
Electronic Thesis and Dissertation Repository

8-24-2015 12:00 AM

Planting the Chalcone Reductase Family Tree: Identification and Characterization of Chalcone Reductase Genes in Soybean

Caroline Julia Sepiol
The University of Western Ontario

Supervisor
Dr. Sangeeta Dhaubhadel
The University of Western Ontario

Graduate Program in Biology
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science
© Caroline Julia Sepiol 2015

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



Part of the [Agricultural Science Commons](#), [Agriculture Commons](#), [Biology Commons](#), [Genetics and Genomics Commons](#), [Molecular Biology Commons](#), [Plant Biology Commons](#), and the [Plant Breeding and Genetics Commons](#)

Recommended Citation

Sepiol, Caroline Julia, "Planting the Chalcone Reductase Family Tree: Identification and Characterization of Chalcone Reductase Genes in Soybean" (2015). *Electronic Thesis and Dissertation Repository*. 3180. <https://ir.lib.uwo.ca/etd/3180>

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.

**PLANTING THE CHALCONE REDUCTASE FAMILY TREE: IDENTIFICATION AND
CHARACTERIZATION OF *CHALCONE REDUCTASE* GENES IN SOYBEAN**

Thesis format: Monograph

by

Caroline Sepiol

Graduate Program in Biology

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

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Caroline Sepiol 2015

Abstract

Soybean (*Glycine max* [L.] Merr) is an important crop grown in Canada, generating \$2.4 billion in sales. Though this number may be promising, soybean farmers lose about \$50 million worth of yield annually due to root and stem rot disease caused by *Phytophthora sojae*. Many strategies have been developed to combat the infection; however, these methods are prohibitively expensive. A 'cost effective' approach to this problem is to select a trait naturally found in soybean that can increase resistance. One such trait is the increased production of root glyceollins. One of the key enzymes exclusively involved in glyceollin synthesis is chalcone reductase (CHR). To identify all *GmCHR* gene family members in the soybean genome, a search was conducted in Phytozome which revealed 16 putative *GmCHR*s. Among these, 7 *GmCHR*s were selected for further study as they contain all active site residues, and are transcribed. All candidate *GmCHR*s localize to both the nucleus and cytoplasm. Amongst the candidate *GmCHR*s, there are four root-specific *GmCHR*s which are induced upon stress. In addition, three QTLs have been found which contain *GmCHR* loci. Overall, these findings support the hypothesis that chalcone reductase is an important component in manipulating glyceollin content and could eventually be used to improve resistance against *P. sojae*.

Keywords:

Soybean, *Phytophthora sojae*, chalcone reductase, resistance, quantitative trait loci, stress

Acknowledgments

The completion of my Master's project would have not been possible without the support of many people. Firstly, I would like to express my gratitude to my supervisor, Dr. Sangeeta Dhaubhadel for her assistance and guidance throughout my thesis project. She has provided me with useful comments on my progress and, she has always encouraged me to plan and conduct experiments independently. Furthermore, I would like thank Dr. Jim Karagiannis, my co-supervisor and my advisory committee members, Dr. Susanne Kohalmi and Dr. Mark Gijzen, for their helpful advice during my graduate studies.

I would like to send my thanks to all the previous and current members of the Dr. Dhaubhadel Lab: Dr. Jaeju Yu, Ling Chen, Dr. Mehran Dastmalchi, Arun Kumaran Anguraj Vadivel, Hemanta Mainali, Kishor Duwadi, Pravesh Lama and Arjun Sukumaran for creating a wonderful working environment and, providing me with critical feedback. Thanks to Chelsea Ishmeal from Dr. Mark Gijzen's lab for dedicating her time to inoculate soybean stems. Additionally thanks to Patrick Chapman and Alex Molnar for analyzing RNA-seq data and constructing the tissue-specific heatmap.

To the rest of the members of Agriculture Canada (*you know who you are*), thank you for your friendships, coffee breaks, and being there to listen whenever I had something to celebrate or complain about.

Last but not least, I am deeply thankful to my parents for encouraging me throughout my entire academic career. Their unconditional love and perspective always made me believe that I could succeed.

List of Abbreviations

*Standard SI units not listed

2,5-DKR	2,5-diketo-D-gluconic acid reductase B
2HDR	2'-hydroxydaidzein reductase
2HID	2-hydroxyisoflavone dehydratase
3,9 DPO	3,9'-dihydroxypterocarpan 6a-monooxygenase
4CL	4-coumarate-Co:A-ligase
AgNO ₃	silver nitrate
AKR	aldo-keto reductase
ALDH	alcohol dehydrogenase
ALR	aldose reductase
ARR	arabinose reductase
<i>att</i>	attachment
bp	base pair
C4H	cinnamate-4-hydroxylase
cDNA	complementary DNA
CDS	coding DNA sequences
CFP	cyan fluorescent protein
CHI	chalcone isomerase
CHR	chalcone reductase
CHS	chalcone synthase
COR	codeinone reductase
cv.	cultivar
eGFP	enhanced green fluorescent protein
G2DT	glycinol 2-dimethylallyltransferase

G4DT	glycinol 4-dimethylallyltransferase
GFP	green fluorescent protein
GmAKR	soybean aldo-keto reductase
GmCHR	soybean chalcone reductase
GMT or UGT73F2	glycosyltransferase
GS	glyceollin synthase
HRP	horseradish peroxidase
I2'H	isoflavone 2'-hydroxylase
IFS	isoflavone synthase
kDa	kilo-Dalton
LB	lysogeny broth
LOD	logarithm of odds
M6DH	morphine 6-dehydrogenase
M6PR	mannose-6-phosphate
MES	2-N-morpholin-ethanesulfonic acid
MT or GmMT7	malonyltransferase
NLS	nuclear localization signal
NTC	no template control
PAL	phenylalanine ammonia lyase
PFS	prostaglandin F synthase
POR	polyketide reductase
PTS	pterocarpan synthase
PVDF	polyvinylidene fluoride
qPCR or qRT-PCR	quantitative PCR
QTL	quantitative trait loci

RNAi	RNA interference
ROS	reactive oxygen species
<i>Rps</i>	resistance to <i>Phytophthora sojae</i>
RT-PCR	reverse transcription-polymerase chain reaction
SDS	sodium dodecyl sulfate
SEM	standard error of the mean
SNP	single nucleotide polymorphism
SSR	simple short repeats
UGT	glycosyltransferase
UTR	Untranslated region
VIGS	virus induced gene silencing
XYR	xylose reductase
YFP	yellow fluorescent protein

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
List of Abbreviations.....	iv
Tables of Contents.....	vii
List of Figures.....	x
List of Tables.....	xiii
List of Appendices.....	xiv
Chapter 1: Introduction.....	1
1.1 Soybean and its economy.....	1
1.2 The culprit: <i>Phytophthora sojae</i> and soybean root and stem rot disease.....	2
1.3 Overview of soybean defense strategies.....	3
1.3.1 Defense strategies.....	3
1.3.2 Complete and partial resistance in soybeans.....	6
1.4 Cultivar development approaches.....	7
1.4.1 Traditional soybean breeding.....	7
1.4.2 “Next generation” breeding.....	7
1.4.3 Quantitative trait loci.....	8

1.4.4	QTLs linked to <i>P. sojae</i> resistance and association studies in soybean.....	9
1.5	Soybean phytoalexins: synthesis of the isoflavonoid glyceollin.....	10
1.6	Chalcone reductase.....	13
1.7	Soybean genome and genomic duplication.....	17
1.8	Hypothesis.....	17
1.9	Objectives.....	18
Chapter 2: Materials and methods		19
2.1	Biological materials.....	19
2.1.1	Plant materials and growth conditions.....	19
2.1.2	Bacterial strains.....	19
2.2	<i>In silico</i> analysis.....	20
2.3	Gene expression analysis	20
2.4	Identification of QTL(s) and QTL markers linked to <i>P. sojae</i> resistance.....	21
2.5	Plasmid construction	21
2.5.1	Cloning into the Gateway entry vector, pDONRZeo.....	21
2.5.2	Cloning into destination vectors	22
2.5.3	DNA sequencing.....	23
2.6	Transient expression of protein in <i>N. benthamiana</i> leaves	26
2.7	Confocal microscopy.....	26
2.8	Protein extraction and western blot analysis	27

2.9	RNA extraction, cDNA synthesis and RT/qRT-PCR	28
2.10	Stress treatment	28
2.11	Statistical analyses.....	29
2.12	Hairy root transformation.....	29
2.12.1	Preparing soybean cotyledons.....	29
2.12.2	Preparing <i>A. rhizogenes</i> cultures.....	29
2.12.3	Transformation of soybean cotyledons	31
2.12.4	Harvesting hairy roots.....	31
Chapter 3: Results		33
3.1	Soybean genome contains 16 putative GmCHRs	33
3.2	<i>GmCHR</i> family members display tissue-specific gene expression.....	39
3.3	GmCHR localizes in the nucleus and cytoplasm in <i>N. benthamiana</i>	45
3.4	Root-specific <i>GmCHRs</i> are induced upon stress	47
3.5	Identification of QTLs and QTL markers linked to <i>P. sojae</i> resistance	48
3.6	Accumulation of root-specific GmCHR is higher in <i>P. sojae</i> resistant soybean cultivar 54	
3.7	RNAi silencing of root-specific GmCHRs	54
Chapter 4: Discussion		60
4.1	Soybean genome contains 16 putative <i>GmCHR</i> genes	60
4.2	<i>GmCHRs</i> show tissue-specific expression patterns	61

4.3	GmCHR localizes in the nucleus and cytoplasm	62
4.4	Root-specific <i>GmCHRs</i> are induced upon stress	63
4.5	QTLs linked to <i>P. sojae</i> resistance contain <i>GmCHR</i> loci	64
4.6	Knockdown of root-specific <i>GmCHRs</i>	65
Chapter 5: Conclusions and future directions		67
References		69
Appendices		84
Curriculum Vitae		99

List of Figures

Figure 1.1. <i>P. sojae</i> and soybean root and stem rot disease.....	4
Figure 1.2. The general phenylpropanoid pathway showing isoflavonoid biosynthesis.	11
Figure 1.3. Evolutionary tree showing representative enzymes of AKR families 1–5.....	14
Figure 1.4. Ribbon diagram of chalcone reductase (<i>Glycine max</i>).....	15
Figure 1.5. Chalcone reductase active site.....	16
Figure 2.1. Gateway vectors for subcellular localization and silencing.	25
Figure 2.2. Set-up of AgNO ₃ stress treatment on soybean hypocotyl.....	30
Figure 2.3. Soybean hairy root transformation.	32
Figure 3.1. Phylogenetic tree showing GmCHRs cluster together with known plant CHRs.	35
Figure 3.2. Identification of important amino acid residues in candidate GmCHRs.....	37
Figure 3.3. Expression profile of <i>GmCHR</i> genes.	40
Figure 3.4. Subcellular localization of the GmCHR family.....	43
Figure 3.5. Accumulation of GmCHR2A-YFP in <i>N. benthamiana</i> leaves.....	46
Figure 3.6. Gene expression of <i>GmCHRs</i> in response to <i>P. sojae</i> infection.	49
Figure 3.7. Effects of AgNO ₃ on etiolated soybean cv. Harosoy 63 hypocotyls.	50
Figure 3.8. Expression of root-specific <i>GmCHRs</i> in response to AgNO ₃ treatment.	51
Figure 3.9. Genomic distribution of <i>GmCHR</i> genes, QTL and QTL markers on soybean chromosomes.	55
Figure 3.10. Accumulation of root-specific <i>GmCHRs</i> in soybean cv. Conrad and OX760-6.....	57
Figure 3.11. Generation of <i>GmCHR</i> silenced hairy roots.....	58
Figure 3.12. Accumulation of root-specific <i>GmCHR</i> in soybean cv. Harosoy 63 hairy root system.....	59

List of Tables

Table 2.1. List of primers used for full-length cloning, gene expression, and gene silencing of <i>GmCHRs</i>	24
Table 3.1. Genetic characteristics of the soybean chalcone reductase gene family	41
Table 3.2. Pairwise coding region and amino acid sequence comparison in the soybean chalcone reductase family.....	42
Table 3.3. Characteristics of QTLs linked to <i>P. sojae</i> resistance which contain <i>GmCHRs</i>	52
Table 3.4. Characteristics of QTL markers linked to <i>P. sojae</i> resistance that flank <i>GmCHR</i>	53

List of Appendices

Appendix A. Complete list of GmCHRs, genetic characteristics and reason of elimination	84
Appendix B. Complete Alignment of GmCHRs, CHRs from other plant species and ARKs	85
Appendix C. Number of reads from RNA-seq data collected from http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29163	87
Appendix D. qPCR primer efficiencies for root-specific gene expression for AgNO ₃ treatment and silencing	88
Appendix E. Complete list of QTLs found in the SoyBase and Soybean Breeder's Toolbox	90
Appendix F. Complete list of QTL markers found in the literature search	93

Chapter 1: Introduction

1.1 Soybean and its economy

Soybean (*Glycine max* [L.] Merr) is one of the most important leguminous crops in the world. Soybean seeds are an excellent source of protein, oils and micronutrients such as calcium and iron, making it an attractive and profitable crop for human consumption. The seeds also contain several bioactive compounds such as saponins, bioactive peptides (Yoshikawa et al., 2000), and isoflavones (Dixon, 2004; Cederroth and Nef, 2009) which are known to have several health benefits. The consumption of these bioactives are heavily studied and have shown to reduce the risk of hormone-dependant cancers, cardiovascular disease, and inhibit the infectivity of HIV. Out of all the bioactives, isoflavonoids accumulate highest in the seeds and these metabolites are often commercialized into supplements. Several studies support that isoflavonoid supplements reduce the risk of breast cancer in women and prostate cancer in males (Gutierrez-Gonzalez et al., 2009; Korde et al., 2009) and alleviate post-menopausal ailments such as hot-flashes (Strom et al., 2001). Soybeans also have emerged into industrial products including soap, cosmetics and waxes. Soybean oil is a major source of biodiesel in the United States aside from corn oil (Fargione et al., 2008; Candeia et al., 2009). In Canada, soybeans are primarily used as cattle feed (Dorff, 2009).

From this versatility and popularity of the soybean, world-wide soybean production has reached 250 million hectares, a 14-fold increase from 50 years ago. The main producers of soybean are the United States (36%), Brazil (36%), Argentina (18%), China (5%) and India (4%) (FAOSTAT, 2015). In Canada, soybean was the 4th largest crop grown in 2014, seeding to 5.5

million acres, producing 6 million tonnes and generating \$2.4 billion in profits (Huston, 2015). Though these numbers appear promising, soybean farmers encounter about \$50 million of soybean yield loss annually in Canada, and \$1-2 billion worldwide due to root and stem rot disease caused by *Phytophthora sojae* (Council, 2012; Murdoch, 2012).

1.2 The culprit: *Phytophthora sojae* and soybean root and stem rot disease

P. sojae is an oomycete and soil-borne plant pathogen that causes soybean root and stem rot disease (Figure 1.1A). The disease is predominant in most soybean growing areas and it is a major contributor to soybean crop loss. *P. sojae* can affect the soybean plant at any stage of development making it difficult to manage infection (Kaufmann and Gerdemann, 1958; Schmitthenner, 1985). Symptoms of infection include: brown lesions in the roots, chlorosis of the leaves and overall plant wilting (Figure 1.1B and C).

P. sojae is closely related to brown algae and is often referred to as water molds (Erwin and Ribeiro, 1996). Nevertheless, earlier studies mistakenly characterized *P. sojae* as a fungus due its morphology (Erwin and Ribeiro, 1996). *P. sojae* produces both mycelium and hyphae which are characteristic of fungi. Yet, several distinct differences exist between fungi and oomycetes such as in cell wall composition, and ploidy level during dominant state, to name a few. In oomycetes, the cell wall consists of beta-glucans while the fungal cell wall contains chitin. In the vegetative state, fungi are haploid whereas oomycete are diploid (Cooke et al., 2000; Latijnhouwers et al., 2003).

P. sojae produces three types of asexual oospores: chlamydospore, sporangia and zoospore, and produces sexual oospores (Tyler, 2007). Any of these oospores can survive for a number of years in a dormant state, and can overwinter in plant debris (Dorrance et al., 2003). Under high soil

moisture levels, the oospores produce zoospores that move in water and attach to soybean roots (Tooley and Grau, 1984; Morris and Ward, 1992). *P. sojae* uses isoflavones such as daidzein and genistein as chemoattractants to locate and navigate to the root. Once attached, the zoospores lose their flagella, and use proteolytic enzymes to breakdown the cell wall of the root. The hypha begins to form in the intercellular space of the cortex to the xylem to extract the nutrients from the plant (Enkerli et al., 1997). Then the pathogen uses the xylem to quickly spread through the entire plant causing the plant to wilt. The cycle is completed when the oospores of the *P. sojae* return to the soil as the plant dies and rots.

Many strategies have been developed to reduce infection such as calcium application, improved soil drainage, fungicides, soil tillage, and seed treatments. These strategies have not only proven to be ineffective, but also place selective pressures on *P. sojae* leading to resistance (Li et al., 2010). An alternative approach to this problem is selecting a cultivar of soybean with an innate resistance to *P. sojae* infection.

1.3 Overview of soybean defense strategies

1.3.1 Defense strategies

The roots of the plant are critical in overall function however, they are also particularly vulnerable to pests and pathogens. These belowground organs are essential in water and nutrient uptake, while anchoring the plant and supporting aboveground organs. The rhizosphere of plants contains both advantageous and pathogenic organisms, as a result, plants have evolved a better defense system to protect themselves (De Coninck et al., 2015).

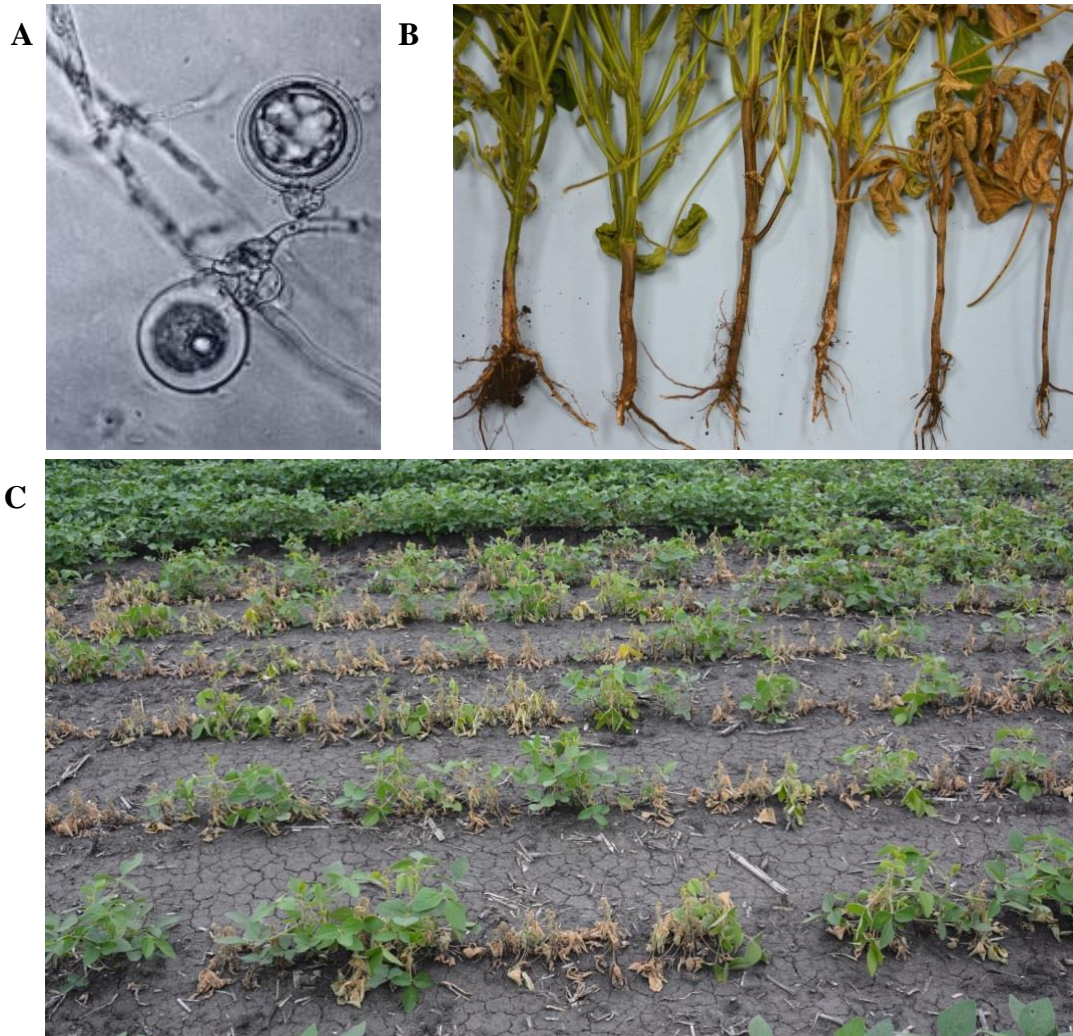


Figure 1.1. *P. sojae* and soybean root and stem rot disease. (A) Oospores of *P. sojae* produced in culture (Adapted from: Schmitthenner, 2001), (B) progression from left to right of soybean root and stem rot disease on soybeans (Adapted from: Byamukama, 2013), and (C) soybean root and stem rot in the field (Adapted from: Byamukama, 2013).

