate-binding motifs at C-terminus (Bhinija et al., 2022; Wang et al., 2006). IFRs are characterized by multiple legume species. For example, *Medicago sativa* (alfalfa (Wang et al., 2006a)), *Pisum sativum* (pea (Sun et al., 1991)), *Cicer arietinum* (chickpea (Rípodas et al., 2013)), *Lotus japonicus* (García-Calderón et al., 2020), and *Phaseolus vulgaris* (common bean (Schlieper et al., 1990)) each possess a single IFR that was shown to play a role in responding to various stresses. Furthermore, within soybean, a single monomeric and cytosolic IFR was identified and characterized in soybean that catalyzes the conversion of



2'-hydroxyformononetin to an unknown product (Cheng et al., 2015). In addition, this soybean cytosolic IFR displayed a conversion of 2'-hydroxydaidzein to 2'-hydroxydihydrodaidzein in the presence of NADPH (Fischer et al., 1990). However, there is still no knowledge about IFRs in the synthesis of isoflavones, especially the conversion of 2'-hydroxydaidzein to 2'-hydroxy-2,3- dihydrodaidzein in soybean, representing a gap in current understanding of these enzymatic reactions within the glyceollin biosynthesis pathway. In addition, studies showed that the soybean genome contains numerous multi-gene families that resulted from two whole-genome duplications (Schmutz et al., 2010). It is expected that several genes involved in soybean isoflavonoid biosynthetic pathway belong to multigene families such as *chalcone synthase* (*CHS*:14; Anguraj Vadivel et al. (2018), *chalcone reductase* (*CHR*:12; Sepiol et al. (2017), *chalcone isomerase* (*CHI*:9; Dastmalchi and Dhaubhadel (2015), *isoflavone synthase* (*IFS*: 2; Jung et al. (2000), *prenyltransferase* (*PT*; Sukumaran et al. (2018) with very high sequence similarity at the amino acid level, but different tissue-specific expression patterns and roles during development or in response to stress. However, there is no knowledge about the biological function of all *GmIFRs* in soybean. In this study, I report the identification of the putative *GmIFRs* in soybean, investigate their subcellular location, protein expression, and assay enzyme activity. My results demonstrate that *GmIFRs* have three conserved motifs, root tissue expression, localization in cytoplasm.

The insights provided in this study shed light on the process of glyceollin production, an essential part of soybean defense against stresses such as *P. sojae* infection. The study identifies and characterizes the *GmIFRs* in soybeans, which play a crucial role in the

second step of glyceollin biosynthesis as a defense mechanism in soybean crops. Understanding the function and characteristics of IFRs can potentially lead to improved soybean crops with enhanced resistance to stresses. The research also highlights the complexity of soybean genome, indicating the presence of multiple gene families involved in isoflavonoid biosynthesis. Further exploration of *GmIFRs* and their functions in glyceollin biosynthesis could contribute to developing soybean varieties better equipped to reduction of severity of *P. sojae* infection.

#### **4.1 Seven *GmIFR* candidates were identified in soybean**

All members of the *GmIFR* gene family were identified by searching the annotated soybean genome on Phytozome *G. max* Wm82.a4. v1. Using a keyword search together with BLAST searches, a total of 97 *GmIFR* candidates were identified. This number represents the current number of annotated *GmIFRs*. However, this total number may change with additional deposits to the database. Out of these 97, in phylogenetic tree 10 *GmIFRs* were clustered in two clades with previously identified and characterized IFRs in the other legumes (**Figure 3.1**). These 10 *GmIFRs* showed similarity in protein sequences and have a common ancestor (**Figure 3.1; 3.2**). *GmIFR* candidates were selected based on their transcript expression in root tissue, and specifically, for proteins that have all three conserved motifs including one NADPH binding motif and two substrate binding motifs. Finally, based on all identification analyses, only seven *GmIFRs* (*GmIFR1A*, *GmIFR1B*, *GmIFR1E*, *GmIFR1D*, *GmIFR11A*, *GmIFR11D*, and *GmIFR11E*) were further processed as putative candidates. *GmIFR11B*, *GmIFR1C*, and *GmIFR11C* had no expression in root tissue were removed for further studies. Interestingly,

*GmIFR11C* represented a shorter transcript and protein sequences than other *GmIFR* candidates, which maybe the reason for observing the lack of root expression and all other studied tissues and absence of EASXXYPXV substrate binding motif at C-terminus of protein sequence (**Figure 3.2**). However, it is possible that *GmIFR11B*, *GmIFR1C*, and *GmIFR11C* may possess weak enzymatic activity or may be evolved in new catalytic features other than isoflavone reductase activity.

One of these 7 *GmIFR* candidates (Glyma.01G172600 or *GmIFR1A*) was previously identified and characterized in soybean (Cheng et al., 2015). The six remaining *GmIFR* candidates (*GmIFR11A*, *GmIFR1B*, *GmIFR1D*, *GmIFR11D*, *GmIFR1E*, and *GmIFR11E*) have not been identified and characterized in soybean before my study. These candidates could potentially be involved in a process related to the production of glyceollin through isoflavone oxidoreductase activity dependent NADPH cofactors.

## **4.2 *GmIFRs* were localized in the cytoplasm**

The Wolf PSORT program predicted a cytoplasmic localization of *GmIFR* candidates except *GmIFR1D* (Horton, 2006). All candidates of the *GmIFR* family displayed cytoplasmic localization in *N. benthamiana* leaf epidermal cells under confocal microscope except for *GmIFR1D*, which exhibited localization to the endoplasmic reticulum (**Figure 3.5**). Consequently, *GmIFR1D* was excluded from further consideration as an IFR candidate in soybean. These findings are consistent with the cytoplasmic localization of *GmIFR1A* which was the first and only identified and characterized IFR in soybean (Cheng et al., 2015). Also, Fischer et al. (1990) had reported IFR is a cytosolic oxidoreductase enzyme. Furthermore, other enzymes involved