Plants possess both physical and molecular levels of defense against pathogens (Kombrink and Somssich, 1995). At the roots and shoots, there are two protective physical barriers, the exodermis and the endodermis (Geldner, 2013). These two cell layers control osmotic pressure and ions, but also hinder the invasion of microorganisms. Both the layers contain suberin, analiphatic polyester of fatty acids, phenolics and alcohols which prevents pathogen admission into the roots. Studies have shown that higher suberin content in the roots delays the penetration of *P. sojae*, thus reducing the susceptibility to the pathogen in soybean (Ranathunge et al., 2008).

At the molecular level, the defense mainly consists of compounds such as secondary metabolites and protective enzymes (Wink, 1988). A key example of defense at the molecular level is glucosinolates which are produced in the *Brassicaceae* family (Pedras and Hossain, 2011). These enzymes are stored in separate cellular compartments and upon physical damage the cells are disrupted causing the myrosinases to hydrolyze glucosinolates to produce cytotoxic compounds. In the *Fabaceae* family, which soybeans belong to, the production of isoflavonoids is a key component in molecular plant defense. Isoflavonoid production is constitutive; however upon pathogen attack, daidzein, an isoflavonoid, is a substrate to the production of a phytoalexin, glyceollin, which has anti-bacterial and anti-oxidative properties.

Physical and molecular levels of defense can be categorized into two types: inducible or innate (Arnason and Bernards, 2010). Induced defense involves the *de novo* synthesis of secondary metabolites, activation of preformed metabolites and/or strengthening of existing or new protective barriers. Innate defenses consist of physical barriers or constitutively produced secondary metabolites with antimicrobial properties (Arnason and Bernards, 2010). Both innate and induced defenses are combined within the plant and comprise the resistance to a pathogen.

1.3.2 Complete and partial resistance in soybeans

Resistance to pathogen infection in plants can either be complete or partial (Schmitthenner, 1999). Complete resistance or race-specific resistance in soybean is conferred by resistance to *P. sojae* (*Rps*) genes, which counteract the virulence genes within *P. sojae*. This relationship is similar to effector-triggered immune response in other pathosystems. To date, there are fourteen *Rps* genes and more than 55 identified races of *P. sojae* (Gao et al., 2005). However, *P. sojae* continues to diversify and as a result, the classification system has become cumbersome.

All *Rps* genes provide complete and absolute immunity against incompatible races of *P. sojae* and it has been shown that 51% of commercially-available soybean cultivars contain at least one *Rps* gene (Gordon et al., 2006). *Rps1a* was the first resistance gene to be widely used in the USA in 1960s (Gao et al., 2005). Subsequently *Rps1c*, *Rps1k*, *Rps3a* and *Rps6* were deployed in the following years (Dorrance et al., 2003). Out of the fourteen *Rps* genes, *Rps1k* has the strongest resistance against a large number of North American *P. sojae* races and has remained stable. Due to this reason *Rps1k* has been widely commercialized in the last two decades (Gao et al., 2005). Complete resistance puts immense selective pressure on *P. sojae* as such continuous commercialization of the *Rps* genes has resulted in the appearance of new races of *P. sojae* which can overcome the known *Rps* genes.

Due to this problem, many investments have been made towards developing partially resistant cultivars. Partial resistance, referred to as field resistance, is a multi-gene trait which involves various defense components, such as suberin content and isoflavonoid levels, giving the plant an advantage towards fighting the pathogen. Cultivars with this type of resistance contain fewer rotten roots than completely susceptible cultivars, show delayed disease progression, and are

effective against all races of *P. sojae* (Schmitthenner, 1985). A key trait of these strong partial resistance cultivars is the increased production of root isoflavonoids, or more specifically, glyceollins. Further supporting the importance of glyceollins are studies showing that when the isoflavonoid pathway is silenced or compromised, it reduces the plants' ability to fight off the pathogen attack (Subramanian et al., 2005; Graham et al., 2007; Lozovaya et al., 2007). For instance, it was found that the silencing of *CHR* genes not only increased the progression of the *P. sojae* pathogen, but also suppressed hyper-sensitive cell death, which in turn exacerbated the progression of the disease. Further partial resistance alone will not completely protect the crop yield. Additional efforts have been made to breed cultivars with both complete and partial resistance to eradicate the cases of soybean root rot disease.

1.4 Cultivar development approaches

1.4.1 Traditional soybean breeding

The earliest method of breeding novel cultivars was through traditional or conventional plant breeding. To attain the ideal cultivar numerous crosses and back-crosses were performed, proving the inefficiency of traditional breeding. Since the domestication of the soybean about 4,500 years ago, ~400 different soybean lines have been bred via traditional breeding and registered in Canada (CFIA, 2015).

1.4.2 “Next generation” breeding

An efficient and more modern approach of breeding is through identifying the genes that govern ideal traits in soybean. This is addressed through molecular approaches such as positional cloning (Rommens et al., 1989), linkage mapping and insertional mutagenesis (Bechtold et al.,

1993). Such methods, however, are limited by the genome size and lack of transposons (Pflieger et al., 2001). An alternative strategy is using the candidate gene approach. The candidate gene approach relies on the relationship between genetic variation within genes and phenotypes of interest.

There are three steps involved in identifying candidate genes and the association to the trait of interest. Firstly, candidate genes are selected based on their proposed molecular and physiological function. Genes, which are associated with quantitative trait loci (QTL) or involved in a biochemical pathway related to the characteristic of interest, are usually selected as candidate genes. Then, candidate genes are screened for polymorphisms that can potentially alter the expression. Finally, statistical testing is performed to determine the association between the genes and phenotype (Pflieger et al., 2001).

1.4.3 Quantitative trait loci

Quantitative trait loci (QTL) are regions of DNA which often correlate with a variant trait. These regions can span over several megabase pairs and can contain several hundreds to thousands of genes (Dupuis and Siegmund, 1999). The first step of identifying a QTL is by conducting crosses of two parental lines with different traits (resistant vs susceptible). Genetic markers are determined between the parental lines. These genetic markers are the same markers used for the genotyping and include single nucleotide polymorphisms, tandem repeats and restriction enzyme length polymorphisms which don't affect the phenotype of the parental lines. Following the F_1 cross, with the aid the genetic markers, linkage groups or heritable regions of the F_1 progeny are identified. A back-cross is performed with F_1 progeny and parental lines to produce individuals with different fractions of the parental genome. The genotypes and phenotype of the F_2 progeny

are statistically scored. One of the scores is the logarithm of odds (LOD) which estimates the association between the DNA regions to the phenotype. The higher the LOD score the higher the probability of association between the phenotype and DNA region. Other assigned scores might be through ANOVA, t-test and F-statistics (Darvasi, 1998). In one way or another, identifying a QTL is an effective method in finding genes which can contribute to desired trait such as *P. sojae* resistance.

1.4.4 QTLs linked to *P. sojae* resistance and association studies in soybean

Several studies have examined the genes and regions of DNA which govern *P. sojae* resistance. Out of all the studies, genes which are responsible for isoflavonoid synthesis and suberin content, as well as the *Rps* genes have been shown to be of particular importance to *P. sojae* resistance. Iqbal et al. (2005) studied the accumulation of transcripts in partially resistant and susceptible soybeans during *Fusarium* attack. The study describes a generation of 23 recombinant lines from a cross derived from soybean cv. Essex (susceptible) and Forrest (resistant) and identification of 6 QTLs. These QTLs contain genes involved in cell wall and phenolic synthesis.

Han et al. (2008) performed QTL mapping on 112 recombinant inbred lines from a cross between soybean cv. Conrad (resistant) and OX760-6-1 (susceptible) and identified 3 QTLs. Conrad contains high suberin content which has been linked to *P. sojae* resistance (Thomas et al., 2007; Ranathunge et al., 2008).

Moy et al. (2004) studied patterns of gene expression in soybean cv. Harosoy 63 during *P. sojae* infection. A 4,896-gene microarray was constructed to determine transcript accumulation in both host and pathogen during infection. Genes encoding defense and pathogenesis-related proteins,

and genes involved in the phytoalexin glyceollin biosynthesis were found to be upregulated upon infection.

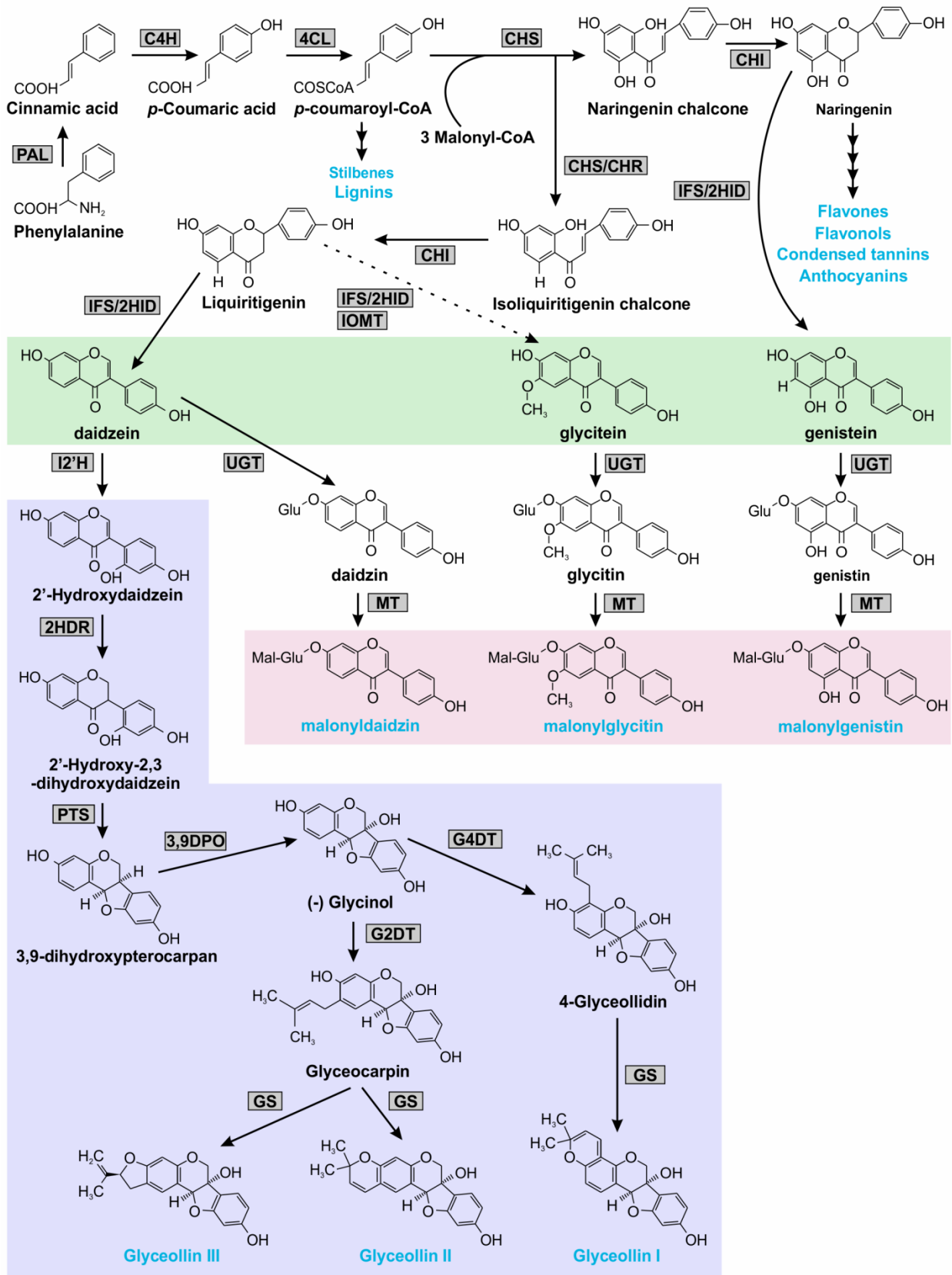
Thus far, there are over 40 QTLs, and over 100 QTL markers that are linked to resistance to *P. sojae* (<http://www.soybase.org/>). The knowledge gained from these studies will aid breeders and scientists in generating the ideal cultivar which is resistant to *P. sojae*.

1.5 Soybean phytoalexins: synthesis of the isoflavonoid glyceollin

In soybeans, phytoalexins are produced via the isoflavonoid pathway (Dakora and Phillips, 1996) which in turn is derived from the general phenylpropanoid pathway. As shown in Figure 1.2, the first committed step is phenylalanine ammonia lyase (PAL) (Habereeder et al., 1989) which eliminates the ammonia group from phenylalanine producing cinnamic acid. The metabolic pathway is further channeled by 4-coumarate CoA ligase (4CL) to produce *p*-coumaroyl-CoA (Li et al., 2014). It is through the activity of either chalcone synthase (CHS) alone or in conjunction with chalcone reductase (CHR) that the pathway branches toward flavonoids and/or isoflavonoid production by joining *p*-coumaroyl-CoA with 3 molecules of malonyl CoA. The first reaction consists of CHS producing naringenin chalcone, a compound that is subsequently converted into the core isoflavone aglycone genistein and other flavonoids with the help of chalcone isomerase (CHI) (Dhaubhadel et al., 2003). The second reaction consists of CHS co-acting with CHR to create isoliquiritigenin chalcone, the building block of the other two core isoflavone aglycones, glycitein and daidzein. The aglycone, daidzein, is the precursor for the phytoalexin glyceollin, an important player in fighting off pathogen attack (Dakora and Phillips, 1996).

Figure 1.2. The general phenylpropanoid pathway showing isoflavonoid biosynthesis. The isoflavonoid pathway produces the three main aglycones, glycitein, daidzein and genistein (boxed in green). Daidzein can either be consecutively converted into malonyldaidzien (pink box) or, upon stress, can be converted into 2'-hydroxydaidzein (induced pathway, purple box) leading into the synthesis of glyceollins. PAL phenylalanine ammonia lyase, C4H cinnamate-4-hydroxylase, 4CL 4-coumarate-CoA-ligase, CHS chalcone synthase, CHR chalcone reductase, CHI chalcone isomerase, IFS isoflavone synthase, 2HID 2-hydroxyisoflavanone dehydratase, UGT glycosyltransferase, MT malonyltransferase, I2'H isoflavone 2'-hydroxylase, 2HDR 2'-hydroxydaidzein reductase, PTS pterocarpan synthase, 3,9 DPO 3,9-dihydroxypterocarpan 6a-monooxygenase, G4DT glycinol 4-dimethylallyltransferase, G2DT glycinol 2-dimethylallyltransferase, GS glyceollin synthase. Gene of study is boxed in black.

Adapted from: Anguraj Vadivel et al. (2015).



1.6 Chalcone reductase

The enzyme CHR is critical for daidzein biosynthesis which ultimately leads to the production of the phytoalexin glyceollin in soybean. CHRs (formally known as polyketide reductase and NAD(P)H dependant 6'-deoxychalcone synthase) belong to the aldo-keto reductase family 4 in the aldo-keto reductase superfamily (Figure 1.3). All members of this superfamily fold into a monomeric, $(\alpha/\beta)_8$ barrel structure (Figure 1.4) (Jez et al., 1997; Jez et al., 1997). The aldo-keto reductase superfamily is divided into 15 sub-families; however, all AKRs contains a common NAD(P)(H) binding site that is located in a deep, large and hydrophobic pocket at the C-terminus end, as well as a catalytic tetrad of Asp-53, Tyr-58, Lys-87, and His-120 (Bomati et al., 2005). As shown in Figure 1.5, CHRs predominantly contain hydrophobic and aromatic residues that line the unoccupied entrance to the active site cavity molded by Pro-29, Ala-57, Trp-89, Phe-130, and Phe-132. Largely polar residues define the base of this catalytic surface and include Asp-53, Tyr-58, Lys-87, His-120, Trp-121, and Asn-167. Beside the facts described above, very little information is known about CHR enzyme since it acts on intermediates for CHS (Bomati et al., 2005).

CHR activity was first demonstrated in crude extracts of *Glycyrrhiza enchinata* (Ayabe et al., 1988). Up to now, CHR-like enzymes have been identified in a variety of leguminous plant species, including *Medicago sativa* (Ballance and Dixon, 1995) *Sesbania rostrata* (Goormachtig et al., 1999), *Pueraria montana* var. *lobata* (He et al., 2011), *Glycyrrhiza glabra* (Hayashi et al., 2013), and *Lotus japonicus* (Shimada et al., 2006).

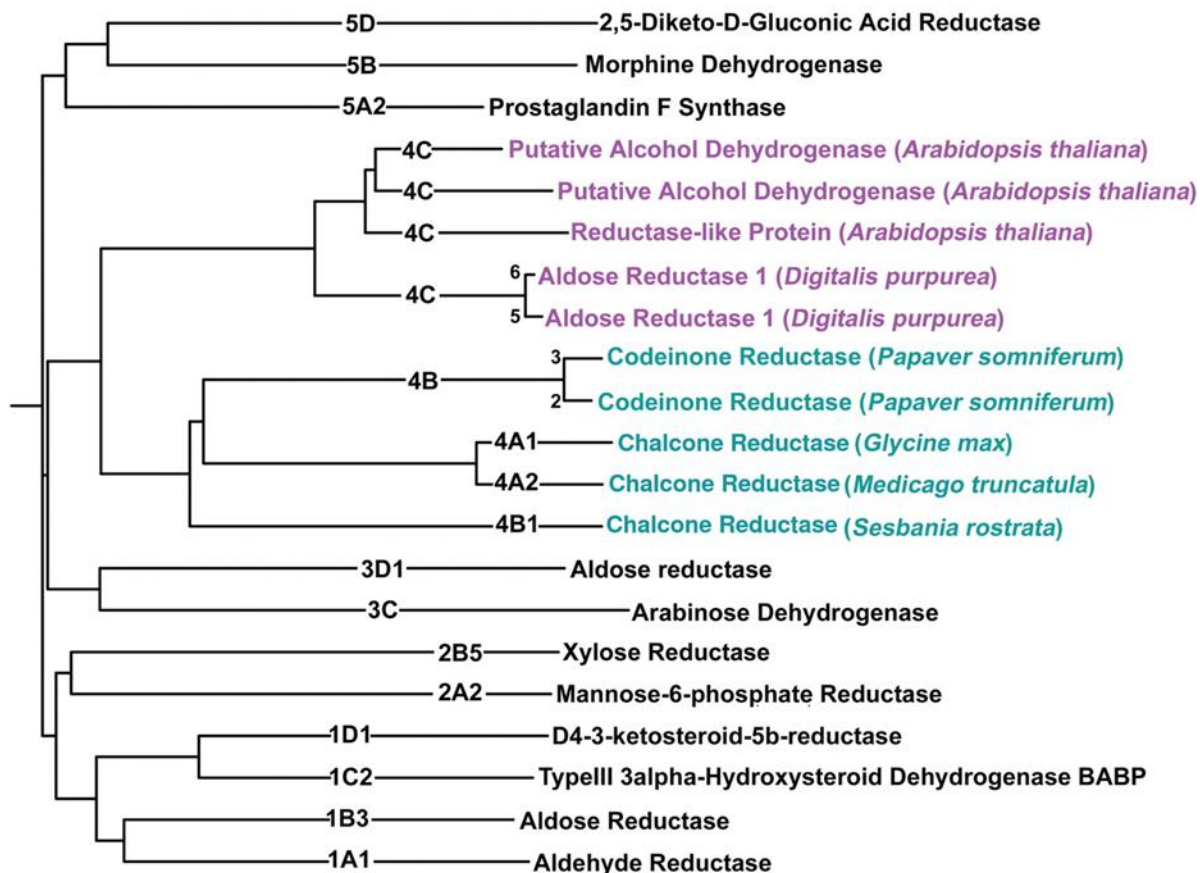


Figure 1.3. Evolutionary tree showing representative enzymes of AKR families 1–5. Purple and green represents enzymes of the AKR family 4. AKR4 family enzymes of putative functions are highlighted in purple while chalcone reductase and codeinone reductase is highlighted in green (Adapted from: Bomati et al., 2005).

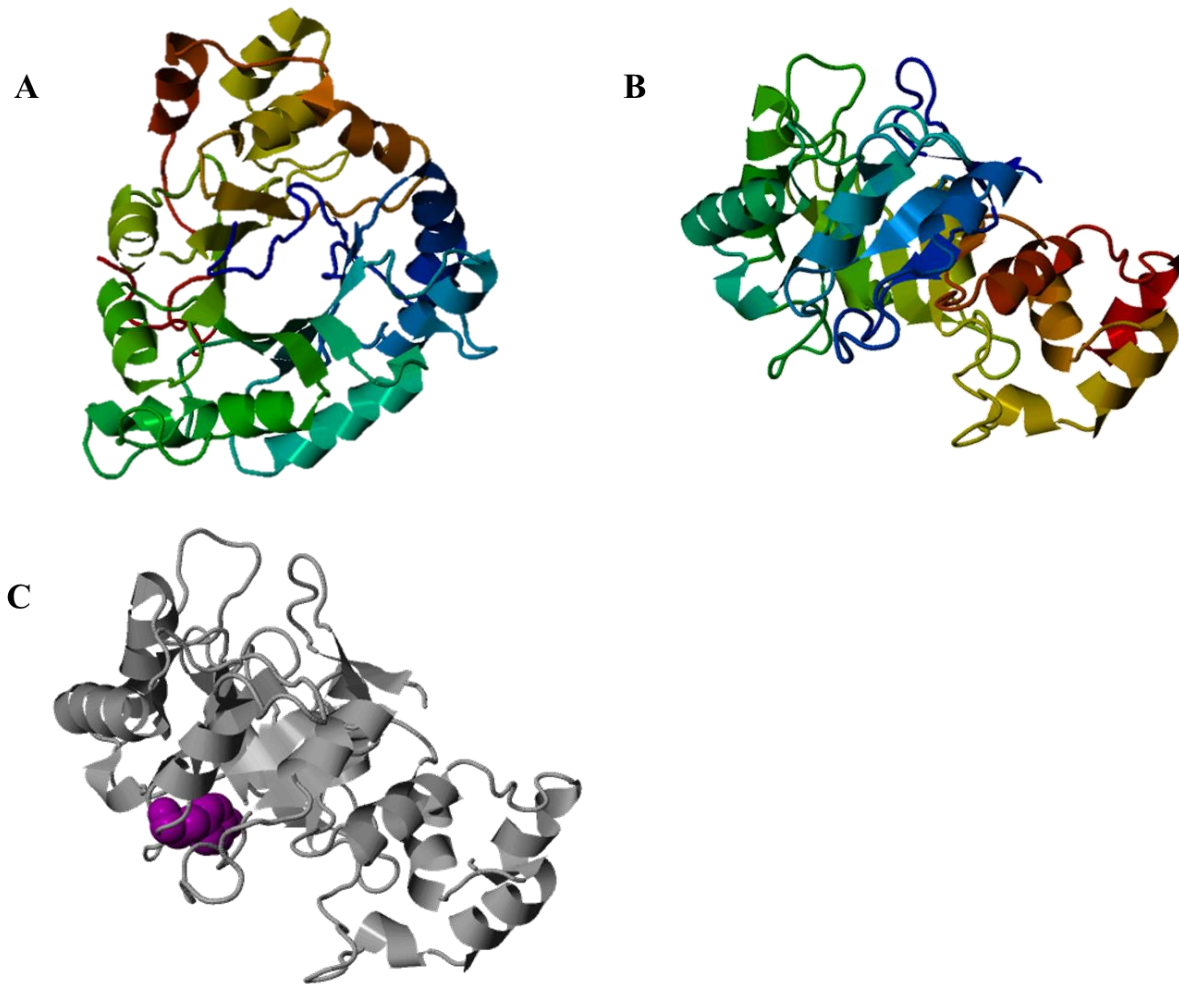


Figure 1.4. Ribbon diagram of chalcone reductase (*Glycine max*). Using Protein Model Portal (<http://www.proteinmodelportal.org/>), a putative representation of chalcone reductase was generated. (A) Top view, (B) side view and the (C) active site of chalcone reductase with substrate (purple) is shown.

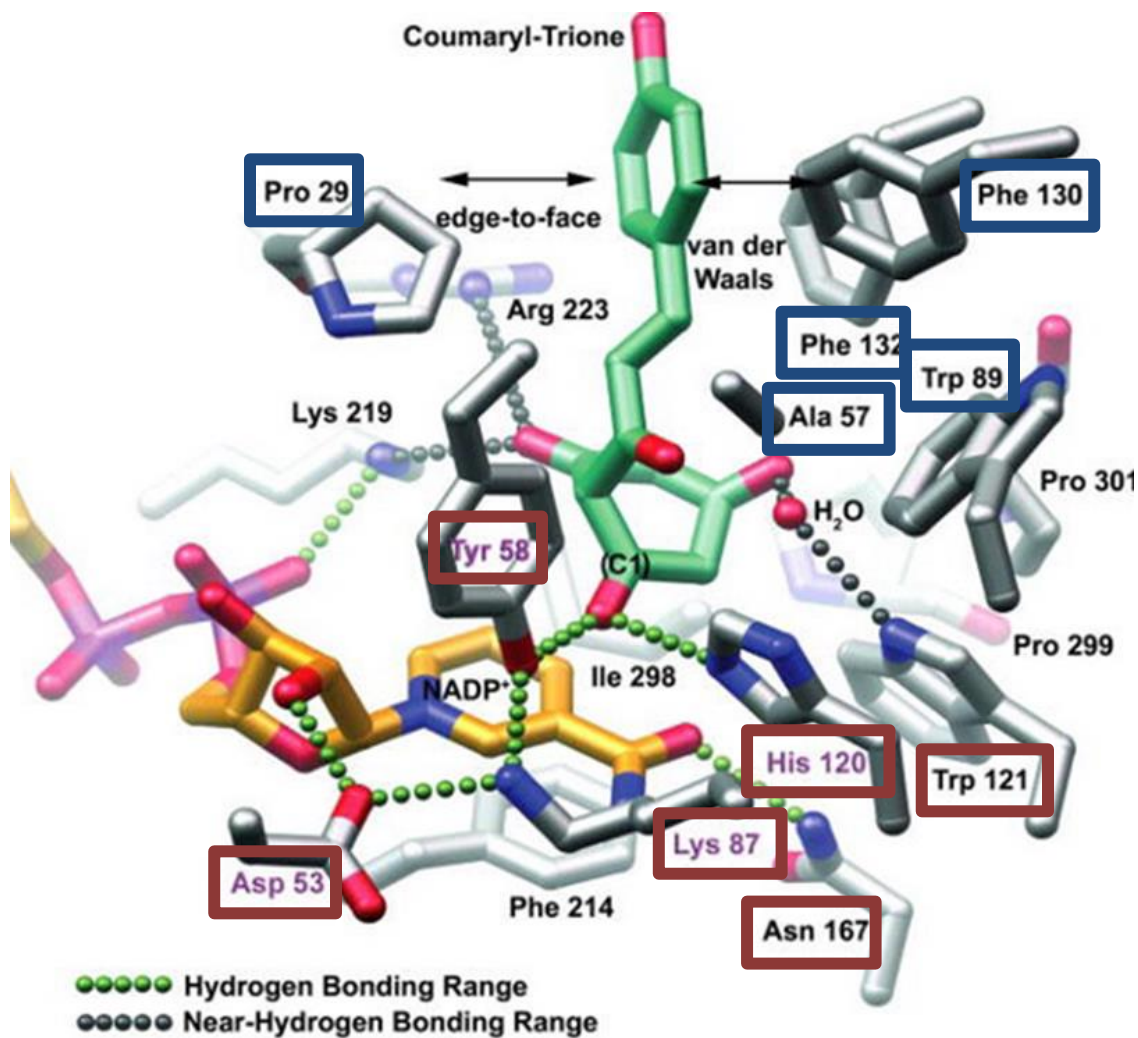


Figure 1.5. Chalcone reductase active site. Catalytic residues, Asp-53, Tyr-58, Lys-87, His-120, Trp-121 and Asn-167 are shown in red boxes. Entrance residues, by Pro-29, Ala-57, Trp-89, Phe-130 and Phe-132 are shown in blue boxes. NADP⁺ is shown in gold, and the proposed substrate is in green. Proposed hydrogen bonding patterns are shown in small green and black circles (Adapted from: Bomati et al., 2005).

Graham et al. (2007) identified 4 soybean *CHRs* through the EST database search (<http://compbio.dfci.harvard.edu/tgi/>) and silenced all of them in soybean hairy roots. The silenced roots produced reduced levels of isoflavonoids, and in turn, were susceptible to *P. sojae* infection.

1.7 Soybean genome and genomic duplication

The soybean genome project was first privately initiated through the Department of Energy Joint Genome Institute Community Sequencing Program in 2006. Sequencing was completed in early 2008 and released in Phytozome, a comparative platform for plant genomics, however; it wasn't until 2010 that the sequencing results were published (Schmutz et al., 2010).

The soybean genome contains 978 megabase pairs with 56,044 protein-coding loci and 88,647 transcripts (<http://phytozome.jgi.doe.gov/pz/portal.html>) which is 70% more genes than in *Arabidopsis thaliana* (Schmutz et al., 2010). Soybean is a palaeopolyploid whose genome experienced two whole duplication events approximately 59 and 13 million years ago. Because of these two duplication events, 75% of genes in soybean have multiple copies. In the context of the genes involved in the isoflavonoid pathway, it is found that *CHS* has 9 family members (Akada and Dube, 1995; Dhaubhadel et al., 2007), *isoflavone synthase (IFS)* has 2 members (Jung et al., 2000) and *CHI* has 12 members (Dastmalchi and Dhaubhadel, 2015).

1.8 Hypothesis

It is hypothesized that there are multiple *CHR* genes (*GmCHR*) present in the soybean genome, and that there are specific members of the *GmCHR*(s) family that have a role in root-specific phytoalexin production.

1.9 Objectives

Thus the objectives of the present research are:

- a) To identify the members of *GmCHR* gene family in soybean.
- b) To determine the subcellular localization of all GmCHRs.
- c) To determine the root-specific *GmCHR*(s) and inducible *GmCHR*(s) upon stress.
- d) To identify QTL(s) and QTL markers linked to *P. sojae* resistance that contain *GmCHR* loci.
- e) To functionally characterize root-specific *GmCHR* family member(s) by RNAi using the soybean hairy root system.

Chapter 2: Materials and methods

2.1 Biological materials

2.1.1 Plant materials and growth conditions

Seeds of *Nicotiana benthamiana* were sprinkled onto wet PRO-MIX BX MYCORRHIZAETM soilTM (Rivière-du-Loup, Canada) in a small tray (5''W x 7''L x 2''D). The tray was placed in a growth room set on a 16h light/8h dark cycle at 23°C/18°C, respectively, with 60%-70% relative humidity and light intensity of 100-150 $\mu\text{mol m}^{-2} \text{s}^{-1}$. After a week, individual seedlings were transferred into sterilized pots and watered regularly. The nutrient mixture of nitrogen, phosphorous, and potassium (20-20-20) was applied once a week.

Seeds of soybean cv. OX760-6, Harosoy 63 and Conrad were sterilized with 70% ethanol containing 3% H₂O₂ for 1 min and rinsed with water at least 6 times prior to planting in soil. The pots were placed in a growth room set on a 16h light/8h dark cycle at 23°C /18°C, respectively.

2.1.2 Bacterial strains

Escherichia coli DH5 α and *Agrobacterium tumefaciens* GV3101 were used for cloning and transient expression in tobacco, respectively. For hairy root transformation *Agrobacterium rhizogenes* K599 was used. For all bacterial transformations, electroporation was carried out in a Gene Pulser® Cuvette (BioRad Laboratories) with 0.1 cm electrode gap using MicroPulserTM (BioRad Laboratories). The electroporation setting used for *E. coli* was 1.80 kV and for *Agrobacterium* sp. was 2.18 kV for 5-6 milliseconds.

2.2 *In silico* analysis

To identify *CHR* gene family members in the soybean genome, a search was conducted in the annotated *G. max* Wm82.a2.v1 genome of Phytozome (<http://phytozome.jgi.doe.gov/pz/portal.html>) (Goodstein et al., 2012). The keywords “aldo-keto” and “aldo/keto” were used to find all the soybean aldo-keto reductases (GmAKRs). To ensure no GmAKRs were missed in the keyword search, each GmAKR was used as a query for a protein BLAST search again.

For generating a phylogenetic tree, protein sequences were aligned using CLUSTALW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) and a Neighbor-joining tree based with 1000 bootstrap replications was created using MEGA6 (Tamura et al., 2013). The Poisson method was selected to calculate the evolutionary distance of the phylogenetic tree and pairwise deletion was selected for gaps/missing data treatment. A candidate gene list was prepared where all the GmCHRs clustered together with other known CHRs on the phylogenetic tree. To determine whether all candidate GmCHRs contain residues deemed important for catalytic activity, the protein sequences of the candidate GmCHR were aligned using CLUSTALW followed by BOXSHADE 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Critical residues were manually spotted based on Bomati et al. (2005). The subcellular localization of candidate GmCHRs were predicted using WoLF-PSORT (<http://wolfpsort.org>), and the chromosome location and CDS information were obtained from Phytozome (<http://phytozome.jgi.doe.gov/pz/portal.html>). The molecular weight of candidate GmCHRs was calculated using ProtParam software (<http://web.expasy.org/protparam>).

2.3 Gene expression analysis

A publicly available RNA-seq database containing transcriptome sequencing of soybean (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29163>) was mined for the expression profiles of 11 *GmCHR* gene family members. The relative expression was normalized across the libraries corresponding to each tissue. A heatmap for *GmCHR* transcripts was generated in R.

2.4 Identification of QTL(s) and QTL markers linked to *P. sojae* resistance

The QTLs and QTL markers from the year 2003 to 2014 corresponding to *P. sojae* resistance were mined from the SoyBase and Soybean Breeder's Toolbox (<http://soybase.org/>). To ensure no QTLs or QTL marker were missed in the search a literature search was also conducted. Relative position of transcribed *GmCHR*s, and QTL markers were mapped onto the chromosomes. QTLs regions in base pairs were noted from the *G. max* genome assembly on Soybase.org.

2.5 Plasmid construction

2.5.1 Cloning into the Gateway entry vector, pDONRZeo

All the primers for cloning purposes contained the *attB1* adaptor sequence (5'-GGGG ACA AGT TTG TAC AAA AAA GCA GGC T-3' for forward primers) and the *attB2* adaptor sequence (5'-GGGG AC CAC TTT GTA CAA GAA AGC TGG GT-3' for reverse primers) for Gateway cloning. To clone *GmCHR* gene family members, gene-specific primers were designed (Table 2.1) and used in a PCR reaction using cDNA synthesized from RNA isolated from different soybean tissues. PCR products were run on a 1% agarose gel and stained with RedSafe (iNtROn Biotechnology). The gels were visualized on a Bio-Rad Gel Doc. PCR products were gel purified using EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc.) and recombined

into pDONRZeo vector (Invitrogen) using BP clonase enzyme (Invitrogen). The recombinant plasmids were then transformed into *E. coli* DH5 α and plated on lysogeny broth (LB) agar plates containing zeocin (50 $\mu\text{g}/\text{mL}$). Positive colonies were screened by colony PCR using gene-specific primers. Positive colonies that contained the right size amplicons were 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.) and sequence verified.

Otherwise stated, all PCR amplifications consisted of a denaturation step at 95°C for 1 minute, annealing step for 30 seconds, extension step at 72°C for 1 minute, with a total of 35 cycles. Before the start of the amplification, there was an the initial denaturation step of 95°C for 5 minutes and after the last cycle, a 5 minute extension was carried out at 67°C.

2.5.2 Cloning into destination vectors

For subcellular localization, the sequence confirmed pDONRZeo-GmCHR plasmids were recombined into pEarleyGate101 (Invitrogen) using LR clonase (Invitrogen), transformed into *E. coli* DH5 α and plated on LB agar plates containing kanamycin (50 $\mu\text{g}/\text{mL}$). *E. coli* colonies that contained pEarleyGate101-GmCHRs plasmids were screened by colony PCR using gene-specific primers (Table 2.1). The positive colonies were selected and grown overnight in 3 mL of LB with kanamycin (50 $\mu\text{g}/\text{mL}$) at 37°C. Plasmid DNA was extracted from the overnight culture using the EZ-10 Spin Column Plasmid DNA Kit (Bio Basic Canada Inc) and transformed into *A. tumefaciens* GV3101 using electroporation. The transformed *Agrobacterium* was grown on LB agar plates containing rifampicin (10 $\mu\text{g}/\text{mL}$), gentamycin (50 $\mu\text{g}/\text{mL}$), and kanamycin (50 $\mu\text{g}/\text{mL}$). Finally, positive colonies were verified by colony PCR. This cloning effort resulted in vectors that produced a translational fusion of each GmCHR with yellow fluorescent protein (YFP).

For RNAi silencing of *GmCHRs*, gene-specific primers (Table 2.1) were designed to clone a 5'end of a *GmCHR* transcript for the targeted silencing of a specific gene family member. The fragments were cloned into pDONRZeo as described in the section 2.3.1, and recombined into the silencing vector, pK7GWIWG2D(II), using LR clonase. This silencing vector contains a separate cassette that produces green fluorescent protein (GFP), as the selectable marker. The recombinant plasmids were transformed into *A. rhizogenes* K599 and plated onto spectinomycin plate (50 µg/mL). The destination vector maps used in my study are shown in Figure 2.1.

2.5.3 DNA sequencing

The DNA constructs were sequenced either at the Robarts Research Institute (Western University) or at the Southern Crop Protection and Food Research Center, Agriculture and Agri-Food Canada, London, Ontario.

Table 2.1. List of primers used for full-length cloning, gene expression, and gene silencing of *GmCHR*s.

Gene	Primer name	Sequence (5'-3')	Amplicon size	Purpose for:	
<i>GmCHR2A</i>	GmCHR2AF	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GGT TAT AGT AIT TGT AAT ACA ATG GCG GCT	967	Subcellular localization and RT-PCR	
	GmCHR2AR	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TAT TTC ATC ATC CCA GAG ATC AGT GAC			
<i>GmCHR9C</i>	GmCHR9CF	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC ATG GAA CCA AAG GCA ATC CAA	945		
	GmCHR9CR	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC ATC CCA AAA ATC TTG TGG AGT TTT GT			
<i>GmCHR14</i>	GmCHR14F	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC CTT GTC AAC CCT TTG AGA GTT AGA ATG	1002		
	GmCHR14R	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TAT TTG ATC ATC CCA GAG ATC AGC			
<i>GmCHR16A</i>	GmCHR16AF	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC TCA GCA GAC CCA AAA GAA GAT ATG	981		
	GmCHR16AR	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC ATC CCA AAG CTC TTC CAA GG			
<i>GmCHR16B</i>	GmCHR16BF	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC AGT TAG AAA CAA GGA GAA AGA AAG AAG ATG	993		
	GmCHR16BR	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC AAC GTC TCC ATC CCA AAG TTC			
<i>GmCHR18</i>	GmCHR18F	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC ATG GCT GCC ACC ACC TTA GT	945		
	GmCHR18R	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TTC TTC ATC CCA TAG ATC AAG AAG AGG			
<i>GmCHR20</i>	GmCHR20F	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GAC ACA TCG CAA CAT CGT ATC ATG	984		
	GmCHR20R	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC AGT TTC CCC ATC AAA TAG ATC AGG			
<i>GmCHR2A</i>	qGmCHR2AF	CGG GGA GGT GCT TCT TGG TTA TAG	150	qPCR	
	qGmCHR2AR	CTC CTT AGT GTC TTT CTT GCA			
<i>GmCHR14</i>	qGmCHR14F	CCA CCT ACC TCC AAT TGC TGA C	167		
	qGmCHR14R	GGG AIT TCA ATA GCA GCA GCC			
<i>GmCHR18</i>	qGmCHR18F	ATC CCA CCT ACC CAC ATC GT	175		
	qGmCHR18R	GGA GGA GGG GAG AGT GAC TT			
<i>GmCHR20</i>	qGmCHR20F	CAC TCG TAT CTG TGA CAC CGT GT	201		
	qGmCHR20R	GTT GTA ACA GCA CTT TGG GGA CGT			
<i>GmCHR2A</i>	siCHR2AF	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC CCG GGA GGT GCT TCT TGG TTA TAG	150		Silencing
	siCHR2AR	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC CTC CTT AGT GTC TTT CTT GCA TGT GAA GTC			
<i>GmCHR14</i>	siCHR14F	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GTA TAA CCC ACC TAC CTC CAA TTG CTG	167		
	siCHR14R	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC GGG AIT TCA ATA GCA GCA GCC AIT C			
<i>GmCHR18</i>	siCHR18F	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GAC TAC TAA TCC CAC CTA CCC ACA TCG	175		
	siCHR18R	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC GCG GCA TCT TCC GCT GG			
<i>GmCHR20</i>	siCHR20F	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC CAC TCG TAT CTG TGA CAC CGT GT	201		
	siCHR20R	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC GTT GTA ACA GCA CTT TGG GGA CGT			