in the isoflavonoid biosynthesis such as *GmCHI* (Dastmalchi and Dhaubhadel, 2015), glycosyltransferase (*UGT73F2*) and malonyltransferase (*GmMT7*; (Dhaubhadel et al., 2008) are also localized to the nucleus and the cytoplasm. The ER localization of *GmIFR1D* was confirmed by ER organelle markers. *GmIFR1D* was found within the ER, even though it lacks both a signal peptide and a transmembrane domain in its protein sequence. Previous studies showed that isoflavonoid biosynthesis involves multiple cytochrome P450s such as C4H, IFS (Dastmalchi et al., 2016), I2'H, and 3,9DPO (Khatri et al., 2022) that are ER localized. Cytochrome P450s are able to interact with soluble enzymes such as CHS, CHR and CHI in the phenylpropanoid and isoflavonoid, and form a multi-enzyme complex at the surface of the ER (Dastmalchi et al., 2016). So, it seems likely that *GmIFR1D* may interact with other proteins resident in the ER like P450s, or it may be escorted into the ER by transporter proteins within the organelle.

### **4.3 No product formation for *GmIFR* candidates compared to 2'-hydroxydaidzein standard**

IFRs are enzymes rely on NADPH and belong to the oxidoreductase family, participating in the synthesis of phytoalexins in leguminous plants (Fischer et al., 1990; Graham et al., 1990). *MsIFR* was first identified as a key NADPH-dependent reductase enzyme involved in the latter part of the medicarpin biosynthetic pathway as a phytoalexin in alfalfa which converts 2'-hydroxyformononetin to (3*R*)-vestitone in response to stress (Wang et al., 2006a). The IFRs from pea and chickpea were reported to be active toward 7,2'-dihydroxy-4' ,5'-methylenedioxyisoflavone and 2'-hydroxyformononetin,

respectively (Schlieper et al., 1990; Sun et al., 1991). Also, the first identified and characterized IFR in soybean which is named *GmIFR1A* in my study has oxidoreductase activity with 2'-hydroxyformononetin (Cheng et al., 2015). Therefore, the *GmIFR* candidates that I identified in soybean are likely to be able to reduce 2'-hydroxydaidzein in glyceollin pathway. For the functional characterization of *GmIFR* candidates, *GmIFR1A*, *GmIFR11A*, *GmIFR11D*, *GmIFR1E*, and *GmIFR11E* with His<sub>6</sub>-tags were expressed in *E. coli* Rosetta-gami (DE3), purified using IMAC chromatography, and utilized for enzyme assays in presence of the NADPH cofactor. Unfortunately, analysis of the chromatograms obtained from HPLC revealed no product formation for any of the investigated His<sub>6</sub> tagged-*GmIFR* candidates when compared to the standard (2'-hydroxydaidzein). These findings indicate the necessity for additional research aimed at clarifying the functional roles and enzymatic activities of these *GmIFR* candidates. The absence of product formation in the current assay may be attributed to various factors, including the need to change enzyme assay reaction conditions such as reaction buffer and the mg of enzymes and cofactor, or assay *GmIFRs* functional activity without tagged 6x Histidine.

## 5 Conclusions and future directions

Identification of the *GmIFR* gene family members in glyceollin pathway in soybean is a key step in exploring potential gene candidates for developing soybean cultivars with partial resistance. Overall, I have identified 10 *GmIFR* candidates using *in silico* and phylogenetic analysis including *GmIFR1A*, *GmIFR11A*, *GmIFR1B*, *GmIFR11B*, *GmIFR1C*, *GmIFR11C*, *GmIFR1D*, *GmIFR11D*, *GmIFR1E*, and *GmIFR11E*. Subsequently, *GmIFR11B*, *GmIFR1C*, *GmIFR11C* were eliminated based on analysis of root tissue transcript expression from the publicly available RNA-Seq database and evaluating the presence of three conserved motifs on protein sequence across all 10 candidates. Finally, seven *GmIFRs* (*GmIFR1A*, *GmIFR11A*, *GmIFR1B*, *GmIFR1D*, *GmIFR11D*, *GmIFR1E*, and *GmIFR11E*) were introduced as putative IFR candidates in soybean for characterization studies. All 7 *GmIFRs* localize to the cytoplasm, and *GmIFR1D* had an ER localization and was removed for protein expression and functional assay. This decision was made because the glyceollin pathway primarily takes place in the cytoplasm. All six cytoplasmic *GmIFRs* were cloned into pET160 DEST to add His<sub>6</sub> tag and TEV site to protein expression. Despite several attempts, amplification of *GmIFR1B* was not successful. *GmIFR1A*, *GmIFR11A*, *GmIFR1D*, *GmIFR11D*, *GmIFR1E*, and *GmIFR11E* proteins were expressed in Rosetta-gami, purified, and confirmed in terms of molecular weight on SDS-polyacrylamide gel. Then, the enzymatic activity of all five purified *GmIFR* proteins were analyzed using HPLC. Enzymatic assays of the five selected *GmIFR* candidates did not yield successful results. However, it is crucial to verify enzymatic activity to properly characterize the *GmIFR* family. Identification and subcellular localization studies proved the hypothesis of this research

that soybeans contain multiple *GmIFR* gene family members. However, the analysis of functional activity *GmIFR* candidates was not successful to confirm the second part of the hypothesis that some of *GmIFR* candidates may show enzyme activity to convert 2'-hydroxydaidzein to 2'-hydroxy-2,3- dihydrodaidzein. To confirm enzymatic activity, it is essential to adjust and optimize both the enzymatic reaction conditions such as reaction buffer and the mg of enzymes and cofactor, the HPLC parameters like increasing in a total runtime, decreasing the controlled temperature, and run the reaction with 2'-hydroxy-2,3- dihydrodaidzein product as a positive control. Additionally, because the presence of the His<sub>6</sub> tag can sometimes affect protein function (Carson et al., 2007), it is advisable to conduct enzyme assays using *GmIFRs* without the histidine tag in future. Furthermore, root expression of *GmIFRs* under *P. sojae* infection can offer to corroborate RNA-Seq data for public databases (**Figure 3.3b**). Following the confirmation of enzymatic activity and transcript expression in root tissue during *P. sojae* infection, soybean resistance to *P. sojae* infection should be tested in *GmIFR* silenced soybean plants. Protein-protein interaction should be performed to examine any likely interactions between *GmIFR* members with other proteins in the glyceollin metabolon, and as a strong confirmation to ER localization of *GmIFR1D*. Better understanding of *GmIFR* localization, expression and catalytic activity can result in more precise manipulation of this gene family and produce a more resistant cultivar of soybean.