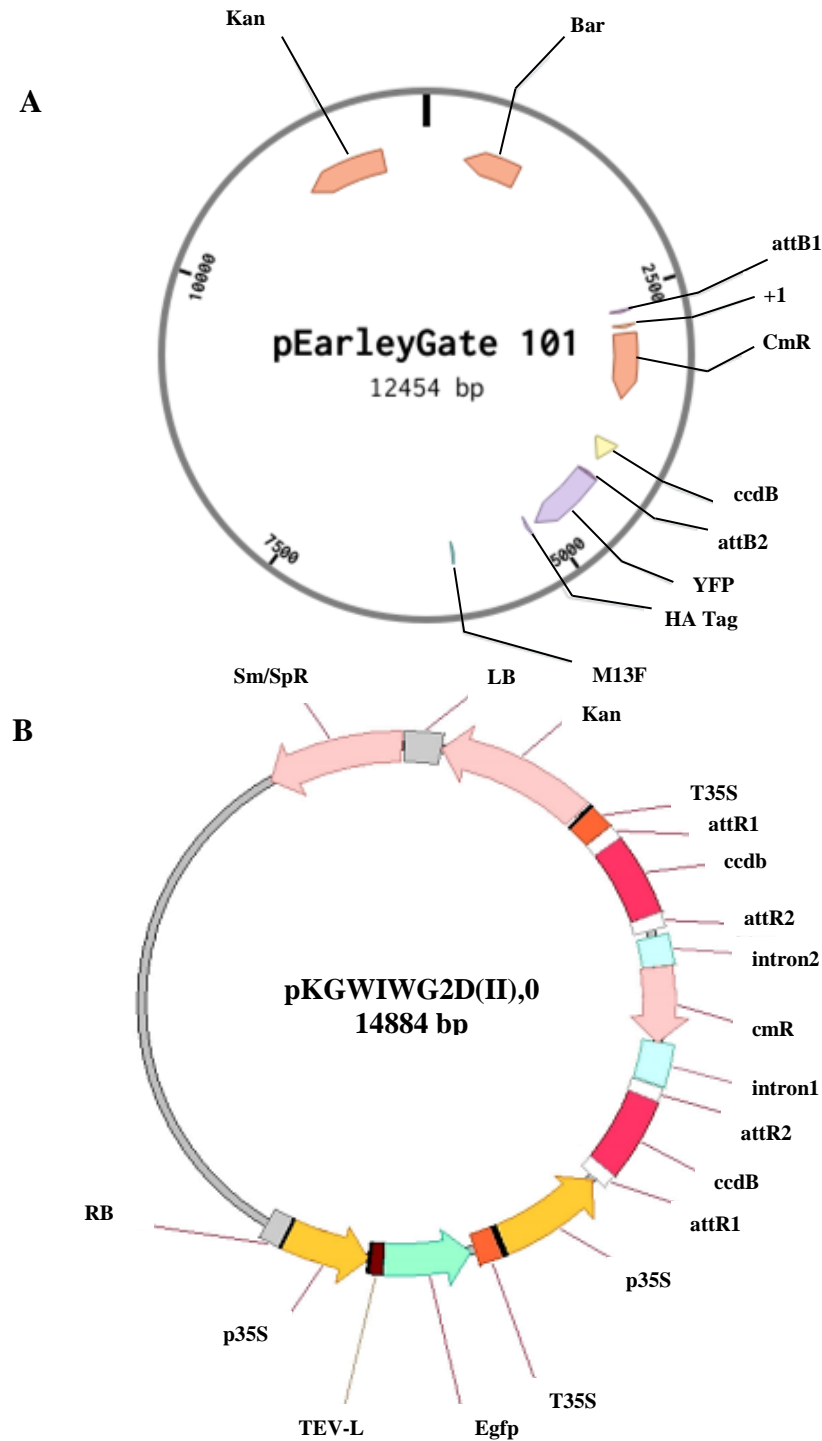


Figure 2.1. Gateway vectors for subcellular localization and silencing. (A) Subcellular localization vector pEarleyGate101 and (B) RNAi silencing vector, pK7GWIWG2D(II). Vector adapted from: <https://benchthumb.s3.amazonaws.com/snapshot/45V8MkrCip.png> and http://www.uoguelph.ca/~jcolasan/pdfs/gateway_protocols_and_plasmids.pdf

2.6 Transient expression of protein in *N. benthamiana* leaves

The subcellular localization of GmCHRs was studied by infiltrating *A. tumefaciens* GV3101 carrying pEarleyGate101-GmCHRs in *N. benthamiana* leaves (Sparkes et al., 2006). A single colony of *A. tumefaciens* GV3101 was inoculated into infiltration culture medium (LB broth containing 10 mM 2-N-morpholino-ethanesulfonic acid (MES) pH 5.6, and 100 μ M acetosyringone) supplemented with kanamycin (50 μ g/mL), rifampicin (10 μ g/mL), and gentamycin (50 μ g/mL) and grown at 28°C until the OD₆₀₀ reached 0.5-0.8. The culture was centrifuged in a microfuge tube at 3000 rpm for 30 minutes at room temperature. Then the pellet was resuspended in Gamborg's solution (3.2 g/L Gamborg's B5 and vitamins, 20 g/L sucrose, 10 mM MES pH5.6, and 200 μ M acetosyringone) to a final OD₆₀₀ of 1 and incubated at room temperature for 1 hour with gentle agitation. The leaves of 4-6 weeks old *N. benthamiana* were infiltrated by placing the tip of the syringe against underside of the leaf. For co-infiltration, the *A. tumefaciens* GV3101 containing the pEarleyGate100 with a nuclear localization signal and cyan fluorescent protein (pEarleyGate100-NLS-CFP) and pEarleyGate101-GmCHR constructs were mixed in equal volumes and then infiltrated into the leaves. The plants were returned to the growth room at normal growth condition as described in section 2.1.1. The protein expression was visualized by confocal microscopy.

2.7 Confocal microscopy

Epidermal cell layers of *N. benthamiana* leaves were visualized using Leica TCS SP2 inverted confocal microscope. For YFP visualization, an excitation wavelength of 514 nm was used and emissions were collected between 525-545 nm. For visualization of CFP, an excitation wavelength of 434 nm was used and emissions were collected between 460-490 nm. For co-

localization, 'Sequential Scan Tool' was utilized which records the image in sequential order instead of acquiring them simultaneously in different channels.

2.8 Protein extraction and western blot analysis

Leaf samples of *N. benthamiana* were used to extract total soluble proteins. Fresh leaves (0.5 g) were ground in liquid nitrogen and re-suspended in extraction buffer (25 mM Tris-HCL pH 8.0, 1 mM EDTA pH 8.0, 20 mM NaCl with Protease Inhibitor Cocktail (Sigma Aldrich). The samples were centrifuged for 30 min at 13,000 rpm at 4°C and the supernatant was collected. The concentration of total soluble protein was determined by a Bradford Assay (Bradford, 1976).

Thirty micrograms of total soluble proteins were loaded on a 7.5% sodium dodecyl sulfate (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 Semi-Dry Electrophoretic Transfer Cell (Bio-Rad) at 20 V for 30 minutes. The membrane was washed with TBS+0.1% Tween 20 three times for 15 minutes followed by blocking in TBS+1% BSA and 0.1% Tween 20 at 4°C overnight. The fusion proteins with YFP were detected using an anti-GFP (1:5000 dilution) mouse primary antibody and conjugated horseradish peroxidase (HRP) goat anti-mouse (1:5000 dilution) secondary antibody. The bound immune-complexes were detected with ECL Prime Western Blot detection reagents (GE Health Care Life Sciences) and exposed in the MicroChemi (DNR Bio-Imaging System).

2.9 RNA extraction, cDNA synthesis and RT/qRT-PCR

Total RNA was extracted from 100 mg of tissue using the RNeasy Plant Mini Kit (Qiagen) following manufacturer's instruction with some modification. An on-column DNaseI (Promega) treatment was used to digest DNA. Subsequently, RNA was quantified with a NanoDrop 1000 spectrophotometer (ThermoScientific), and its quality was evaluated from its A_{260}/A_{280} ratio and by electrophoresing the RNA on a 1% agarose gel in 1X TAE buffer (40 mM Tris, 40 mM acetate, 1 mM EDTA, pH 8.2). Thermoscript RT-PCR System (Life Technologies) was used to synthesize cDNA from 1 μ g of total RNA in a 20 μ L reaction. For RT-PCR, 1 μ L of undiluted RT reaction was used as a template. PCR amplification was performed using gene-specific primers (Table 2.1) using the condition described in the section 2.5.1. For qRT-PCR, the RT reaction (cDNA) template was diluted 20 times and reaction was performed using gene-specific primers (Table 2.1) and SsoFast EvaGreen Supermix Kit (BioRad) in a Bio-Rad C1000 Thermal Cycler with the CFX96™ Real-Time PCR System. *CON4* was used as a reference gene to normalize the expression.

2.10 Stress treatment

Stems of seven-day old seedlings of soybean cv. L76-1988 were inoculated with *P. sojae* race 7. The stems of the infected plants were collected at 24, 48 and 72 hours post-inoculation.

For the AgNO₃ treatment, soybean cv. Harosoy 63 was grown in water-soaked vermiculite in the dark at 25°C for 6 days. Prior to the treatment, 10 etiolated seedlings per treatment were transferred into glass trays, after which, 5-10 drops of 10 μ L of either water (control) or 1 mM of AgNO₃ were placed onto the hypocotyl of each seedling (Figure 2.2). The trays were transferred

back in the dark to 25°C. Samples were collected at 6, 12, 24, 48 and 72 hours after the treatment.

2.11 Statistical analyses

Statistical analyses were performed using GraphPad Prism 6.0. Significant differences between means in gene expression were calculated using multiple t-test per row. Statistical significance was set at $p < 0.05$.

2.12 Hairy root transformation

2.12.1 Preparing soybean cotyledons

Seeds of soybean cv. Harosoy 63 were planted in vermiculite. On day 3 or 4, the soybean plants were sprayed with water to remove the seed cover and vermiculite which were attached to the cotyledons. Hairy root transformation was performed on 6 day old seedlings.

2.12.2 Preparing *A. rhizogenes* cultures

The day before hairy root transformation, *A. rhizogenes* K599 with either no vector, empty silencing vector, or a vector containing pK7GWIWG2D(II)-GmCHR were inoculated in 5 mL of LB broth (no vector) or LB broth with spectinomycin (50 µg/mL) and were grown overnight to an OD₆₀₀ of 0.4 to 1.0 at 28°C. The culture was centrifuged at 2500 g for 20 min at 4°C and, re-suspended to a final OD₆₀₀ of 0.5 in ice cold 10 mM MgSO₄ (Kereszt et al., 2007).



Figure 2.2. Set-up of AgNO_3 stress treatment on soybean hypocotyl. Seeds of soybean cv. Harosoy 63 were grown in dark at 24°C for 6 days in water-soaked vermiculite. The seedlings were then placed onto a tray and the roots were covered with water-soaked cheese cloth. The hypocotyls were inoculated with either water (control) or with 1 mM AgNO_3 (treatment).

2.12.3 Transformation of soybean cotyledons

Humidity chambers were prepared by placing layers of wetted filter paper on the lid of a petri dish. Cotyledons were twisted off of the seedling and collected in batches of 50-60 in a glass beaker. The cotyledons were then surface sterilized with 70% ethanol and dried on filter paper in the biological safety cabinet. The cotyledons were cut with a sterilized razor blade near the end of petiole as shown in Figure 2.3 the cut was horizontally made through the major vein. The cut cotyledons were arranged in rows of 2-3-3-2, a total of 10 cotyledons per chamber. A culture of *A. rhizogenes* was drawn in a sterile syringe with 18 gauge needle and a drop of the culture was placed onto the wounded site. The petri dishes were sealed, placed in complete darkness at 25°C for two days and then transferred into continuous light for 3 weeks at 23°C.

2.12.4 Harvesting hairy roots

After 3 weeks hairy roots were screened for silenced transgenic lines. The transgenic lines were identified by a selectable fluorescent marker, GFP using a fluorescent microscope. Images were taken under visible and UV light and were recorded using camera software Image-pro Express 6.0. Control K599 transformed hairy roots were harvested also. All roots from each corresponding construct, K599 only, empty silencing vector and silenced *GmCHR*, were pooled together, weighed, and frozen in liquid nitrogen and stored at -80°C

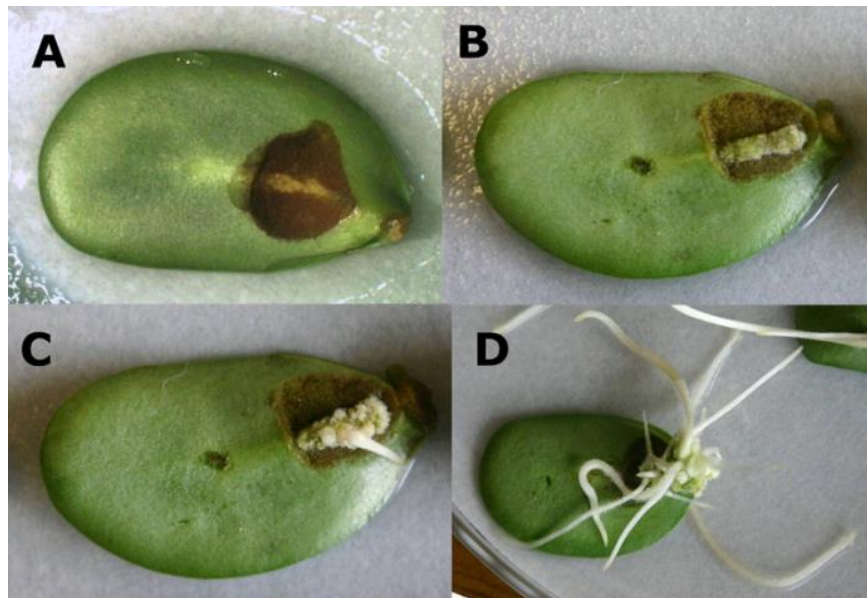


Figure 2.3. Soybean hairy root transformation. (A) Two day old soybean cotyledon (post-inoculation) was cut near the petiole end and a culture of *A. rhizogenes* K599 was placed onto the cut site. (B,C) After 7 days hairy roots began to form. (D) After 3 weeks, hairy roots were ready for harvesting (Adapted from: Subramanian et al., 2005).

Chapter 3: Results

3.1 Soybean genome contains 16 putative GmCHRs

To identify all the *GmCHR* gene family members in soybean, a keyword search was conducted in the annotated *Glycine max* Wm82.a2.v1 genome on Phytozome. Using the keyword “aldo-keto” and “aldo/keto”, protein databases, KOG, Pfam and PANTHER identified 34 and 68 *GmAKRs*, respectively. Both keyword searches were compared and compiled to ensure no duplicates. Each *GmAKR* sequence was then used as a query for a BLAST search, until no new *GmAKRs* were identified. With this process, no new *GmAKRs* were found in the BLAST searches which concluded in the total of 68 *GmAKRs*.

The AKR superfamily consists of 15 families; where CHRs are grouped into the AKR4 family (Jez et al., 1997). It has been previously shown that CHRs from several different plant species cluster separately from other AKR family members (Figure 1.3) (Bomati et al., 2005). To identify *GmCHRs* in soybean, the 68 candidate *GmAKRs* were combined with previously characterized AKRs from Bomati et al. (2005), with the assumption that the *GmCHRs* would cluster together with other known CHRs. As shown in Figure 3.1, 16 putative soybean CHRs clustered together with other known CHRs (Appendix A). *GmCHR2A*, *GmCHR15*, *GmCHR2B*, *GmCHR20*, *GmCHR18* and *GmCHR14* cluster closely together with CHRs from *M. sativa* and CHR *P. montana* var. *lobata*, POR from *L. japonicus* and *G. glaba*. Other inter-species clustering consists of *GmCHR9A*, *GmCHR9B*, *GmCHR9C*, *GmCHR9E*, *GmCHR16A*, *GmCHR16B*, and *GmCHR16C* with CHR from *S. rostrata*. *GmCHR9D* and *GmCHR12* cluster together closely without any CHR from other plant species. Codeinone reductase, a non-CHR, is

also found in the same cluster as CHR however no other GmCHRs cluster together with this group.

CHR's active site is primarily molded by the catalytic tetrad Asp-53, Tyr-58, Lys-87, and His-120 with the additional amino acids Trp-121, and Asn-167 based on *M. sativa* CHR (GenBank accession number AAB41555.1) (Bomati et al., 2005). To identify whether all the putative GmCHRs contain critical amino acids residues that are required for CHR function, 16 putative GmCHRs, known CHRs from other plant species, and a selection of AKRs were aligned and the amino acid residues were searched manually. This process identified the following alterations in 5 candidate GmCHRs: 1) in GmCHR2B, Tyr-58 is missing; 2) in GmCHR9B, Asp-53, Tyr-58, Lys-87, His-120, and Trp-121 are missing; 3) GmCHR7 do not contain any of the critical residues; 4) GmCHR9E is missing His-120, Trp-121 and Asn-167 and 5) GmCHR16C does not contain Asn-167 (Figure 3.2). Based on these results, GmCHR2B, GmCHR9B, GmCHR9E, GmCHR7 and GmCHR16C were eliminated for further study, leaving 11 putative GmCHRs.

Detailed information on each putative *GmCHR*, including gene location, coding sequence length and predicted subcellular localization is shown Table 3.1. Pairwise amino acid sequence identity comparison of the 11 *GmCHR* gene family ranged from 50.11%-97%. However, pairwise nucleotide sequence identity comparison ranged from 37%-96% (Table 3.2).

Figure 3.1. Phylogenetic tree showing GmCHRs cluster together with known plant CHRs.

The protein sequences of putative GmAKRs and known AKRs from other plant species were aligned using CLUSTALW and the phylogenetic tree was constructed by Neighbor-joining method using MEGA6. The green branch indicates CHR specific AKR, black circles (●) and white circles (○) indicate putative GmCHRs and CHRs from other plants species, respectively. POR polyketide reductase, COR codeinone reductase, AKR aldo-keto reductase, XYR xylose reductase, M6PR mannose-6-phosphate reductase, M6DH morphine 6-dehydrogenase, ALHD alcohol dehydrogenase, ALR aldose reductase, ARR arabinose reductase, 2,5 DKG 2,5-diketo-D-gluconic acid reductase B, and PFS prostaglandin F synthase. Branch lengths are measured as substitutions per site. Accession numbers for known AKRs (shown in the parentheses) are according to Uniprot. Common nomenclature for GmCHRs is also shown in parenthesis.

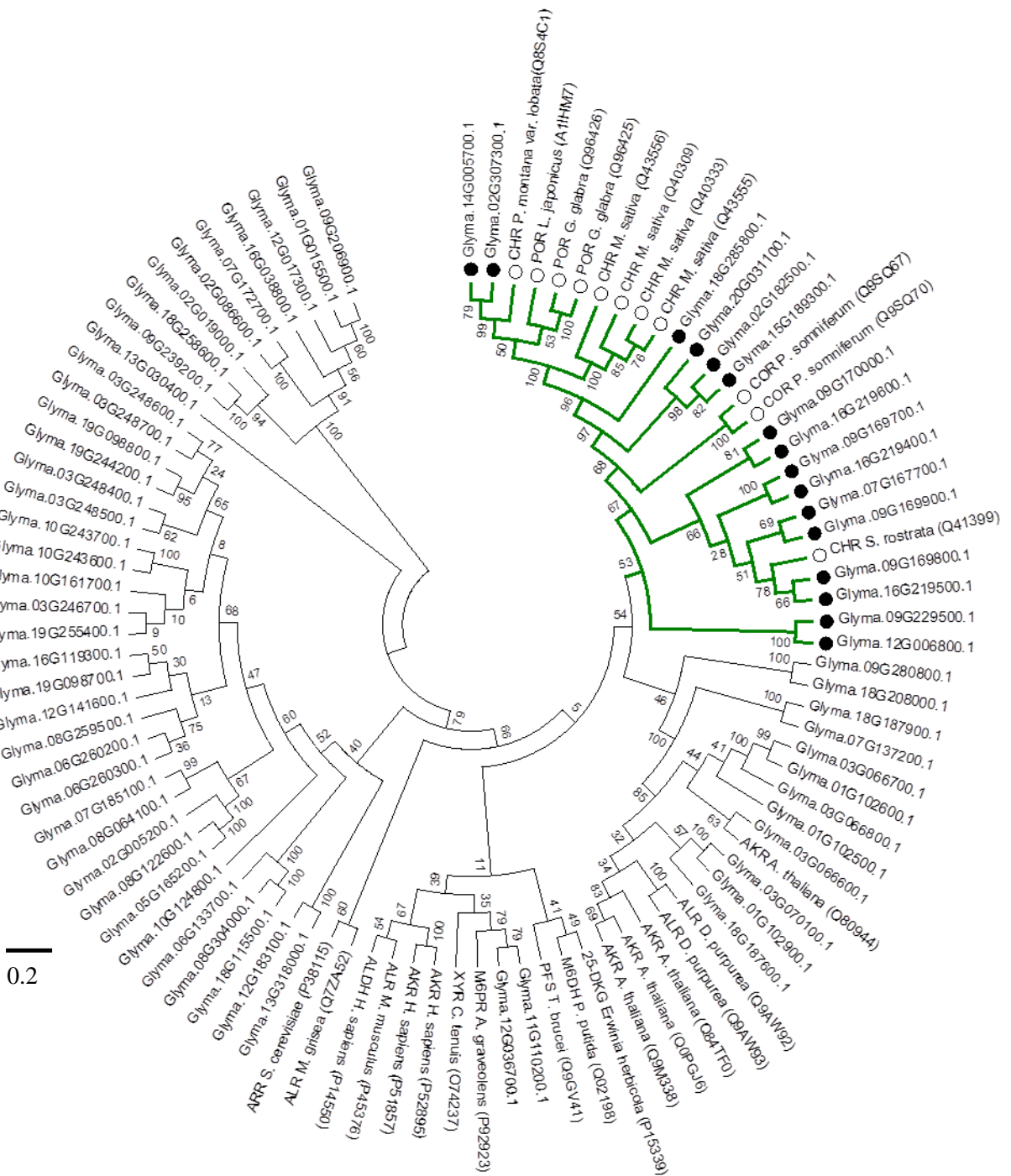


Figure 3.2. Identification of important amino acid residues in candidate GmCHRs.