## References

- Abotaleb, M., S.M. Samuel, E. Varghese, S. Varghese, P. Kubatka, A. Liskova, and D. Büsselberg. 2019. Flavonoids in cancer and apoptosis. *Cancers*. 11:1-28.
- Akashi, T., T. Aoki, and S.-i. Ayabe. 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 1. *Federation of European Biochemical Societies Letters*. 431:287-290.
- Akashi, T., K. Sasaki, T. Aoki, S.-i. Ayabe, and K. Yazaki. 2008. 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 Physiology*. 149:683-693.
- Ali, N. 2010. Soybean processing and utilization. In *The Soybean: Botany, Production and Uses*. Commonwealth Agricultural Bureaux International, Cambridge. pp. 345–374.
- Anguraj Vadivel, A.K., K. Krysiak, G. Tian, and S. Dhaubhadel. 2018. Genome-wide identification and localization of chalcone synthase family in soybean (*Glycine max* [L]Merr). *BioMed Central Plant Biology*. 18:325-338.
- Babiychuk, E., S. Kushnir, E. Belles-Boix, M. Van Montagu, and D. Inzé. 1995. Arabidopsis thaliana NADPH oxidoreductase homologs confer tolerance of yeasts toward the thiol-oxidizing drug diamide. *Journal of Biological Chemistry*. 270:26224-26231.
- Baraibar, M., and L. Deutsch. 2023. The soybean through world history: lessons for sustainable agrofood systems. Taylor and Francis, London. pp. 267 pp.
- Basson, A.R., S. Ahmed, R. Almutairi, B. Seo, and F. Cominelli. 2021. Regulation of Intestinal Inflammation by Soybean and Soy-Derived Compounds. *Foods*. 10:774-804.
- Bebber, D.P., and S.J. Gurr. 2015. Crop-destroying fungal and oomycete pathogens challenge food security. *Fungal Genetics and Biology*. 74:62-64.
- Bhinija, K., P.S. Huehne, S. Mongkolsuk, S. Sitthimonchai, and J. Satayavivad. 2022. A short-chain dehydrogenase/reductase (SDR) detection for the *isoflavone reductase* gene in *Bulbophyllum* and other orchids. *South African Journal of Botany*. 144:295-304.



- Biswas, B., and P.M. Gresshoff. 2014. The Role of Symbiotic Nitrogen Fixation in Sustainable Production of Biofuels. *International Journal of Molecular Sciences*. 15:7380-7397.
- Bizuneh, G.K. 2021. The chemical diversity and biological activities of phytoalexins. *Advances in Traditional Medicine*. 21:31-43.
- Block, H., B. Maertens, A. Spriestersbach, N. Brinker, J. Kubicek, R. Fabis, J. Labahn, and F. Schäfer. 2009. Immobilized-metal affinity chromatography (IMAC): a review. In *Methods in Enzymology*. R.R. Burgess and M.P. Deutscher, editors. Academic Press. pp. 439-473.
- Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*. 72:248-254.
- Broun, P. 2005. Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in *Arabidopsis*. *Current Opinion in Plant Biology*. 8:272-279.
- Candeia, R.A., M.C.D. Silva, J.R. Carvalho Filho, M.G.A. Brasilino, T.C. Bicudo, I.M.G. Santos, and A.G. Souza. 2009. Influence of soybean biodiesel content on basic properties of biodiesel–diesel blends. *Fuel*. 88:738-743.
- Carson, M., D.H. Johnson, H. McDonald, C. Brouillette, and L.J. Delucas. 2007. His-tag impact on structure. *Acta crystallographica*. 63:295-301.
- Cederroth, C.R., and S. Nef. 2009. Soy, phytoestrogens and metabolism: A review. *Molecular and Cellular Endocrinology*. 304:30-42.
- Chen, J., C. Ullah, M. Reichelt, F. Beran, Z.-L. Yang, J. Gershenzon, A. Hammerbacher, and D.G. Vassão. 2020. The phytopathogenic fungus *Sclerotinia sclerotiorum* detoxifies plant glucosinolate hydrolysis products via an isothiocyanate hydrolase. *Nature Communications*. 11:3090.
- Cheng, Q., N. Li, L. Dong, D. Zhang, S. Fan, L. Jiang, X. Wang, P. Xu, and S. Zhang. 2015. Overexpression of soybean isoflavone reductase (*GmIFR*) enhances resistance to *Phytophthora sojae* in soybean. *Frontiers in Plant Science*. 6:1024-1035.
- Chinta, S.J., A. Ganesan, P. Reis-Rodrigues, G.J. Lithgow, and J.K. Andersen. 2013. Anti-inflammatory role of the isoflavone diadzein in lipopolysaccharide-stimulated microglia: implications for parkinson's disease. *Neurotoxicity Research*. 23:145-153.
- Choi, I.-Y., D.L. Hyten, L.K. Matukumalli, Q. Song, J.M. Chaky, C.V. Quigley, K. Chase, K.G. Lark, R.S. Reiter, M.-S. Yoon, E.-Y. Hwang, S.-I. Yi, N.D. Young, R.C. Shoemaker, C.P. van Tassell, J.E. Specht, and P.B. Cregan. 2007. A soybean

- transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics*. 176:685-696.
- Dakora, F.D., and D.A. Phillips. 1996. Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins. *Physiological and Molecular Plant Pathology*. 49:1-20.
- Dale Young, N. 1999. A cautiously optimistic vision for marker-assisted breeding. *Molecular Breeding*. 5:505-510.
- Daniel, S., and W. Barz. 1990. Elicitor-induced metabolic changes in cell cultures of chickpea (*Cicer arietinum* L.) cultivars resistant and susceptible to *Ascochyta rabiei*. II. Differential induction of chalcone-synthase-mRNA activity and analysis of in-vitro-translated protein patterns. *Planta*. 82:279-286.
- Dastmalchi, M., M.A. Bernards, and S. Dhaubhadel. 2016. Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. *The Plant Journal*. 85:689-706.
- Dastmalchi, M., and S. Dhaubhadel. 2015. Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. *Planta*. 241:507-523.
- de Ronne, M., C. Labbé, A. Lebreton, H. Sonah, R. Deshmukh, M. Jean, F. Belzile, L. O'Donoghue, and R. Bélanger. 2020. Integrated QTL mapping, gene expression and nucleotide variation analyses to investigate complex quantitative traits: a case study with the soybean-*Phytophthora sojae* interaction. *Plant Biotechnology Journal*. 18:1492-1494.
- de Ronne, M., P. Santhanam, B. Cinget, C. Labbé, A. Lebreton, H. Ye, T.D. Vuong, H. Hu, B. Valliyodan, D. Edwards, H.T. Nguyen, F. Belzile, and R. Bélanger. 2022. Mapping of partial resistance to *Phytophthora sojae* in soybean PIs using whole-genome sequencing reveals a major QTL. *The Plant Genome*. 15:1-16.
- Dhaubhadel, S. 2011. Regulation of isoflavonoid biosynthesis in soybean seeds. In *Biochemistry, Chemistry and Physiology*. T.-B. Ng, editor. In Tech, Europe. pp. 243-258.
- Dhaubhadel, S., M. Farhangkhomee, and R. Chapman. 2008. Identification and characterization of isoflavonoid specific glycosyltransferase and malonyltransferase from soybean seeds. *Journal of Experimental Botany*. 59:981-994.
- Dhaubhadel, S., M. Gijzen, P. Moy, and M. Farhangkhomee. 2006. Transcriptome analysis reveals a critical role of *CHS7* and *CHS8* genes for isoflavonoid synthesis in soybean seeds. *Plant Physiology*. 143:326-338.

- Dhaubhadel, S., B.D. McGarvey, R. Williams, and M. Gijzen. 2003. Isoflavonoid biosynthesis and accumulation in developing soybean seeds. *Plant Molecular Biology*. 53:733-743.
- Ding, M., A. Pan, J.E. Manson, W.C. Willett, V. Malik, B. Rosner, E. Giovannucci, F.B. Hu, and Q. Sun. 2016. Consumption of soy foods and isoflavones and risk of type 2 diabetes: a pooled analysis of three US cohorts. *European Journal of Clinical Nutrition*. 70:1381-1387.
- Dong, S., D. Yu, L. Cui, D. Qutob, J. Tedman-Jones, S.D. Kale, B.M. Tyler, Y. Wang, and M. Gijzen. 2011. Sequence variants of the *Phytophthora sojae* RXLR effector Avr3a/5 are differentially recognized by *Rps3a* and *Rps5* in soybean. *Plos One*. 6:1-8.
- Dong, X., W. Xu, R.A. Sikes, and C. Wu. 2013. Combination of low dose of genistein and daidzein has synergistic preventive effects on isogenic human prostate cancer cells when compared with individual soy isoflavone. *Food Chemistry*. 141:1923-1933.
- Dorrance, A.E. 2018. Management of *Phytophthora sojae* of soybean: a review and future perspectives. *Canadian Journal of Plant Pathology*. 40:210-219.
- Dorrance, A.E., and S.A. McClure. 2001. Beneficial effects of fungicide seed treatments for soybean cultivars with partial resistance to *Phytophthora sojae*. *Plant Disease*. 85:1063-1068.
- Dorrance, A.E., S.A. McClure, and S.K. St. Martin. 2003. Effect of partial resistance on *Phytophthora* stem rot incidence and yield of soybean in Ohio. *Plant Disease*. 87:308-312.
- Du Bois, C.M. 2018. The story of soy. Reaktion Books, London.
- Du Bois, C.M., C.-B. Tan, and S. Mintz. 2008. The world of soy. University of Illinois Press, Chicago pp. 208–233 pp.
- Durrant, W.E., and X. Dong. 2004. Systemic acquired resistance. *Annual Review of Phytopathology*. 42:185-209.
- Earley, K.W., J.R. Haag, O. Pontes, K. Opper, T. Juehne, K. Song, and C.S. Pikaard. 2006. Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal*. 45:616-629.
- Enstone, D.E., C.A. Peterson, and F. Ma. 2002. Root endodermis and exodermis: structure, function, and responses to the environment. *Journal of Plant Growth Regulation*. 21:335-351.
- Fischer, D., C. Ebenau-Jehle, and H. Grisebach. 1990. Phytoalexin synthesis in soybean: Purification and characterization of NADPH:2'-hydroxydaidzein oxidoreductase