Candidate GmCHRs obtained from the phylogenetic analysis were aligned using CLUSTALW with other known AKRs and CHRs from other plant species. Critical and other residues were noted: entrance of the catalytic site (□), AKR catalytic tetrad (□), unique amino acids from COR(□), CHR active site with AKR catalytic tetrad (□). GmCHRs which are missing critical residue are indicated through gray shading. Only an abridged version of the alignment is shown. Full version of the alignment is shown in Appendix B.

2, 5-DKG 46 D I A V N Y E N E S E V G R A V R A S - - - - S V D R D E L I V A S K I P G R Q H G R A E A V D S I R G S I - - D R L G
 XYR 47 D I A E I Y E N E K E V G D G V K R A I D E G L V - K R E E I F L T S K L W N Y H D P K N V E T A L N K T L - A D L K
 ALDH 45 D I A A I Y E N E P E L G E A L K E D V G P G K A V P R E E L F V T S K L W N T K H H P E D V P A L R K T L - A D L Q
 ALR 47 D I A R I Y E N E K E I G V V L K K L F D D G V V - K R E D L F T S K L W S T D H A P E D V P V A L D K T L - E D L Q
 GmCHR9D 47 D I A K I Y G S E P A L G K A I N E A I C E G E I - E R E G I F L T S K L W G S D H H D P - - V S A L K Q T L - E N L G
 GmCHR12 47 D I A K I Y G S E P A L G K A I N E A I C E E E V - E R E D I F L T S K L W G S D H H N P - - V S A L K Q T L - E N L G
 GmCHR2E 36 D I A S - - - - - D L Q L G L V K S I E E L F F T S K L W P T N N F P H L V L P A L Q K S I - Q K F N
 GmCHR15 52 D I A S I Y G S E Q P L G E A T A E A L Q L G L I G S R D E L F T S K L W C T D N F P H L V L P A L Q K T L - R S L K
 GmCHR20 57 D I A A C Y G S E Q A L G E A T A E A L R V G L I A S R D E L F T S K L W C C D N H P H L V L P A L Q N S I - R S L K
 GmCHR18 54 D I A S P Y G S E Q S V G E A T A E A L K Q G L I A S R D Q L F T T K L W V T D N H A H T I L P A L Q K S I - R T L Q
 CHR P.montana 53 D I A A P Y G S E Q A L G E A L K E A V D L G L V - S R Q D L F V T S K L W V T D N H P H L V P A L R K S I - K T L Q
 GmCHR2A 54 D I A A P Y G S E Q A L G E A L K E A T H L G L V - T R Q D L F V T S K L W V T E N H P H L V P A L R K S I - K T L Q
 GmCHR14 54 D I A A P Y G S E Q A L G E A L K E A I H L G L V - S R Q D L F V T S K L W V T E N H P H L V L P A L R N H - - - - -
 CHR M.sativa 53 D I A A P Y G S E Q A L G E A L K E A T E L G L V - T R E L F V T S K L W V T E N H P H L V I P A L Q K S I - K T L Q
 CHR M.sativa 53 D I A A P Y G S E Q A L G E A L K E A T E L G L V - T R D D L F V T S K L W V T E N H P H L V I P A L Q K S I - K T L Q
 CHR M.sativa 53 D I A A P Y G S E Q A L G E A L K E A T E L G L V - T R E L F V T S K L W V T E N H P H L V I P A L Q K S I - K T L Q
 CHR M.sativa 53 D I A A P Y G S E Q A L G E A L K E A T E L G L V - T R E L F V T S K L W V T E N H P H L V I P A L Q K S I - K T L Q
 POR L.japonicus 54 D I A A P Y G S E Q A L G E A L K E A L D L G L V - S R E L F V T S K L W V T E N H P H L V P A L Q K S I - K T L Q
 POR G.glabra 55 D I A A P Y G S E T A L G E A L K E A R D L G L V - T R E D L F V T S K L W V T E N H P H L V P A L R K S I - E T L Q
 POR G.glabra 54 D I A A P Y G S E T A L G E A L K E A R D L G L V - T R E D L F V T S K L W V T E N H P H L V P A L R K S I - E T L Q
 GmCHR7 1
 COR1 51 D I A A P Y G S E E C L G E A T A E A L Q L G L I K S R D E L F T S K L W C A D A H A D L V L P A L Q N S I - R N L K
 COR2 51 D I A A P Y G S E E C L G E A T A E A L Q L G L I K S R D E L F T S K L W C A D A H A D L V L P A L Q N S I - R N L K
 GmCHR9C 51 D I A A P Y G S E E P L G Q A T A L A D Q L I K S R N E L F V T T K L W C T D A H P G L V L P A L E S S I - Q R L G
 GmCHR16C 51 D I A T I Y G S E E A L G L A A Q A Q R Q G L I K N R G E I F V T I K L W C S D S H P D L V I P A L K K S I - Q R L G
 CHR S.rostrata 51 D S A S V Y G T E E A I G M A V S K A T E Q G L I K S R D E V F T S K P W N T D A H D L I V P A L K T I L - K K L G
 GmCHR9E 51 D I A A V Y G T E E A I G L A V A N A T E K G L I K S R D E V F T S K P W N T D A R R D L I V P A L K T I L - K Y H *
 GmCHR9B 1
 GmCHR16B 51 D I A A V Y G T E E A I G L A V A K A I D K G L I K S R D E V F T S K P W N T D A H R D L I V P A L K T I L - K K L G
 GmCHR9A 51 D I A S I Y G - - - - - N L G L I K S R D K V F I T S K L W T D A H P D L V P I L K T H C R K K L G
 GmCHR16A 51 D I A S I Y G S E E S L G K A V A K A L E L G L I N S R E L F T S K L W S T D A H P D L V P A L K T S I - Q K L G

2, 5-DKG 115 - - V G R W L D T W R G M I D A R E A G L V S I G V S N I T E P M L K T L - D E T G V P A V N O V E L H P Y I P O A
 XYR 142 Y E I V P I L E T W R A L E K L V A A G K I K S I G V S N I P G A L L D L L R G A T I K P A V L Q V E H H P Y L Q O P
 ALDH 135 Y D S T H Y K E T W R A L E A L V A K G I V Q A L G I S N N S R Q I D D I L S V A S V R P A V L Q V E C H P Y L A Q N
 ALR 132 L I E T D I P G T W R A M E A L Y D S G K A F A I G V S N I T L K K L S D L L D V A R I P P A V N O V G C H P S C A Q T
 GmCHR9D 131 F E K L D E T T W A G M E K C L E M G L C R C I G V S N I S S K K T E C L L D Y A S T P P A V N O V E M H P M W R O G
 GmCHR12 131 F E K L D E T T W A G M E K C L E M G L C R C I G V S N I S S K K T E C L L D Y A S T P P A V N O V E M H P M W R O G
 GmCHR2B 132 T - T F D L R G V W K G M E E C O K L G L A K S I G V S N I T C K K L E D L L S F A T I P P S V N O V E M N P A W H O K
 GmCHR15 140 T - S F D L K G V W K G M E E C O K L G L T K C I G V S N I S C K K L E N L L S F A T I P P S I N O V E M N P T W O O K
 GmCHR20 145 V - F F D M K G V W A M E E C H K L G L T K S I G V S N I S C K K L E N L L S F A T I P P S V N O V E M N L A W O O K
 GmCHR18 142 V - F F D M K G V W S M E E C Q L G L T K A I G V S N I S I K K L E K L S F A T I P P A V N O V E M N L G W O O Q
 CHR P.montana 140 L - F F D V K G V W E M Q E C O K L G L T K A I G V S N I S V K K L Q N L L S V A T I R P V V N O V E M N L A W O O K
 GmCHR2A 141 L - F F D V K G V W S M E E C O K L G L T K A I G V S N I S V K K L Q N L L S V A T I R P V V D Q V E M N L A W O O K
 GmCHR14 136 L - F F D V K G V W S M E E C O K L G L T K A I G V S N I S V K K L Q N L L S V A T I R P V V D Q - - - - - K
 CHR M.sativa 140 L - F F D V K G V W S M E E S I K L G L T K A I G V S N I S V K K L E N L L S V A T I P P A V N O V E M N L A W O O K
 CHR M.sativa 140 L - F F D V K G V W S M E E S I K L G L T K A I G V S N I S V K K L E N L L S V A T I P P A V N O V E M N L A W O O K
 CHR M.sativa 140 L - F F D V K G V W S M E E S I K L G L T K A I G V S N I S V K K L E N L L S V A T I P P A V N O V E M N L A W O O K
 CHR M.sativa 140 L - F F D V K G V W S M E E S I K L G L T K A I G V S N I S V K K L E N L L S V A T I P P A V N O V E M N L A W O O K
 POR L.japonicus 141 L - F F D V K G V W S M E E C K L G L T K A I G V S N I S V K K L Q N L L S V A N I L P A V N O V E M N L A W O O K
 POR G.glabra 142 L - F F D V K G V W S M E E S I K L G L T K A I G V S N I S V K K L Q N L L S V A T I R P A V N O V E M N L A W O O K
 POR G.glabra 141 L - F F D V K G V W S M E E S I K L G L T K A I G V S N I S V K K L Q N L L S V A T I R P A V N O V E M N L A W O O K
 GmCHR7 1
 COR1 139 L - F M D Y K S V W A M E E C Q L G F T R A I G V S N I S C K K L Q E L M A A A K I P P V V N O V E M S P T L H O K
 COR2 139 L - F M D Y K S V W A M E E C Q L G F T R A I G V S N I S C K K L Q E L M E T A N S P P V V N O V E M S P T L H O K
 GmCHR9C 139 L - F L D M K G T W E D M E R C S K L G L A K S I G V S N I G V K K L S E I L Q N A R V P P A L V O V E M N A W O O E
 GmCHR16C 138 L - F F D M K G T W E R M E E C S K L G L T K S I G T - - - - - E M N V A W O O G
 CHR S.rostrata 140 L - F F D I E G T W R A M E E C V E L G L A K S I G I C N I G T K K I T K L L E I A T I P P A V N O V E M N P S W O O G
 GmCHR9E 1
 GmCHR9B 1
 GmCHR16B 140 L - F F D I E G T W R A M E E C Y K L G L A K S I G I C N I G I K K L T K L L E I A T I P P A V N O V E M N P S W O O G
 GmCHR9A 128 L P S F D M K G I W E A M E E C V E L G L A K S I G V S N I G I K K L T Q L E N A T I P P A V N O V E I S P A W O O G
 GmCHR16A 140 L P S F D M K G I W E A M E E C V E L G L A K S I G V S N I G I K K L S Q L L E N A T I P P A V N O V E M S P A W O O G

3.2 *GmCHR* family members display tissue-specific gene expression

To determine tissue-specific expression pattern of *GmCHR*s, publicly accessible RNA-seq data was used. RNA-seq data was derived from high throughput sequencing of total RNA isolated from various soybean tissues at different developmental stages such as different stages of developing embryos, the seed coat, the pod wall, roots, stems, leaves, flowers and flower buds ([http:// www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29163](http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29163)). The relative expression was normalized across the libraries corresponding to each tissue. A heatmap was produced based on expression level of each *GmCHR* for each tissue.

No RNA-seq data was found for *GmCHR9A* and *GmCHR9D* suggesting that they may not be transcribed. *GmCHR12* was detected in the RNA-seq data however the transcript levels were low (Appendix C). Due to this reason, *GmCHR9A*, *GmCHR9D* and *GmCHR12* were not included into the heatmap.

Out of 8 putative *GmCHR*s remaining on the heatmap, *GmCHR2A*, *GmCHR14*, *GmCHR18* and *GmCHR20* were expressed at higher level in roots as compared to other tissues (Figure 3.3). Majority of the *GmCHR*s were either expressed in the seedlings or dry seeds.

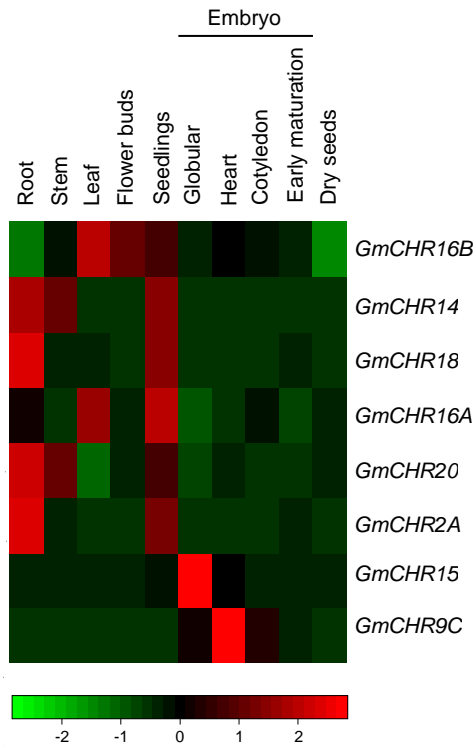


Figure 3.3. Expression profile of *GmCHR* genes. Soybean RNA-seq data (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29163>) were normalized and a heatmap was constructed to evaluate the transcript levels of *GmCHR* across different tissues. Transcript abundance is indicated by a gradient from red (high) to green (low).

Table 3.1. Genetic characteristics of the soybean chalcone reductase gene family

Name	Gene	Locus name	Locus Range	Splice variants	Coding sequence (nt)	Protein Molecular Weight (kDa)	Predicted Sub-localization	Remarks
<i>GmCHR2A</i>		Glyma.02G307300	Chr02: 48163443..48164865	1	948	35.59	Cytoplasm	
<i>GmCHR9A</i>		Glyma.09G169700	Chr09: 39437065..39439064	1	846	31.86	Cytoplasm	No-RNA seq data available
<i>GmCHR9C</i>		Glyma.09G170000	Chr09: 39453516..39456085	1	957	35.48	Cytoplasm	
<i>GmCHR9D</i>		Glyma.09G229500	Chr09: 45341437..45343797	1	969	36.49	Cytoplasm	No-RNA seq data available
<i>GmCHR12</i>		Glyma.12G006700.	Chr12: 508878..511003	1	948	35.85	Chloroplast	Unable to clone
<i>GmCHR14</i>		Glyma.14G005700	Chr14: 463009..64543	1	981	36.77	Nucleus	
<i>GmCHR15</i>		Glyma.15G189300	Chr15: 20075936..20082802	1	948	35.43	Cytoplasm	Unable to clone
<i>GmCHR16A</i>		Glyma.16G219400	Chr16: 37672779..37675776	3	963	36.15	Cytoplasm	
<i>GmCHR16B</i>		Glyma.16G219500	Chr16: 37677371..37670638	1	969	36.04	Cytoplasm	
<i>GmCHR18</i>		Glyma.18G285800	Chr18: 56611460..56612978	2	948	35.14	Cytoplasm	
<i>GmCHR20</i>		Glyma.20G031100	Chr20: 3790362..3793599	1	966	35.91	Cytoplasm	

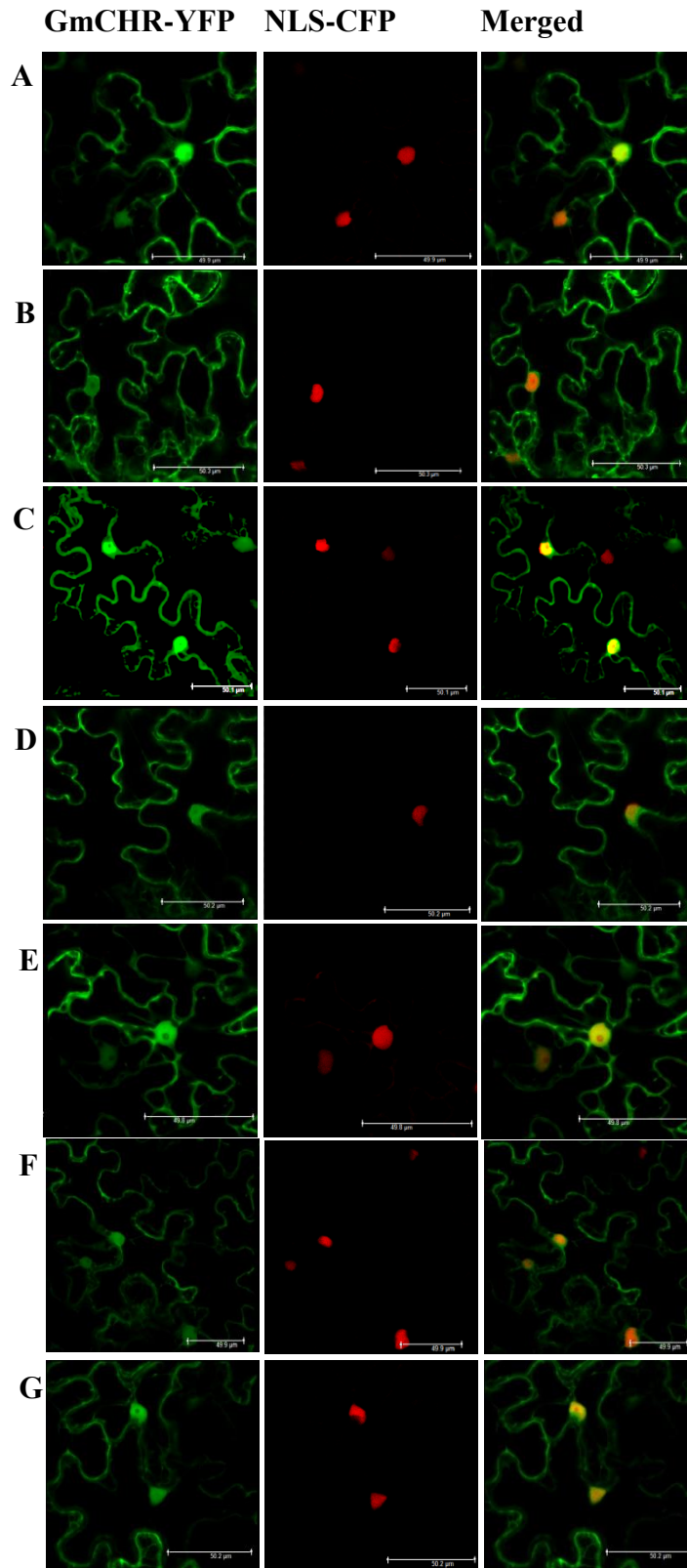
CHR chalcone reductase; nt nucleotide; kDa kilo-Dalton

Table 3.2. Pairwise coding region and amino acid sequence comparison in the soybean chalcone reductase

	<i>GmCHR2A</i>	<i>GmCHR9B</i>	<i>GmCHR9C</i>	<i>GmCHR9D</i>	<i>GmCHR12</i>	<i>GmCHR14</i>	<i>GmCHR15</i>	<i>GmCHR16A</i>	<i>GmCHR16B</i>	<i>GmCHR18</i>	<i>GmCHR20</i>
<i>GmCHR2A</i>		55.3	53.27	50.32	50.11	95.35	64.45	57.49	56.22	69.09	66.14
<i>GmCHR9B</i>	47.23		63.74	54.65	53.67	56.85	57.79	72.51	96.75	54.88	54.33
<i>GmCHR9C</i>	46.67	56.67		52.43	51.9	56.33	57.38	67.92	64.58	55.06	56.43
<i>GmCHR9D</i>	38.41	44.63	43.81		97.67	52.33	53.16	53.16	54.16	50.63	52.85
<i>GmCHR12</i>	38.41	44.3	44.13	95.87		51.94	53.16	53.48	54.01	50.42	52.43
<i>GmCHR14</i>	92.7	47.23	45.91	37.14	37.14		66.54	58.79	58.14	72.35	69.25
<i>GmCHR15</i>	58.1	52.77	50.48	41.59	41.59	57.14		57.91	58.23	66.35	74.79
<i>GmCHR16A</i>	47.62	66.45	62.89	46.03	45.4	47.5	54.29		72.9	58.54	57.42
<i>GmCHR16B</i>	46.35	96.09	56.92	44.44	44.13	45.34	51.75	66.56		56.16	53.93
<i>GmCHR18</i>	64.76	45.8	47.3	37.78	37.14	64.76	59.37	50.16	45.4		66.88
<i>GmCHR20</i>	55.56	46.25	50.0	39.05	38.73	55.14	70.48	49.38	44.55	58.41	
	Nucleotide										

Sequence identity (%) between seven chalcone reductase family members at the nucleotide and amino acid levels; nucleotide sequences correspond to the coding region sequences for each gene, acquired from the Joint Genome Institute assembly of Glycine max genome; amino acid sequences are deduced from the coding sequences; nucleotide and amino acid sequences were aligned by CLUSTALW and percentage identities were calculated from the subsequent multiple sequence alignment.