- from elicitor-challenged soybean cell cultures. *Archives of Biochemistry and Biophysics*. 276:390-395.
- Francis, D.M., and R. Page. 2010. Strategies to optimize protein expression in *E. coli*. *Current Protocols in Protein Science*. 61:1-29.
- Gang, D.R., H. Kasahara, Z.Q. Xia, K. Vander Mijnsbrugge, G. Bauw, W. Boerjan, M. Van Montagu, L.B. Davin, and N.G. Lewis. 1999. Evolution of plant defense mechanisms. Relationships of phenylcoumaran benzylic ether reductases to pinoresinol-lariciresinol and isoflavone reductases. *Journal of Biological Chemistry*. 274:7516-7527.
- Gao, H., N.N. Narayanan, L. Ellison, and M.K. Bhattacharyya. 2005. Two classes of highly similar coiled coil-nucleotide binding-leucine rich repeat genes isolated from the *Rps1-k* locus encode *phytophthora* resistance in soybean. *Molecular Plant Microbe Interactions*®. 18:1035-1045.
- García-Calderón, M., C.M. Pérez-Delgado, P. Palove-Balang, M. Betti, and A.J. Márquez. 2020. Flavonoids and isoflavonoids biosynthesis in the model legume *Lotus japonicus*; connections to nitrogen metabolism and photorespiration. *Plants*. 9:774.
- Geldner, N. 2013. The Endodermis. *Annual Review of Plant Biology*. 64:531-558.
- Graham, T.L. 1990. Role of constitutive isoflavone conjugates in the accumulation of glyceollin in soybean infected with *Phytophthora megasperma*. *Molecular Plant Microbe Interactions*. 3:157.
- Graham, T.L. 1991. Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates 1. *Plant Physiology*. 95:594-603.
- Graham, T.L., M.Y. Graham, S. Subramanian, and O. Yu. 2007. RNAi silencing of genes for elicitation or biosynthesis of 5-deoxyisoflavonoids suppresses race-specific resistance and hypersensitive cell death in *Phytophthora sojae* infected tissues. *Plant Physiology*. 144:728-740.
- Graham, T.L., J.E. Kim, and M.Y. Graham. 1990. Role of constitutive isoflavone conjugates in the accumulation of glyceollin in soybean infected with *Phytophthora megasperma*. *Molecular Plant-Microbe Interactions*. 3:157-166.
- Guo, L., R.A. Dixon, and N.L. Paiva. 1994. Conversion of vestitone to medicarpin in alfalfa (*Medicago sativa* L.) is catalyzed by two independent enzymes. Identification, purification, and characterization of vestitone reductase and 7,2'-dihydroxy-4'-methoxyisoflavanol dehydratase. *Journal of Biological Chemistry*. 269:72-78.
- Hardham, A.R. 2007. Cell biology of plant-oomycete interactions. *Cellular Microbiology*. 9:31-39.

- Horton, P. 2006. Protein subcellular localization prediction with WoLF PSORT. *In* Proceedings of Asian Pacific Bioinformatics Conference. 39-48.
- Hu, Q., C. Qu, X. Xiao, W. Zhang, Y. Jiang, Z. Wu, D. Song, X. Peng, X. Ma, and Y. Zhao. 2021. Flavonoids on diabetic nephropathy: advances and therapeutic opportunities. *Chinese Medicine*. 16:74.
- Hua, C., L. Linling, X. Feng, W. Yan, Y. Honghui, W. Conghua, W. Shaobing, L. Zhiqin, H. Juan, W. Yuping, C. Shuiyuan, and C. Fuliang. 2013. Expression patterns of an *isoflavone reductase-like* gene and its possible roles in secondary metabolism in *Ginkgo biloba*. *Plant Cell Reports*. 32:637-650.
- Huang, M., H. Hu, L. Ma, Q. Zhou, L. Yu, and S. Zeng. 2014. Carbon-carbon double-bond reductases in nature. *Drug Metab Rev*. 46:362-378.
- Husain, Q. 2017. High yield immobilization and stabilization of oxidoreductases using magnetic nanosupports and their potential applications: an update. *Current Catalysis*. 6:168-187.
- Hussain, G., L. Zhang, A. Rasul, H. Anwar, M.U. Sohail, A. Razzaq, N. Aziz, A. Shabbir, M. Ali, and T. Sun. 2018. Role of plant-derived flavonoids and their mechanism in attenuation of Alzheimer's and Parkinson's diseases: an update of recent data. *Molecules*. 23:814.
- Hymowitz, T. 1970. On the domestication of the soybean. *Economic Botany*. 24:408-421.
- Hyten, D.L., S.B. Cannon, Q. Song, N. Weeks, E.W. Fickus, R.C. Shoemaker, J.E. Specht, A.D. Farmer, G.D. May, and P.B. Cregan. 2010. High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BioMed Central Genomics*. 11:1-38.
- Iqbal, J., B.A. Abbasi, T. Mahmood, S. Kanwal, B. Ali, S.A. Shah, and A.T. Khalil. 2017. Plant-derived anticancer agents: A green anticancer approach. *Asian Pacific Journal of Tropical Biomedicine*. 7:1129-1150.
- Jahan, M.A., B. Harris, M. Lowery, A.M. Infante, R.J. Percifield, and N. Kovinich. 2020. Glyceollin transcription factor GmMYB29A2 regulates soybean resistance to *Phytophthora sojae*. *Plant Physiology*. 183:530-546.
- Jaudan, A., S. Sharma, S.N.A. Malek, and A. Dixit. 2018. Induction of apoptosis by pinostrobin in human cervical cancer cells: Possible mechanism of action. *Plos One*. 13:1-23.
- Jing, M., B. Guo, H. Li, B. Yang, H. Wang, G. Kong, Y. Zhao, H. Xu, Y. Wang, W. Ye, S. Dong, Y. Qiao, B.M. Tyler, W. Ma, and Y. Wang. 2016. A *Phytophthora sojae* effector suppresses endoplasmic reticulum stress-mediated immunity by

- stabilizing plant Binding immunoglobulin Proteins. *Nature Communications*. 7:11685-11702.
- Jung, W., O. Yu, S.-M.C. Lau, D.P. O'Keefe, J. Odell, G. Fader, and B. McGonigle. 2000. Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nature Biotechnology*. 18:208-212.
- Kasteel, M., T. Ketelaar, and F. Govers. 2023. Fatal attraction: How *Phytophthora* zoospores find their host. *Seminars in Cell and Developmental Biology*. 148-149:13-21.
- Kaur, S., M.K. Samota, M. Choudhary, M. Choudhary, A.K. Pandey, A. Sharma, and J. Thakur. 2022. How do plants defend themselves against pathogens-Biochemical mechanisms and genetic interventions. *Physiology and Molecular Biology of Plants*. 28:485-504.
- Keller, H.E. 1995. Objective lenses for confocal microscopy. In *Handbook of Biological Confocal Microscopy*. J.B. Pawley, editor. Springer United State, Boston, MA. 111-126.
- Khatri, P., O. Wally, I. Rajcan, and S. Dhaubhadel. 2022. Comprehensive Analysis of Cytochrome P450 Monooxygenases Reveals Insight Into Their Role in Partial Resistance Against *Phytophthora sojae* in Soybean. *Frontiers in Plant Science*. 13:862314-862334.
- Kim, H.J., H.-J. Suh, J.H. Kim, S. Park, Y.C. Joo, and J.-S. Kim. 2010a. Antioxidant activity of glyceollins derived from soybean elicited with *Aspergillus sojae*. *Journal of Agricultural and Food Chemistry*. 58:11633-11638.
- Kim, I.-S., C.-H. Kim, and W.-S. Yang. 2021. Physiologically Active Molecules and Functional Properties of Soybeans in Human Health—A Current Perspective. *International Journal of Molecular Sciences*. 22:4054.
- Kim, S.G., S.T. Kim, Y. Wang, S.-K. Kim, C.H. Lee, K.-K. Kim, J.-K. Kim, S.Y. Lee, and K.Y. Kang. 2010b. Overexpression of rice isoflavone *reductase-like* gene (*OsIRL*) confers tolerance to reactive oxygen species. *Physiologia Plantarum*. 138:1-9.
- Kim, S.T., K.S. Cho, S.G. Kim, S.Y. Kang, and K.Y. Kang. 2003. A rice *isoflavone reductase-like* gene, *OsIRL*, is induced by rice blast fungal elicitor. *Molecules and Cells*. 16:224-231.
- Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*. 35:1547-1549.
- Kumar, V.K.V., A.R.A. Rani, and G. Chauhan. 2010. Nutritional value of soybean. Commonwealth Agricultural Bureaux International.

- Li, L., F. Lin, W. Wang, J. Ping, J.C. Fitzgerald, M. Zhao, S. Li, L. Sun, C. Cai, and J. Ma. 2016. Fine mapping and candidate gene analysis of two loci conferring resistance to *Phytophthora sojae* in soybean. *Theoretical and Applied Genetics*. 129:2379-2386.
- Li, S., H. Xu, J. Yang, and T. Zhao. 2019. Dissecting the genetic architecture of seed protein and oil content in soybean from the yangtze and huaihe river valleys using multi-locus genome-wide association studies. *International Journal of Molecular Sciences*. 20:3041.
- Li, X., Y. Han, W. Teng, S. Zhang, K. Yu, V. Poysa, T. Anderson, J. Ding, and W. Li. 2010. Pyramided QTL underlying tolerance to *Phytophthora* root rot in mega-environments from soybean cultivars 'Conrad' and 'Hefeng 25'. *Theoretical and Applied Genetics*. 121:651-658.
- Libault, M., A. Farmer, T. Joshi, K. Takahashi, R.J. Langley, L.D. Franklin, J. He, D. Xu, G. May, and G. Stacey. 2010. An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *The Plant Journal*. 63:86-99.
- Liu, C.-J., D. Huhman, L.W. Sumner, and R.A. Dixon. 2003. Regiospecific hydroxylation of isoflavones by cytochrome P450 81E enzymes from *Medicago truncatula*. *The Plant Journal*. 36:471-484.
- Loon, L.C.v., M. Rep, and C.M.J. Pieterse. 2006. Significance of inducible defense-related proteins in infected plants. *Annual Review of Phytopathology*. 44:135-162.
- Ma, W., L. Yuan, H. Yu, B. Ding, Y. Xi, J. Feng, and R. Xiao. 2010. Genistein as a neuroprotective antioxidant attenuates redox imbalance induced by  $\beta$ -amyloid peptides 25–35 in PC12 cells. *International Journal of Developmental Neuroscience*. 28:289-295.
- Medic, J., C. Atkinson, and C.R. Hurburgh. 2014. Current knowledge in soybean composition. *Journal of the American Oil Chemists' Society*. 91:363-384.
- Messina, M. 2010a. A brief historical overview of the past two decades of soy and isoflavone research. *The Journal of Nutrition*. 140:1350-1354.
- Messina, M. 2010b. Insights gained from 20 years of soy research. *The Journal of Nutrition*. 140:2289-2295.
- Messina, M., and V. Messina. 2010. The role of soy in vegetarian diets. *Nutrients*. 2:855-888.
- Mideros, S., M. Nita, and A.E. Dorrance. 2007. Characterization of components of partial resistance, *Rps2*, and root resistance to *Phytophthora sojae* in soybean. *Phytopathology*. 97:655-662.

- Ng, T.B., X.J. Ye, J.H. Wong, E.F. Fang, Y.S. Chan, W. Pan, X.Y. Ye, S.C.W. Sze, K.Y. Zhang, F. Liu, and H.X. Wang. 2011. Glyceollin, a soybean phytoalexin with medicinal properties. *Applied Microbiology and Biotechnology*. 90:59-68.
- Özgen, F., S. Schmidt, Q. Husain, and M. Ullah. 2019. Biocatalysis: enzymatic basics and applications.
- Paiva, N.L., R. Edwards, Y. Sun, G. Hrazdina, and R.A. Dixon. 1991. Stress responses in alfalfa (*Medicago sativa* L.) 11. Molecular cloning and expression of alfalfa isoflavone reductase, a key enzyme of isoflavonoid phytoalexin biosynthesis. *Plant Molecular Biology* 17:653-667.
- Pei, R., X. Liu, and B. Bolling. 2020. Flavonoids and gut health. *Current Opinion in Biotechnology*. 61:153-159.
- Peiretti, P.G., M. Karamać, M. Janiak, E. Longato, G. Meineri, R. Amarowicz, and F. Gai. 2019. Phenolic composition and antioxidant activities of soybean (*Glycine max* (L.) Merr.) plant during growth cycle. *Agronomy*. 9:153.
- Pieterse, C.M.J., D.V.d. Does, C. Zamioudis, A. Leon-Reyes, and S.C.M.V. Wees. 2012. Hormonal modulation of plant immunity. *Annual Review of Cell and Developmental Biology*. 28:489-521.
- Rípodas, C., V.D. Via, O.M. Aguilar, M.E. Zanetti, and F.A. Blanco. 2013. Knock-down of a member of the *isoflavone reductase* gene family impairs plant growth and nodulation in *Phaseolus vulgaris*. *Plant Physiology and Biochemistry*. 68:81-89.
- Rong, L., H. Chen, Z. Yang, S. Yuan, and X.a. Zhou. 2020. Research status of soybean symbiosis nitrogen fixation. *Oil Crop Science*. 5:6-10.
- Saha, A., and S. Mandal. 2019. Nutritional benefit of soybean and its advancement in research. *Sustainable Food Production*. 5:6-16.
- Sahoo, D.K., N.S. Abeysekara, S.R. Cianzio, A.E. Robertson, and M.K. Bhattacharyya. 2017. A novel *Phytophthora sojae* resistance *Rps12* gene mapped to a genomic region that contains several *Rps* genes. *Public Library of Science One*. 12:1-14.
- Santaniello, E., P. Ferraboschi, P. Grisenti, and A. Manzocchi. 1992. The biocatalytic approach to the preparation of enantiomerically pure chiral building blocks. *Chemical Reviews*. 92:1071-1140.
- Schlieper, D., K. Tiemann, and W. Barz. 1990. Stereospecificity of hydrogen transfer by fungal and plant NADPH:isoflavone oxidoreductases. *Phytochemistry*. 29:1519-1524.
- Schmitthenner, A. 1985. Problems and progress in control of *Phytophthora* root rot of soybean. *Plant Disease*. 69:362-368.



- Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng, D. Xu, U. Hellsten, G.D. May, Y. Yu, T. Sakurai, T. Umezawa, M.K. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, B. Abernathy, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.-C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, and S.A. Jackson. 2010. Genome sequence of the palaeopolyploid soybean. *Nature*. 463:178-183.
- Schopfer, C.R., G. Kochs, F. Lottspeich, and J. Ebel. 1998. Molecular characterization and functional expression of dihydroxypterocarpan 6a-hydroxylase, an enzyme specific for pterocarpanoid phytoalexin biosynthesis in soybean (*Glycine max* L.). *Federation of European Biochemical Societies Letters*. 432:182-186.
- Sepiol, C.J., J. Yu, and S. Dhaubhadel. 2017. Genome-wide identification of *chalcone reductase* gene family in soybean: insight into root-specific *GmCHRs* and *Phytophthora sojae* resistance. *Frontiers in Plant Science*. 8:2073-2088.
- Shoemaker, R.C., J. Schlueter, and J.J. Doyle. 2006. Paleopolyploidy and gene duplication in soybean and other legumes. *Current Opinion in Plant Biology*. 9:104-109.
- Shoji, T., R. Winz, T. Iwase, K. Nakajima, Y. Yamada, and T. Hashimoto. 2002. Expression patterns of two tobacco *isoflavone reductase*-like genes and their possible roles in secondary metabolism in tobacco. *Plant Molecular Biology*. 50:427-440.
- Simons, R., J.-P. Vincken, N. Roidos, T.F.H. Bovee, M. van Iersel, M.A. Verbruggen, and H. Gruppen. 2011. Increasing Soy Isoflavonoid Content and Diversity by Simultaneous Malting and Challenging by a Fungus to Modulate Estrogenicity. *Journal of Agricultural and Food Chemistry*. 59:6748-6758.
- Slusarenko, A.J., R. Fraser, and L.C. van Loon. 2012. Mechanisms of resistance to plant diseases. Springer Science and Business Media, Dordrecht.
- Sparkes, I.A., J. Runions, A. Kearns, and C. Hawes. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols*. 1:2019-2025.
- Subramanian, S., M.Y. Graham, O. Yu, and T.L. Graham. 2005. RNA interference of soybean isoflavone synthase genes leads to silencing in tissues distal to the transformation site and to enhanced susceptibility to *Phytophthora sojae* *Plant Physiology*. 137:1345-1353.
- Subramanian, S., G. Stacey, and O. Yu. 2006. Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. *The Plant Journal*. 48:261-273.

- Sugano, S., T. Sugimoto, H. Takatsuji, and C.-J. Jiang. 2013. Induction of resistance to *Phytophthora sojae* in soybean (*Glycine max*) by salicylic acid and ethylene. *Plant Pathology*. 62:1048-1056.
- Sugimoto, T., M. Aino, M. Sugimoto, and K. Watanabe. 2005. Reduction of *Phytophthora* stem rot disease on soybeans by the application of CaCl<sub>2</sub> and Ca(NO<sub>3</sub>)<sub>2</sub>. *Journal of Phytopathology*. 153:536-543.
- Sugimoto, T., M. Kato, S. Yoshida, I. Matsumoto, T. Kobayashi, A. Kaga, M. Hajika, R. Yamamoto, K. Watanabe, M. Aino, T. Matoh, D.R. Walker, A.R. Biggs, and M. Ishimoto. 2012. Pathogenic diversity of *Phytophthora sojae* and breeding strategies to develop *Phytophthora*-resistant soybeans. *Breed Sci*. 61:511-522.
- Sukumaran, A., T. McDowell, L. Chen, J. Renaud, and S. Dhaubhadel. 2018. Isoflavonoid-specific *prenyltransferase* gene family in soybean: GmPT01, a pterocarpan 2-dimethylallyltransferase involved in glyceollin biosynthesis. *The Plant Journal*. 96:966-981.
- Sun, Y., Q. Wu, H.D. Van Etten, and G. Hrazdina. 1991. Stereoisomerism in plant disease resistance: Induction and isolation of the 7,2'-dihydroxy-4',5'-methylenedioxyisoflavone oxidoreductase, an enzyme introducing chirality during synthesis of isoflavonoid phytoalexins in pea (*Pisum sativum* L). *Archives of Biochemistry and Biophysics*. 284:167-173.
- Taylor, C.K., R.M. Levy, J.C. Elliott, and B.P. Burnett. 2009. The effect of genistein aglycone on cancer and cancer risk: a review of in vitro, preclinical, and clinical studies. *Nutrition Reviews*. 67:398-415.
- Thines, M. 2014. Phylogeny and evolution of plant pathogenic oomycetes—a global overview. *European Journal of Plant Pathology*. 138:431-447.
- Thomas, R., X. Fang, K. Ranathunge, T.R. Anderson, C.A. Peterson, and M.A. Bernards. 2007. Soybean root suberin: anatomical distribution, chemical composition, and relationship to partial resistance to *Phytophthora sojae*. *Plant Physiology*. 144:299-311.
- Tice, J.A., Steven R. Cummings, Rebecca Smith-Bindman, Laura Ichikawa, William E. Barlow, and Karla Kerlikowske. 2008. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Annals of Internal Medicine*. 148:337-347.
- Toone, E.J. 2010. Advances in enzymology and related areas of molecular biology, volume 75: protein evolution. Wiley Interscience, Durham.
- Traore, S.M., and B. Zhao. 2011. A novel Gateway®-compatible binary vector allows direct selection of recombinant clones in *Agrobacterium tumefaciens*. *Plant Methods*. 7:42.