Figure 3.4. Subcellular localization of the GmCHR family. GmCHR_s were translationally fused with the YFP reporter gene, transformed into *N. benthamiana* leaves by *A. tumefaciens*-mediated transformation: (A) GmCHR2A, (B) GmCHR9C, (C) GmCHR14, (D) GmCHR16A, (E) GmCHR16B, (F) GmCHR18, and, (G) GmCHR20. A nuclear localization signal fused with the CFP reporter gene was used for the co-localization study. Fluorescence was visualized by confocal microscopy. The scale bar: (A) 49.9 μm , (B) 50.3 μm , (C) 50.1 μm , (D) 50.2 μm , (E) 49.8 μm , (F) 49.9 μm and (G) 50.2 μm .



3.3 GmCHR localizes in the nucleus and cytoplasm in *N. benthamiana*

To determine the subcellular localization of the 9 putative *GmCHR* genes, each member was translationally fused with YFP. Despite several attempts, amplification of *GmCHR15* and *GmCHR12* was not successful. Therefore I proceeded with the investigation of subcellular localization of 7 GmCHRs. The GmCHR-YFP fusion protein was transiently expressed in leaf epidermal cells of *N. benthamiana* followed by confocal microscopy. The results revealed both nuclear and cytoplasmic localization for GmCHR2A, GmCHR9C, GmCHR14, GmCHR16A, GmCHR16B, GmCHR18 and GmCHR20. To confirm the presence of the nuclear localization, co-expression of GmCHR-YFP fusion protein with nuclear localization signal containing control (NLS-CFP) was performed. Co-expression of GmCHR-YFP and NLS-CFP showed overlap between the NLS-CFP and GmCHR-YFP signals, thus confirming their nuclear localization (Figure 3.4).

Foreign proteins, such like GmCHR-YFP, when overexpressed, may get cleaved by endogenous proteases *in planta* (Outchkourov et al., 2003). To further confirm the nuclear localization of the intact GmCHR-YFP, *A. tumefaciens* GV3101 containing the pEarleyGate101-GmCHR2A was infiltrated into leaves of *N. benthamiana* and leaf samples were collected from day 1 to day 3 post-infiltration. Total soluble proteins were extracted from the leaf samples and Western blot was performed. The results revealed that the YFP fluorescence observed by the transient expression is the result of the intact GmCHR-YFP (63 kDa) and not from the cleaved YFP (Figure 3.5).

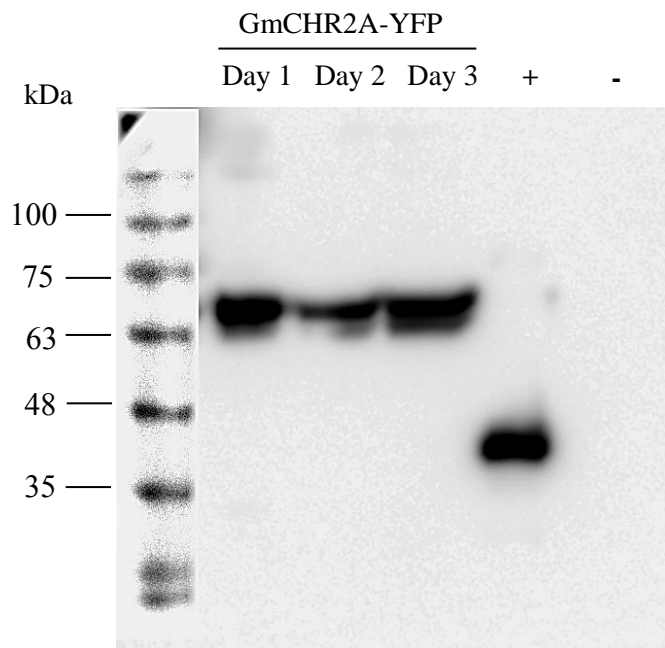


Figure 3.5. Accumulation of GmCHR2A-YFP in *N. benthamiana* leaves. *A. tumefaciens* containing GmCHR2A-YFP construct were transiently expressed in leaves of *N. benthamiana* from Day 1 to Day 3 post infiltration. Total soluble proteins (30 μ g) were separated on SDS-PAGE and transferred to PVDF membrane by electroblotting. GmCHR2A-YFP was detected by sequential incubation of the blot with anti-GFP antibody and HRP conjugated goat anti-mouse antibody, followed by chemiluminescent reaction. (+) and (-) indicate eGFP with hydrophobin tag (37 kDa) as a positive control and total soluble proteins from non-filtrated leaves as negative control, respectively.

3.4 Root-specific *GmCHR*s are induced upon stress

To identify *GmCHR*s genes that are induced upon pathogen attack, soybean cv. L76-1988 stems were infected with *P. sojae* agar culture. Stem samples were collected 24, 48 and 72 h post-infection and expression analysis of 7 putative *GmCHR*s was performed using RT-PCR. The results clearly demonstrated that the expression of root-specific *GmCHR* genes *GmCHR2A*, *GmCHR14*, *GmCHR18* and *GmCHR20* were induced after 24 h and remained induced until 72 h post infection. Uninoculated stems of soybean plants grown at the same time were used as control. Expressions of *GmCHR9C* and *GmCHR16B* were undetectable in both control and infected samples. However, accumulation of *GmCHR16A* remained unchanged in both infected and control samples (Figure 3.6).

Since treatment of AgNO_3 induces phytoalexin production in soybeans, it has been used in the past as to mimic pathogen attack (Ward et al., 1979; Moy et al., 2004; Kubeš et al., 2014). Soybean cv. Harosoy 63 seeds were grown in the dark for seven days. The seedlings were placed onto a tray and the roots were covered with water-soaked cheese cloth. The hypocotyls were inoculated with either water droplets or with 1 mM AgNO_3 . Inoculated tissues were collected at 6, 12, 24, 48, or 72 hours. Upon AgNO_3 treatment, hypocotyls displayed brown lesions at 12 hour and are worsened in the later time points in comparison to the control hypocotyl that displayed no lesions (Figure 3.7). To further investigate the induced gene expression profile of root-specific *GmCHR*s, qPCR was performed. Expression of *GmCHR2A* was significantly induced at 6, 12, 24 and 48 hours. *GmCHR14* was significantly induced at 12, 24, 48 and 72 hours. Likewise, *GmCHR18* was induced at 12, 24, 48 and 72 hours, whereas *GmCHR20* was induced only at 6 and 72 hours post treatment (Figure 3.8). Primer efficiencies for root-specific *GmCHR*s are noted in Appendix D

3.5 Identification of QTLs and QTL markers linked to *P. sojae* resistance

To determine QTLs and QTL markers that are linked to *P. sojae* resistance, a survey of soybean database and a literature search were conducted. A search in the ‘SoyBase and Soybean Breeder’s Toolbox’ from the year 2003 to 2013 identified 55 QTLs that are linked to *P. sojae* resistance in soybean (Appendix E). These 55 QTLs were extensively researched for GmCHR loci, parental lines and LOD scores. Three QTLs, Phytoph 10-3, Phytoph 14-3, Phytoph 8-2, were found which contain GmCHR loci. Highlights of the QTLs included: (1) Phytoph 10-3 contains GmCHR2A locus, spans over 2 megabase pairs and has an LOD score of 29.7; (2) Phytoph 14-3 contains GmCHR18 locus, spans over 43 megabase pairs and has the LOD score of 3.4; (3) Phytoph 8-2 contains GmCHR20 locus, stretches over 31 megabase pairs with the LOD score of 4.5 (Table 3.3 and Figure 3.9). These details strongly suggest that Phytoph 10-3 containing GmCHR2A is involved in *P. sojae* resistance in soybean.

An additional literature search was conducted for QTL markers linked to *P. sojae* resistance in soybean. This process identified over 500 QTL markers. Markers that share the same chromosome as GmCHR were segregated, and exact locations of the QTL markers were noted. Upon filtering through the QTL markers, a total of six markers were found to flank GmCHR2A, GmCHR16A, GmCHR18 and GmCHR20 loci (Appendix F, and Figure 3.9). The details on the marker are summarized in Table 3.4.

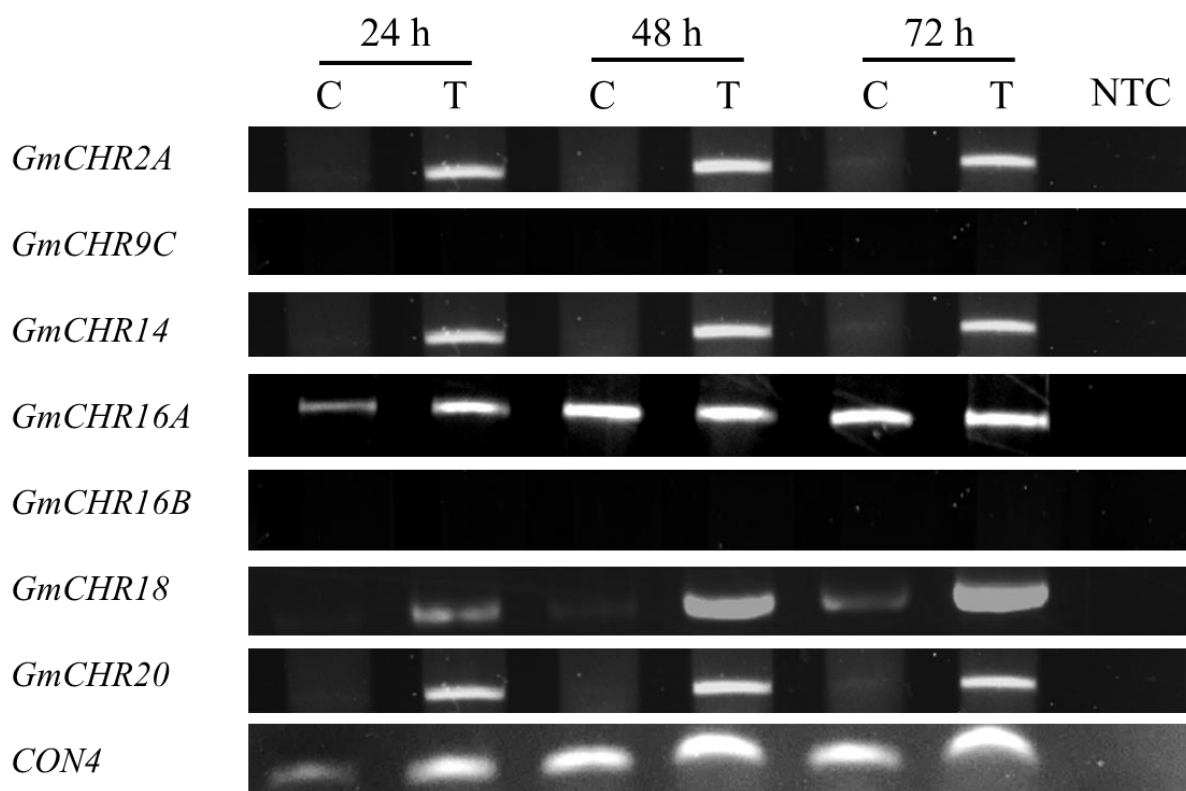


Figure 3.6. Gene expression of *GmCHR*s in response to *P. sojae* infection. Total RNA (1 µg) was used to synthesize cDNA from uninfected (C) or *P. sojae*-infected stems of soybean cv. L76-1988 (T) at the time points as indicated after infection. Expression analysis was conducted by RT-PCR with *GmCHR* gene-specific primers. NTC indicates no template control. *CON4* is shown as a reference gene.

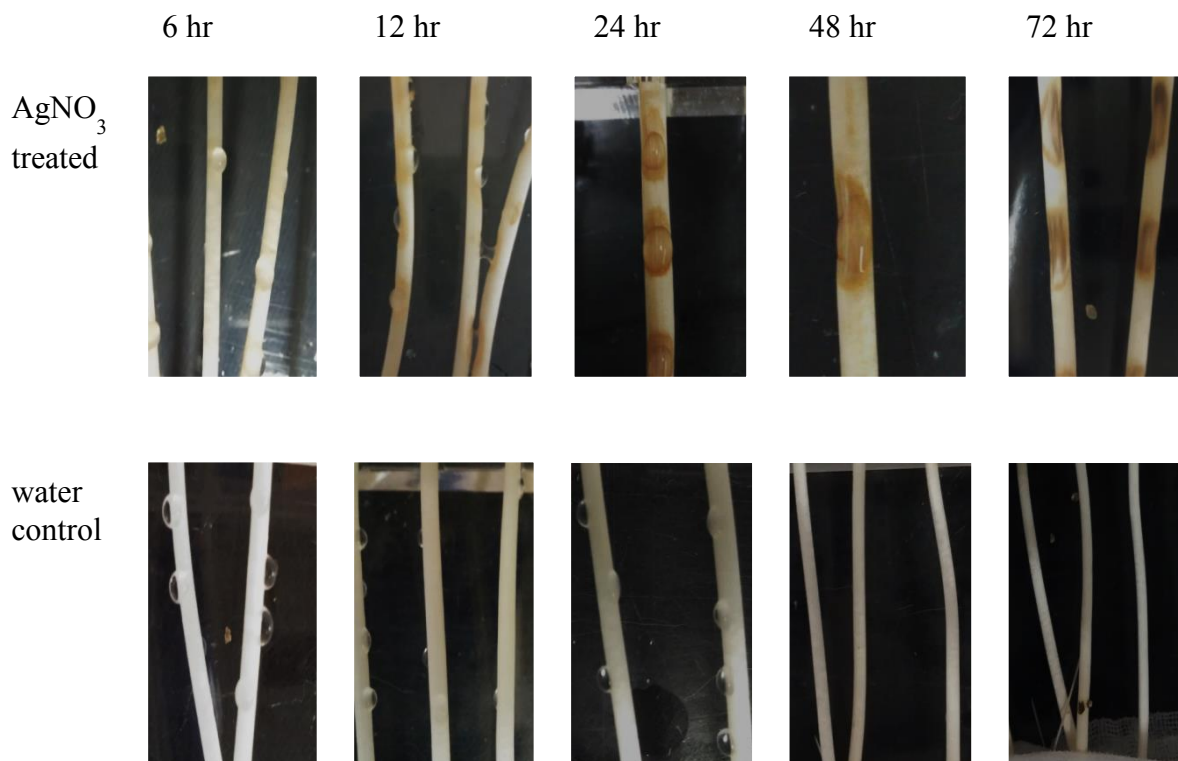


Figure 3.7. Effects of AgNO₃ on etiolated soybean cv. Harosoy 63 hypocotyls. Seeds of soybeans cv. Harosoy 63 were grown in dark at 24°C for 6 days in water-soaked vermiculite. The seedlings were then placed onto a tray and inoculated with either water (control) or 1 mM AgNO₃ (treated). Photographs were taken at the time points as indicated.

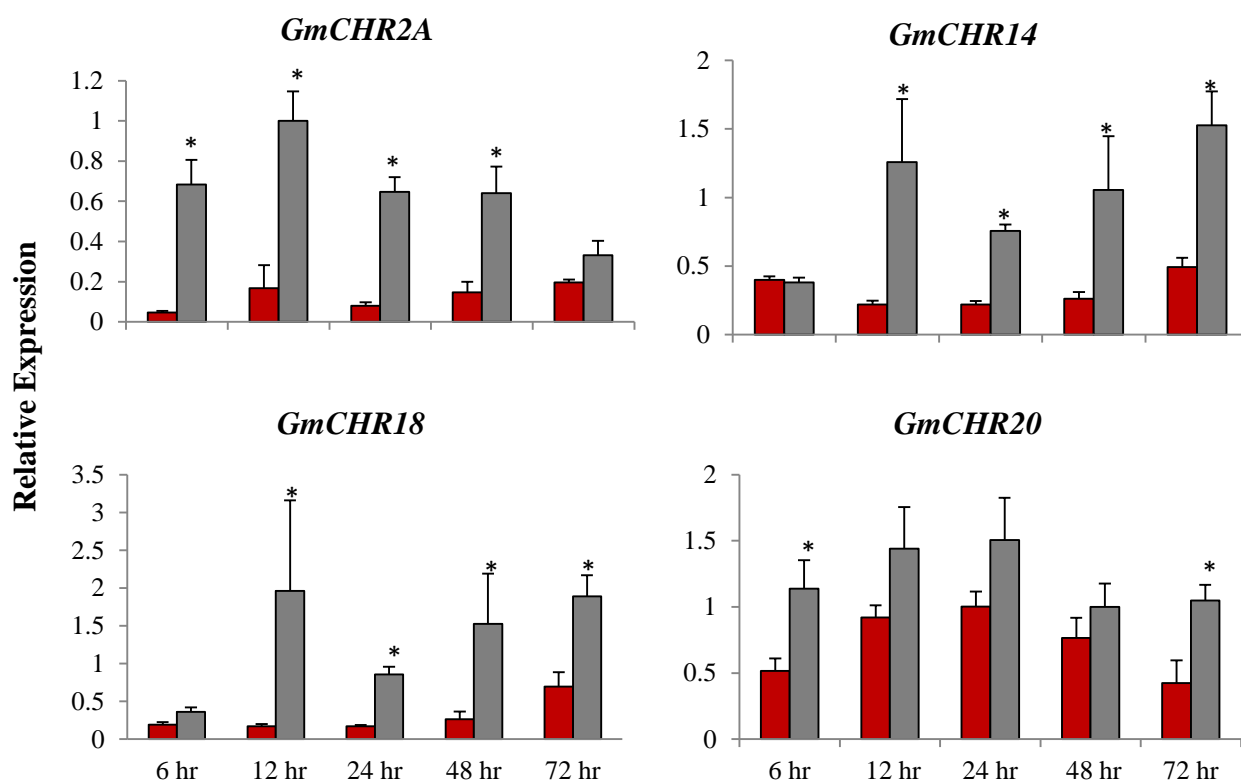


Figure 3.8. Expression of root-specific *GmCHR*s in response to AgNO_3 treatment. Total RNA (1 μg) of soybean cv. Harosoy 63 was used to synthesized cDNA from untreated and AgNO_3 treated hypocotyls Expression analysis was conducted by qPCR with *GmCHR* gene-specific primers. Error bars indicates standard error of the mean (SEM) of two biological and three technical replicates per biological replicates. *CON4* was used as a reference gene. The asterisks (*) denotes significant expression (multiple t-test per row, $p < 0.05$). The red bar represents the control while the gray bar represents the AgNO_3 treatments.

Table 3.3. Characteristics of QTLs linked to *P. sojae* resistance which contain *GmCHRs*

QTL	Chr	QTL Marker Interval (physical map)	<i>GmCHR</i> and CHR location	Parents	Heritability	LOD	Reference
Phytoph 10-3	2	Gm02: 48,345,840 - 46,353,672	<i>GmCHR2A</i> Gm02: 48,163,443 - 48,164,792	Parent 1: Conrad Parent 2: OX760-6-1	n/a	29.77	(Han et al., 2008)
Phytoph 14-3	18	Gm18: 59,499,678- 16,804,048	<i>GmCHR18</i> Gm18: 56,611,421- 56,613,070	Parent1: OX20-8 Parent 2: PI 398841	0.77	3.4	(Wang et al., 2012)
Phytoph 8-2	20	Gm20: 34,100,981- 3,915,838	<i>GmCHR20</i> Gm20: 3,790,428- 3,793,674	Parent 1: V71-370 Parent 2: PI407162	0.89	4.5	(Tucker et al., 2010)

Chr chromosome; LOD logarithm of odds

Table 3.4. Characteristics of QTL markers linked to *P. sojae* resistance that flank *GmCHR*

<i>GmCHR</i> and location	QTL Marker	QTL Marker location	Type of Marker	Parents	Reference
<i>GmCHR2A</i>					
Gm02: 48,163,443- 48,164,792	Satt274	48,345,948- 48,346,001	SSR	Parent 1: Conrad Parent 2: OX760-6-1	(Han et al., 2008)
<i>GmCHR16</i>					
Gm16: 37,678,679- 37,678,780	Satt244	33,327,176- 33,327,379	SSR	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
<i>GmCHR18</i>					
	BARC-039397-07314	56,889,971- 56,889,971	SNP	Parent 1: Conrad Parent2: Sloan	(Wang et al., 2010)
Gm18: 56,611,421- 56,613,070	BARCSOYSSR_18)_1777	54,744,147- 54,744,204	SNP		
	Satt472	53,866,536- 53,866,808	SSR		
<i>GmCHR20</i>					
Gm20: 3,790,324- 3,793,674	Satt614	3,915,838- 3,916,147	SSR	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)

SSR simple short repeats; SNP single nucleotide polymorphism

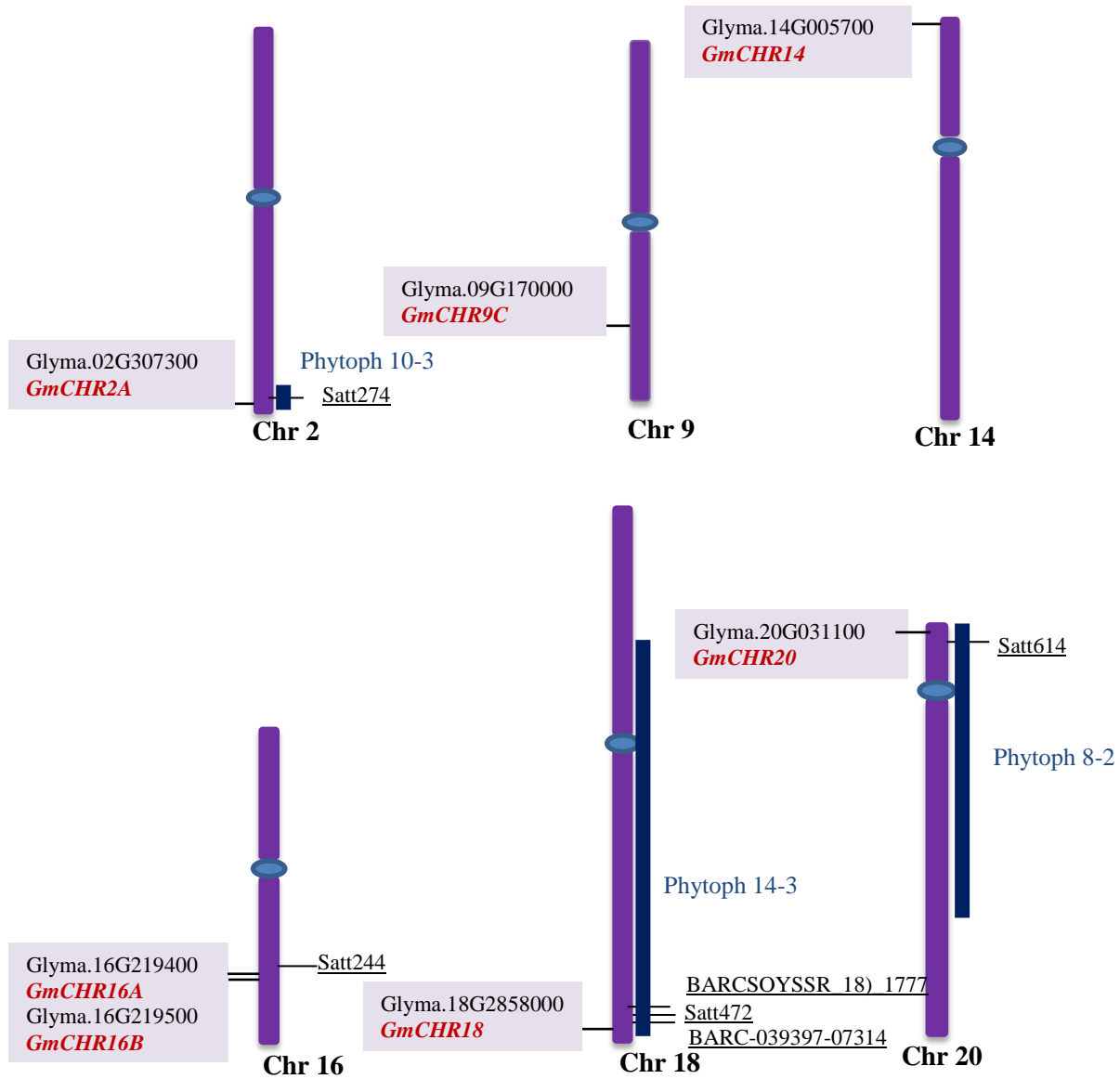
3.6 Accumulation of root-specific GmCHR is higher in *P. sojae* resistant soybean cultivar

Since several QTL markers and QTLs linked to *P. sojae* resistance contained GmCHR gene family members, I wanted to assess the parental cultivars of QTL Phytoph 10-3 for root-specific GmCHR gene expression. Root tissues of two-week old seedlings of soybean cv. Conrad and OX760-6, were used for gene expression analysis using qPCR. As shown in Figure 3.10, expression of GmCHR2A, GmCHR14 and GmCHR18 were significantly higher in the Conrad (*P. sojae* resistant cultivar) compared to the OX760-6 (*P. sojae* susceptible cultivar). No difference in the expression of GmCHR20 was observed between Conrad and OX760-6.

3.7 RNAi silencing of root-specific GmCHRs

For the functional characterization of the root-specific GmCHRs, and how they influence the production of daidzein, RNAi silencing of root-specific GmCHRs was performed. The gene-specific unique sequences for four root-specific GmCHRs were cloned into the RNAi vector pK7GWIWG2D(II). Primers were designed to amplify unique regions of each GmCHRs sequence to facilitate targeted silencing using the soybean hairy root system (Table 2.1). The transgenic roots were selected based on GFP expression as the vector contains a separate GFP cassette (Figure 2.1 and Figure 3.11). The transgenic roots were used to assess the target gene expression before proceeding into measurement of isoflavonoid content. The qPCR analysis of GmCHR2A, GmCHR14, GmCHR18 and GmCHR20 did not show a silencing effect in the silenced lines in comparison to the control (*A. rhizogenes* K599 only) or empty vector (silencing vector only) (Figure 3.12).

Figure 3.9. Genomic distribution of *GmCHR* genes, QTL and QTL markers on soybean chromosomes. Chromosomal locations of *GmCHR*s are indicated based on the location of the genes, length of chromosomes and positions of centromeres. QTL and QTL markers linked to *P. sojae* resistance are mapped against approximate location of *GmCHR*s. The chromosomes are drawn to scale and chromosome numbers are shown beside each chromosome. Dark navy bars indicate QTL regions while underline notations indicate QTL markers. Centromeres are indicated by blue circles on the chromosomes.



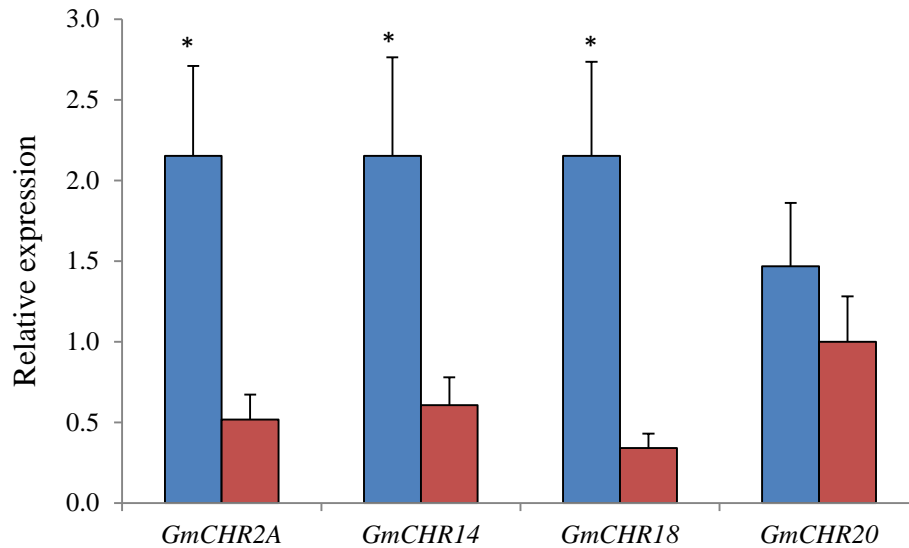


Figure 3.10. Accumulation of root-specific *GmCHR*s in soybean cv. Conrad and OX760-6. Total RNA (1 μ g) was used to synthesize cDNA from soybean cv. Conrad and OX760-6. Expression analysis was conducted by qPCR using gene-specific primers. Blue bars corresponds to the expression of *GmCHR*s in soybean cv. Conrad while, red bars indicate the expression of *GmCHR*s in soybean cv. OX760-6. Error bars indicates SEM of two biological and three technical replicates per biological replicates (multiple t-test per row, $p < 0.05$). *CON4* was used as a reference gene. The asterisks (*) denotes significant difference in expression.

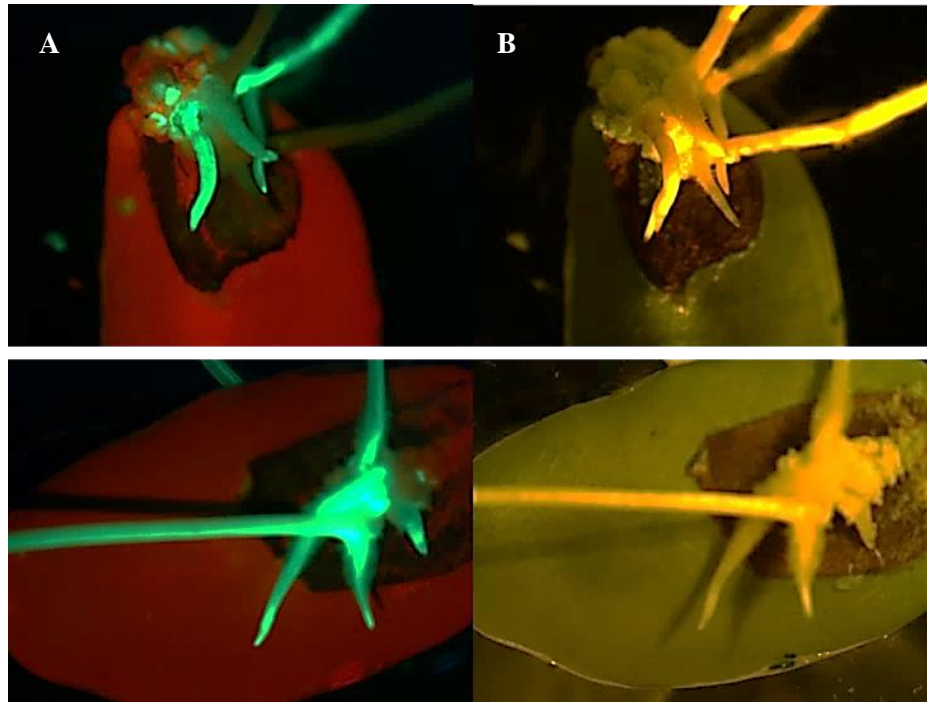


Figure 3.11. Generation of *GmCHR* silenced hairy roots. The transformed roots were identified by a selectable fluorescent marker, green fluorescence protein (GFP). Images were taken under (A) UV light, (B) visible light.

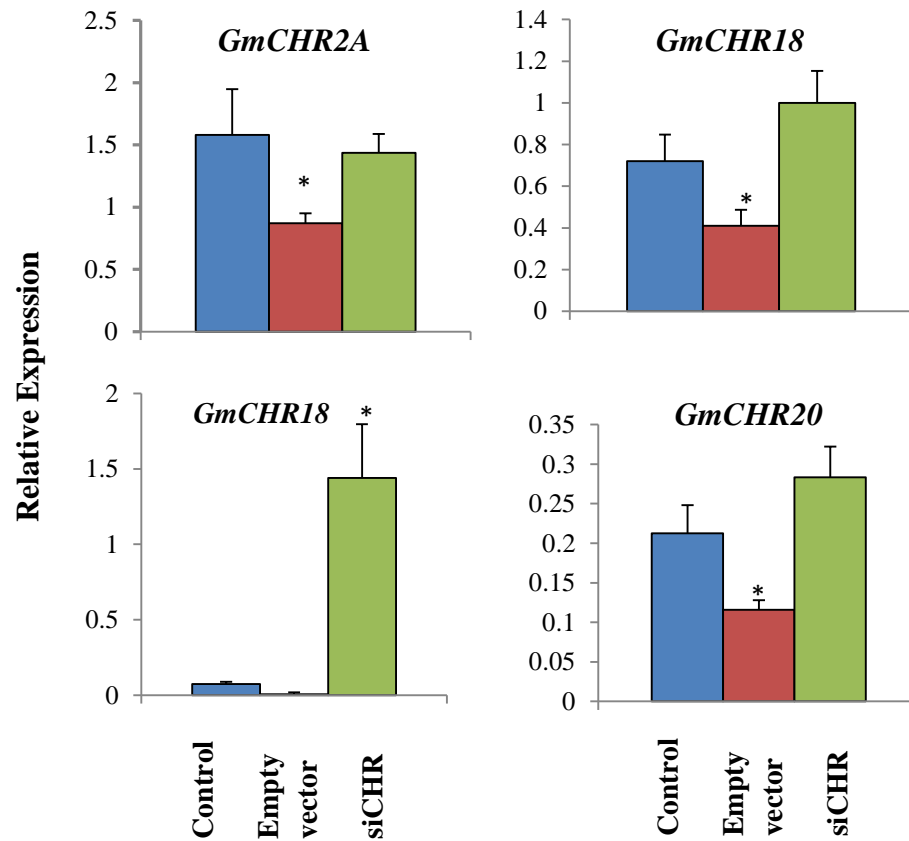


Figure 3.12. Accumulation of root-specific *GmCHR* in soybean cv. Harosoy 63 hairy root system. Total RNA (1 μ g) was used to synthesize cDNA from control (K599 only, blue box), empty vector (silencing vector-only, red box) and silenced lines of either *GmCHR2A*, *GmCHR14*, *GmCHR18* or *GmCHR20* (green box). Expression analysis was conducted by qPCR with *GmCHR* gene-specific primers. Three biological replicates were used. *CON4* were used a reference gene. The asterisks (*) denotes significant expression (multiple t-test per row, $p < 0.05$).

Chapter 4: Discussion

CHRs are legume-specific enzymes, which together with CHS, convert *p*-coumaroyl-CoA and 3 molecules of malonyl-CoA to isoliquiritigenin chalcone, the building block of two core isoflavone aglycones, glycitein and daidzein (Dhaubhadel et al., 2003). The aglycone, daidzein serves as a precursor in the production of phytoalexin glyceollin in soybean. Many studies have reported the expression patterns of genes during infection and the heritability of resistance, however little is known about the importance of *CHR*, the first key enzyme, which directs the flux to the production of phytoalexin glyceollins in soybean. Here I report the identification all putative *GmCHRs* in soybean, investigate their subcellular location, and tissue-specific and pathogen induced gene expression. My results demonstrate that the root-specific *GmCHRs* are induced upon pathogen infection and are located near QTLs and QTL markers linked to *P. sojae* resistance traits.

4.1 Soybean genome contains 16 putative *GmCHR* genes

All members of the *GmCHR* gene family were identified by searching the annotated soybean genome on Phytozome *G. max* Wm82.a2.v1. Using a keyword search together with BLAST searches, 68 *GmAKRs* and 16 putative *GmCHRs* were identified. This total number may change as more aldo-keto reductase sequences are deposited into the database. Many CHR-like enzymes have been reported in a variety of leguminous plants including *M. sativa* (Ballance and Dixon, 1995), *S. rostrata* (Goormachtig et al., 1999), *P. montana* var. *lobata* (He et al., 2011), *G. glabra* (Hayashi et al., 2013), and *L. japonicus* (Shimada et al., 2006). Contrary to my findings, a recent study identified only 2 *CHR* genes in soybean. This study used only one sequence from GenBank (accession number EU921437) to search for the soybean genome database, and identified the two

paralog *GmCHRs*, *GmCHR2A* and *GmCHR14* (Chu et al., 2014). The approach used in my study was more robust and provides confidence as it first identified all the *GmAKRs* based on the current database annotation, and then segregated the *GmCHRs* gene family members using their phylogenetic relationship with other known *CHRs*. The large number of *GmCHRs* could be the result of duplication events in the soybean genome. Soybean is a palaeopolyploid with a genome size of approximately 978 megabase pairs that has undergone at least two whole genome duplications. As a result of the genome duplications, nearly 75% of soybean genes are present in multiple copies. Phylogenetic analysis of *GmCHRs* illustrated that most *GmCHRs* cluster in pairs (Figure 3.1), further supporting ancient genome duplication events (Schmutz et al., 2010).

Out of 16 putative *GmCHRs*, 11 were found to contain conserved critical residues (Figure 3.2). Since *CHRs* are part of the *AKR* family, these enzymes must contain the catalytic site (Bomati et al., 2005). The five *GmCHRs*: *GmCHR2B*, *GmCHR9B*, *GmCHR9E*, *GmCHR16C* and *GmCHR7*, lack one or more catalytic site residues, therefore, I eliminated them from the study. However, it is possible that they may possess weak enzymatic activity or may be evolving new catalytic features.

4.2 *GmCHRs* show tissue-specific expression patterns

GmCHR transcript accumulations showed expression in various types of soybean tissues. The majority of the *GmCHRs* were either expressed in the seedlings, roots or dry seeds. In addition, from the RNA-seq data it was found that only 9 *GmCHRs* were transcribed. Since there was no RNA-seq reads found in the database search for *GmCHR9B* and *GmCHR9D*, I conclude that they may be pseudogenes. Previously, it was found that *CHR* in soybean were moderately expressed in the flowers and weakly expressed in leaves, stems, roots, endosperms and embryos (Liu,

2009). Differential expression of *CHRs* have also been studied in other plant species such as *Astragalus membranaceus* (Xu et al., 2012) and *P. montana* var. *lobata* (He et al., 2011). It was found that *CHRs* from *Astragalus membranaceus* and *P. montana* var. *lobata* were highly expressed in roots and stem, respectively. Based on the soybean RNA-seq data, transcript accumulations from these two plants, I can infer that the isoflavonoid pathway is active in those organs. Out of nine *GmCHRs* identified, *GmCHR2A*, *GmCHR14*, *GmCHR18* and *GmCHR20* transcript accumulation was much higher in root tissue compared to other tissues under study (Figure 3.3). Studies have shown that *CHS7*, *CHS8* (Yi et al., 2010), *IFS1*, *IFS2* (Dhaubhadel et al., 2003) and *CHI* (Dastmalchi and Dhaubhadel, 2015) are also expressed in the roots which infers that they assist in the role of root-specific phytoalexins production. Therefore, I chose these 4 *GmCHRs* for further characterization as they may play a role in resistance to root and stem rot disease caused by *P. sojae*.

4.3 GmCHR localizes in the nucleus and cytoplasm

All members of the *GmCHR* family displayed nuclear and cytoplasmic localization in *N. benthamiana* leaf epidermal cells (Figure 3.4). These findings are consistent with the localization of *GmCHS* which works together with *GmCHR* to produce deoxychalcone. Evidence has shown that 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.

This is a thought-provoking result since only *GmCHR14* is localized in the nucleus. It is possible that there is a carrier protein involved that takes *GmCHRs* to nucleus. Molecules that are smaller than 40 kDa are able to passively diffuse into the nuclear pore complex whereas, larger

molecules (60kDa and higher) require an active transport system, mediated by transport receptors (Fried and Kutay, 2003; D'Angelo et al., 2009). The average molecular weight of GmCHR-YFP is 63 kDa, therefore, it is possible that the fusion protein can go to nucleus by the piggy-back mechanism. Cargo proteins can temporarily bind with other transport proteins, and the complex translocates into or out of the nucleus with the help of importins (Harel and Forbes, 2004).

4.4 Root-specific *GmCHR*s are induced upon stress

To evaluate if *GmCHR* family members respond differently upon pathogen infection, their expression levels were studied by RT-PCR at various time points after *P. sojae* infection. Interestingly, the expression levels of only root-specific *GmCHR*s, *GmCHR2A*, *GmCHR14*, *GmCHR18* and *GmCHR20*, were induced upon infection suggesting that they have a role in defence against *P. sojae* infection (Figure 3.6).

Upon infection, *P. sojae* releases elicitors which stimulate the plant defense response (Jones and Dangl, 2006). As a result, the plant induces the expression of resistance and defense related genes to counteract infection. Studies have shown the upregulation or induction of *CHR*s at the infection site during *Fusarium* attack in soybeans (Iqbal et al., 2005), cadmium treatment in *Medicago truncatula* (Aloui et al., 2012), and *Colletrichum falactum* infection in sugarcane (Selvaraj et al., 2014).

Several studies have used the AgNO₃ treatment to mimic pathogen infection and induce phytoalexin production in soybeans (Moy et al., 2004; Kubeš et al., 2014). The mechanism of this “elicitor effect” is not completely understood, however, there are two possible hypotheses. Firstly, some metals such as quicksilver can affect protective “anti-reactive oxygen species”

(anti-ROS) enzymes (Mithöfer et al., 2004). ROS can participate in the creation of oxylipids which can induce the defense response. Park et al. (2009) mentioned that silver ions can also participate in ROS production. Secondly, silver ions could also block the activity of the plant hormone ethylene. This inhibition could increase the synthesis of secondary metabolites (Zhang and Wu, 2003). The quantitative analysis of root-specific *GmCHR*s in response to AgNO₃ demonstrated a significant increase in transcript accumulation of *GmCHR2A*, *GmCHR14* and *GmCHR18* (Figure 3.8). These root-specific *GmCHR*s respond to the AgNO₃ treatment as early as 12 hours which coincides with findings from Alkharouf et al. (2006). Changes in gene expression within roots upon *Heterodera glycines* (the soybean cyst nematode) attack was investigated using a 6000 gene microarray. It was found that CHR (Genbank BM108162) was induced as soon as 6 and 12 hours upon infection.