- Tyler, B.M. 2007. *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. *Molecular Plant Pathology*. 8:1-8.
- Veitch, N.C. 2007. Isoflavonoids of the leguminosae. *Natural Product Reports*. 24:417-464.
- Wang, X., X. He, J. Lin, H. Shao, Z. Chang, and R.A. Dixon. 2006a. Crystal structure of isoflavone reductase from alfalfa (*Medicago sativa* L.). *Journal of Molecular Biology*. 358:1341-1352.
- Wang, Y., Y. Li, T. Zhang, Y. Chi, M. Liu, and Y. Liu. 2018. Genistein and Myd88 activate autophagy in high glucose-induced renal podocytes in vitro. *Medical Science Monitor*. 24:4823-4831.
- Wang, Y., W. Zhang, Y. Wang, and X. Zheng. 2006b. Rapid and sensitive detection of *Phytophthora sojae* in soil and infected soybeans by species-specific polymerase chain reaction assays. *Phytopathology*. 96:1315-1321.
- Wang, Z., M. Libault, T. Joshi, B. Valliyodan, H.T. Nguyen, D. Xu, G. Stacey, and J. Cheng. 2010. SoyDB: a knowledge database of soybean transcription factors. *BioMedical Central Plant Biology*. 10:14.
- Welle, R., and H. Grisebach. 1988. Induction of phytoalexin synthesis in soybean: Enzymatic cyclization of prenylated pterocarpan to glyceollin isomers. *Archives of Biochemistry and Biophysics*. 263:191-198.
- Wilkins, M.R., E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, and D.F. Hochstrasser. 1999. Protein identification and analysis tools in the ExPASy server. *Methods Molecular Biology*. 112:531-552.
- Yan, H., and B. Nelson. 2019. Adaptation of *Phytophthora sojae* to *Rps* resistance genes over the past two decades in North Dakota. *Plant Health Progress*. 20:88-93.
- Yoosefzadeh-Najafabadi, M., and I. Rajcan. 2023. Six decades of soybean breeding in Ontario, Canada: a tradition of innovation. *Canadian Journal of Plant Science*. 103:333-352.
- Younus, H. 2019. Oxidoreductases: overview and practical applications. In *Biocatalysis: Enzymatic Basics and Applications*. Q. Husain and M.F. Ullah, editors. Springer International Publishing, Berlin. pp. 39-55.
- Yu, D., H.-S. Shin, Y.S. Lee, D. Lee, S. Kim, and Y.C. Lee. 2014. Genistein attenuates cancer stem cell characteristics in gastric cancer through the downregulation of Gli1. *Oncology Reports*. 31:673-678.
- Yu, O., W. Jung, J. Shi, R.A. Croes, G.M. Fader, B. McGonigle, and J.T. Odell. 2000. Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant Physiology*. 124:781-794.

- Zenbayashi, K., T. Ashizawa, T. Tani, and S. Koizumi. 2002. Mapping of the QTL (quantitative trait locus) conferring partial resistance to leaf blast in rice cultivar Chubu 32. *Theoretical and Applied Genetics*. 104:547-552.
- Zhang, F.F., D.E. Haslam, M.B. Terry, J.A. Knight, I.L. Andrulis, M.B. Daly, S.S. Buys, and E.M. John. 2017. Dietary isoflavone intake and all-cause mortality in breast cancer survivors: The Breast Cancer Family Registry. *Cancer*. 123:2070-2079.

## Curriculum Vitae

**Name:** Negin Azizkhani

**Post-secondary Education and Degrees:**

**The University of Western Ontario**  
London, Ontario, Canada  
2022-2024 (MSc)

**Graduate University of Advanced Technology**  
Kerman, Iran  
2014-2017 (MSc)

**Shahrekord University**  
Shahrekord, Iran  
2010-2014 (BSc)

**Honors and Awards:** **Western Graduate Research Scholarship (WGRS)**  
2022-2023

**Related Work Experience**

**Teaching Assistant**  
The University of Western Ontario  
2022-2023

**Research Assistant**  
Agriculture and Agri-Food Canada, London, Ontario, Canada  
2022-2023

**Publications:**

Mirzaei, S and **Azizkhani**, N. (2019). The value of biodiversity in the legume family and its benefits for human lives. In: biodiversity and conservation characterization and utilization of plants, microbes, and natural resources for sustainable development and ecosystem management, 1st ed.; Jeyabalan, S. Devarajan, T. Goh, H.C. Saher, I., Eds.; Publisher; *Apple Academic Press*, pp. 27- 54.

**Azizkhani**, N., Mirzaei, S., Torkzadeh-Mahani, M. (2021). Genome-wide identification and characterization of legume *T2 Ribonuclease* gene family and analysis of *GmaRNS9*, a soybean *T2 Ribonuclease* gene, function in nodulation. *3 Biotech*, 11, 495-516.

**Scientific Presentation:**

**Azizkhani, N., Dhaubhadel, S., Kohalmi SE. (2023).** Identification and characterization of isoflavone reductase family members in soybean. Biology Graduate Research Forum (BGRF). **[Presentation]**.