4.5 QTLs linked to *P. sojae* resistance contain *GmCHR* loci

Several studies have suggested that *Rps* genes (Dorrance and Schmitthenner, 2000; Sandhu et al., 2005), isoflavonoid biosynthetic genes (Subramanian et al., 2005; Graham et al., 2007), and genes involved in suberin production contribute to *P. sojae* resistance in soybean (Ranathunge et al., 2008). Furthermore, there are QTLs and QTL markers linked to *P. sojae* resistance in soybean. However, little has been reported linking the candidate genes with QTLs and phenotype. Here, I have identified a total of 6 QTL markers and 3 QTLs which flank or are approximate to *GmCHR2A*, *GmCHR16A*, *GmCHR18* and *GmCHR20*. Among the 3 QTLs, Phytoph 14-3 and Phytoph 8-2 cover most of the chromosome (Figure 3.9, Table 3.3). QTL regions can generally span several megabase pairs and can contain several hundreds to thousands of genes (Dupuis and Siegmund, 1999). For those two QTLs, further fine-mapping experiments are required to hone the spanning regions (Touzet et al., 1995; Holtan and Hake, 2003).

Statistical association or validation studies can confirm the co-segregation of genes with a QTL. Based on these parameters, QTL Phytoph10-3 is the most reliable QTL found in this study as it is confined to a specific region on the chromosome and contains the highest LOD score (Table 3.3).

Upon finding a QTL which contain the *GmCHR2A* loci, the parental cultivars were assessed for their root-specific *GmCHR* gene expression. Interestingly, expression of *GmCHR2A*, *GmCHR14* and *GmCHR18* were significantly higher in Conrad (resistant cultivar) as compared to OX760-6 (susceptible cultivar) (Figure 3.10). Soybean cv. Conrad and OX760-6 were also used to study suberin content with *P. sojae* infections (Ranathunge et al., 2008). Therefore, it is possible that *P. sojae* resistance in soybean may comprise both isoflavonoid and suberin content.

4.6 Knockdown of root-specific *GmCHR*s

To further study the importance of *GmCHR* in *P. sojae* resistance, silencing of root-specific *GmCHR*s via hairy root transformation was performed. Following transformation, validation of silencing through analysis of *GmCHR* expression showed increased expression of these genes, which is a surprising result (Figure 3.12). I speculate that the RNAi construct did not lead mRNA degradation but rather block the 5' UTR and start codon region for protein translation. If that occurred, it can trigger a feedback loop leading to increased transcription, and in turn, be revealed through qPCR as increased expression. This premise is based on the RNA activation, where endogenous small RNA fragments are shown to upregulate target genes for overexpression studies (Jiao and Slack, 2014). However, there are alternate ways to explore and assign putative gene functions such as virus induced gene silencing (VIGS) (Zhang and Ghabrial, 2006) and clustered regularly interspaced short palindromic repeats (CRISPR) system (Jacobs et

al., 2015). These are novel techniques and have shown tremendous success in complex plants such as soybeans.

Chapter 5: Conclusions and future directions

Soybeans have emerged as a key component in both human and livestock nutrition as well as in industrial products. It is one of the leading crops which drive economical profits over billions of dollars in sales, worldwide. However, \$50 million worth of yields are lost due to stem and root rot disease caused by *P. sojae*. Taking advantage of the plant defense system, breeders and scientists can create a cultivar which can naturally fight off the pathogen attack.

Identification of the *CHR* gene family members in soybean is the first step in exploring potential gene candidates for developing an effective cultivar. Overall, I have identified 16 putative GmCHRs using an *in silico* approach: GmCHR2A, GmCHR2B, GmCHR9A, GmCHR9B, GmCHR9C, GmCHR9D, GmCHR9E, GmCHR7, GmCHR12, GmCHR14, GmCHR15, GmCHR16A, GmCHR16B, GmCHR16C, GmCHR18, and GmCHR20. Proteins sequences of these GmCHRs were manually searched for critical amino acids based on *M. sativa* CHR (GenBank accession number AAB41555.1) and upon the search 5 GmCHR were eliminated: GmCHR2B, GmCHR9B, GmCHR9E GmCHR16C and GmCHR7. RNA-seq data was searched for the reminder 11 *GmCHRs*. No RNA-seq data was found for GmCHR9A and GmCHR9D, suggesting that they can be pseudogenes.

Seven *GmCHRs* were cloned into pEarleyGate101 to create a C-terminal fusion with YFP to determine the subcellular localization. Despite of several attempts, amplification of *GmCHR15* and *GmCHR12* was not successful. All 7 GmCHRs localize in the nucleus and cytoplasm. The nuclear localization could be a result of passive diffusion due to the small molecular weight of the recombinant GmCHR-YFP protein. An alternative approach to test the possibility of passive diffusion is through photoconvertible fluorescent proteins (Nienhaus et al., 2006).

Photoconvertible proteins are useful tools in monitoring molecular and cellular dynamics and will change their fluorophore colour from green to red in response to irradiation from UV light. In this case, new constructs can be generated with a C-terminal fusion of EosFP instead of YFP. Upon the subcellular localization display, a region EosFP can be excited from green to red and, proteins can be monitored for their nuclear localization and rate of diffusion.

GmCHRs display tissue-specific expression patterns giving the possibility of tissue-specific immunity. Since *P. sojae* attacks at the roots, only the root-specific *GmCHR2A*, *GmCHR14*, *GmCHR18* and *GmCHR20* were monitored in this study. Additionally, those same four are induced in both *P. sojae* and AgNO₃ treatments. Future enzymatic assays can ensure which root-specific GmCHR can efficiently create isoliquiritigenin chalcone.

Studies have identified several QTLs and QTL markers that link to *P. sojae* resistance in soybean. Upon sifting the literature, three potential QTL which contains the *GmCHR* loci has been found. This gives rise to the possibility that *GmCHRs* are an important component developing an effective cultivar against *P. sojae*. The parental lines from that particular QTL were accessed for their root-specific gene expression. It was found that expression of *GmCHR2A*, *GmCHR14* and *GmCHR18* were significantly higher in the Conrad (*P. sojae* resistant cultivar) compared to the OX760-6 (*P. sojae* susceptible cultivar).

RNAi silencing of the root-specific *GmCHRs* was unsuccessful in the present study. Recently virus induced gene silencing (VIGS) have been proven to be an effective method of silencing especially for complex plants such as soybeans and common beans.

References

- Akada S, Dube SK** (1995) Organization of soybean chalcone synthase gene clusters and characterization of a new member of the family. *Plant Molecular Biology* **29**: 189-199
- Alkharouf N, Klink V, Chouikha I, Beard H, MacDonald M, Meyer S, Knap H, Khan R, Matthews B** (2006) Timecourse microarray analyses reveal global changes in gene expression of susceptible *Glycine max* (soybean) roots during infection by *Heterodera glycines* (soybean cyst nematode). *Planta* **224**: 838-852
- Aloui A, Dumas-Gaudot E, Daher Z, van Tuinen D, Aschi-Smit S, Morandi D** (2012) Influence of arbuscular mycorrhizal colonisation on cadmium induced *Medicago truncatula* root isoflavonoid accumulation. *Plant Physiology and Biochemistry* **60**: 233-239
- Anguraj Vadivel A, Sukumaran A, Li X, Dhaubhadel S** (2015) Soybean isoflavonoids: role of GmMYB176 interactome and 14-3-3 proteins. *Phytochemistry Reviews*: 1-13
- Arnason JT, Bernardis MA** (2010) Impact of constitutive plant natural products on herbivores and pathogens. *Canadian Journal of Zoology* **88**: 615-627
- Ayabe S-I, Udagawa A, Furuya T** (1988) NAD(P)H-dependent 6'-deoxychalcone synthase activity in *Glycyrrhiza echinata* cells induced by yeast extract. *Archives of Biochemistry and Biophysics* **261**: 458-462
- Ballance GM, Dixon RA** (1995) *Medicago sativa* cDNAs encoding chalcone reductase. *Plant Physiology* **107**: 1027-1028

- Bechtold N, Ellis J, Pelletier G** (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *Methods in Molecular Biology* **316**: 1194-1199
- Bomati EK, Austin MB, Bowman ME, Dixon RA, Noel JP** (2005) Structural elucidation of chalcone reductase and implications for deoxychalcone biosynthesis. *Journal of Biological Chemistry* **280**: 30496-30503
- Bradford MM** (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
- Burnham K, Dorrance A, VanToai T, St Martin S** (2003) Quantitative trait loci for partial resistance to in soybean. *Crop Science* **43**: 1610-1617
- Byamukama E** (2013) *Phytophthora* root & stem rot appearing In soybeans. *In*, <http://igrow.org/agronomy/soybeans/phytophthora-root-stem-rot-appearing-in-soybeans/>
- Candeia RA, Silva MCD, Carvalho Filho JR, Brasilino MGA, Bicudo TC, Santos IMG, Souza AG** (2009) Influence of soybean biodiesel content on basic properties of biodiesel–diesel blends. *Fuel* **88**: 738-743
- Cederroth CR, Nef S** (2009) Soy, phytoestrogens and metabolism: a review. *Molecular and Cellular Endocrinology* **304**: 30-42
- CFIA** (2015) Varieties of Crop Kinds Registered in Canada. *In*, http://www.inspection.gc.ca/active/netapp/regvar/regvar_lookupe.aspx

- Chu S, Wang J, Cheng H, Yang Q, Yu D** (2014) Evolutionary study of the isoflavonoid pathway based on multiple copies analysis in soybean. *BMC Genetics* **15**: 76
- Cooke D, Drenth A, Duncan J, Wagels G, Brasier C** (2000) A molecular phylogeny of *Phytophthora* and related oomycetes. *Fungal Genetics and Biology* **30**: 17-32
- Council CS** (2012) Canadian soybean council industry statistics. *In*, Vol 2014, <http://www.soybeancouncil.ca/IndustryStatistics/tabid/200/language/en-US/Default.aspx>
- D'Angelo MA, Raices M, Panowski SH, Hetzer MW** (2009) Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. *Cell* **136**: 284-295
- Dakora FD, Phillips DA** (1996) Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins. *Physiological and Molecular Plant Pathology* **49**: 1-20
- Darvasi A** (1998) Experimental strategies for the genetic dissection of complex traits in animal models. *Nature Genetics* **18**: 19-24
- Dastmalchi M, Dhaubhadel S** (2015) Soybean chalcone isomerase: evolution of the fold, and the differential expression and localization of the gene family. *Planta* **241**: 507-523
- De Coninck B, Timmermans P, Vos C, Cammue BPA, Kazan K** (2015) What lies beneath: belowground defense strategies in plants. *Trends in Plant Science* **20**: 91-101

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

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

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

Dixon RA (2004) Phytoestrogens. *Annual Review of Plant Biology* **55**: 225-261

Dorff E (2009) The soybean, agriculture's jack-of-all-trades, is gaining ground across Canada. Canadian Agriculture at a Glance. In <http://www.statcan.gc.ca/pub/96-325-x/2007000/article/10369-eng.htm>, Statistics Canada Catalogue no. 96-325-

XIE

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

Dorrance A, Schmitthenner A (2000) New sources of resistance to *Phytophthora sojae* in the soybean plant introductions. *Plant Disease* **84**: 1303-1308

Dupuis J, Siegmund D (1999) Statistical methods for mapping quantitative trait loci from a dense set of markers. *Genetics* **151**: 373-386

- Enkerli K, Mims CW, Hahn MG** (1997) Ultrastructure of compatible and incompatible interactions of soybean roots infected with the plant pathogenic oomycete *Phytophthora sojae*. Canadian Journal of Botany **75**: 1493-1508
- Erwin DC, Ribeiro OK** (1996) Phytophthora diseases worldwide. American Phytopathological Society (APS Press)
- FAOSTAT** (2015) Food and agriculture organization of united nation statistics division. In, <http://faostat3.fao.org/home/E>
- Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P** (2008) Land clearing and the biofuel carbon debt. Science **319**: 1235-1238
- Fried H, Kutay U** (2003) Nucleocytoplasmic transport: taking an inventory. Cellular and Molecular Life Sciences CMLS **60**: 1659-1688
- Gao H, Narayanan NN, Ellison L, Bhattacharyya MK** (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
- Geldner N** (2013) The endodermis. Annual Review of Plant Biology **64**: 531-558
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS** (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research **40**: D1178-1186

- Goormachtig S, Lievens S, Herman S, Van Montagu M, Holsters M** (1999) Chalcone reductase-homologous transcripts accumulate during development of stem-borne nodules on the tropical legume *Sesbania rostrata*. *Planta* **209**: 45-52
- Gordon SG, St. Martin SK, Dorrance AE** (2006) 8 maps to a resistance gene rich region on soybean molecular linkage Group F. *Crop Science* **46**: 168-173
- Graham TL, Graham MY, Subramanian S, Yu O** (2007) RNAi silencing of genes for elicitation or biosynthesis of 5-deoxyisoflavonoids suppresses race-specific resistance and hypersensitive cell death in *Phytophthora sojae* infected tissues. *Plant Physiology* **144**: 728-740
- Gutierrez-Gonzalez JJ, Wu X, Zhang J, Lee J-D, Ellersieck M, Shannon JG, Yu O, Nguyen HT, Sleper DA** (2009) Genetic control of soybean seed isoflavone content: importance of statistical model and epistasis in complex traits. *Theoretical and Applied Genetics* **119**: 1069-1083
- Habereder H, Schröder G, Ebel J** (1989) Rapid induction of phenylalanine ammonia-lyase and chalcone synthase mRNAs during fungus infection of soybean (*Glycine max* L.) roots or elicitor treatment of soybean cell cultures at the onset of phytoalexin synthesis. *Planta* **177**: 58-65
- Han Y, Teng W, Yu K, Poysa V, Anderson T, Qiu L, Lightfoot D, Li W** (2008) Mapping QTL tolerance to *Phytophthora* root rot in soybean using microsatellite and RAPD/SCAR derived markers. *Euphytica* **162**: 231-239

- Harel A, Forbes DJ** (2004) Importin Beta: conducting a much larger cellular symphony. *Molecular Cell* **16**: 319-330
- Hayashi H, Fujii I, Iinuma M, Shibano M, Ozaki K, Watanabe H** (2013) Characterization of a *Glycyrrhizin*-Deficient Strain of *Glycyrrhiza uralensis*. *Biological and Pharmaceutical Bulletin* **36**: 1448-1453
- He X, Blount J, Ge S, Tang Y, Dixon R** (2011) A genomic approach to isoflavone biosynthesis in kudzu (*Pueraria lobata*). *Planta* **233**: 843-855
- Holtan HEE, Hake S** (2003) Quantitative trait locus analysis of leaf dissection in tomato using *Lycopersicon pennellii* segmental introgression lines. *Genetics* **165**: 1541-1550
- Huston M** (2015) SoyCanada. In Manitoba Pulse & Soybean Growers, <http://soycanada.ca/>
- Iqbal MJ, Yaegashi S, Ahsan R, Shopinski K, Lightfoot D** (2005) Root response to *Fusarium solani* f. sp . *glycines*: temporal accumulation of transcripts in partially resistant and susceptible soybean. *Theoretical and Applied Genetics* **110**: 1429-1438
- Jacobs TB, LaFayette PR, Schmitz RJ, Parrott WA** (2015) Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnology* **15**: 16
- Jez JM, Bennett MJ, Schlegel BP, Lewis M, Penning TM** (1997) Comparative anatomy of the aldo-keto reductase superfamily. *Biochemistry Journal* **326 (Pt 3)**: 625-636
- Jez JM, Flynn TG, Penning TM** (1997) A new nomenclature for the aldo-keto reductase superfamily. *Biochemical Pharmacology* **54**: 639-647
- Jiao AL, Slack FJ** (2014) RNA-mediated gene activation. *Epigenetics* **9**: 27-36

Jones JD, Dangl JL (2006) The plant immune system. *Nature* **444**: 323-329

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

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

Kereszt A, Li D, Indrasumunar A, Nguyen CDT, Nontachaiyapoom S, Kinkema M, Gresshoff PM (2007) *Agrobacterium rhizogenes*-mediated transformation of soybean to study root biology. *Nature Protocols* **2**: 948-952

Kombrink E, Somssich IE (1995) Defense responses of plants to pathogens. *Advances in Botanical Research* **21**: 2-34

Korde LA, Wu AH, Fears T, Nomura AM, West DW, Kolonel LN, Pike MC, Hoover RN, Ziegler RG (2009) Childhood soy intake and breast cancer risk in Asian American women. *Cancer Epidemiology Biomarkers & Prevention* **18**: 1050-1059

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

Latijnhouwers M, de Wit PJ, Govers F (2003) Oomycetes and fungi: similar weaponry to attack plants. *Trends in Microbiology* **11**: 462-469

- Lee S, Mian MAR, McHale LK, Wang H, Wijeratne AJ, Sneller CH, Dorrance AE** (2013) Novel quantitative trait loci for partial resistance to *Phytophthora sojae* in soybean PI 398841. *Theoretical and Applied Genetics*. **126**: 1121-1132
- Li XP, Han YP, Teng WL, Zhang SZ, Yu KF, Poysa V, Anderson T, Ding JJ, Li WB** (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
- Li Z-B, Li C-F, Li J, Zhang Y-S** (2014) Molecular cloning and functional characterization of two divergent 4-Coumarate:Coenzyme A Ligases from Kudzu (*Pueraria lobata*). *Biological and Pharmaceutical Bulletin* **37**: 113-122
- Liu GY** (2009) Isolation, sequence identification and tissue expression profile of two novel soybean (*Glycine max*) genes-vestitone reductase and chalcone reductase. *Molecular Biology Reports* **36**: 1991-1994
- Lozovaya VV, Lygin AV, Zernova OV, Ulanov AV, Li S, Hartman GL, Widholm JM** (2007) Modification of phenolic metabolism in soybean hairy roots through down regulation of chalcone synthase or isoflavone synthase. *Planta* **225**: 665-679
- Mithöfer A, Schulze B, Boland W** (2004) Biotic and heavy metal stress response in plants: evidence for common signals. *FEBS Letters* **566**: 1-5
- Morris PF, Ward E** (1992) Chemoattraction of zoospores of the soybean pathogen, *Phytophthora sojae*, by isoflavones. *Physiological and Molecular Plant Pathology* **40**: 17-22

- Moy P, Qutob D, Chapman BP, Atkinson I, Gijzen M** (2004) Patterns of gene expression upon infection of soybean plants by *Phytophthora sojae*. *Molecular Plant-Microbe Interactions* **17**: 1051-1062
- Murdoch M** (2012) Government of Canada invests to help farmers boost soybean exports. In http://www.agr.gc.ca/cb/index_e.php?s1=n&s2=2012&page=n120312, ed, Vol 2014, Agriculture and Agri-Food Canada
- Nguyen V, Vuong T, VanToai T, Lee J, Wu X, Mian M, Dorrance A, Shannon J, Nguyen H** (2012) Mapping of quantitative trait loci associated with resistance to and flooding tolerance in soybean. *Crop Science* **52**: 2481-2493
- Nienhaus GU, Nienhaus K, Hölzle A, Ivanchenko S, Renzi F, Oswald F, Wolff M, Schmitt F, Röcker C, Vallone B, Weidemann W, Heilker R, Nar H, Wiedenmann J** (2006) Photoconvertible Fluorescent Protein EosFP: Biophysical Properties and Cell Biology Applications. *Photochemistry and Photobiology* **82**: 351-358
- Outchkourov NS, Rogelj B, Strukelj B, Jongsma MA** (2003) Expression of sea anemone equistatin in potato effects of plant proteases on heterologous protein production. *Plant Physiology* **133**: 379-390
- Park H-J, Kim JY, Kim J, Lee J-H, Hahn J-S, Gu MB, Yoon J** (2009) Silver-ion-mediated reactive oxygen species generation affecting bactericidal activity. *Water Research* **43**: 1027-1032

- Pedras MSC, Hossain S** (2011) Interaction of cruciferous phytoanticipins with plant fungal pathogens: Indole glucosinolates are not metabolized but the corresponding desulfo-derivatives and nitriles are. *Phytochemistry* **72**: 2308-2316
- Pflieger S, Lefebvre V, Causse M** (2001) The candidate gene approach in plant genetics: a review. *Molecular Breeding* **7**: 275-291
- Ranathunge K, Thomas RH, Fang X, Peterson CA, Gijzen M, Bernards MA** (2008) Soybean root suberin and partial resistance to root rot caused by *Phytophthora sojae*. *Phytopathology* **98**: 1179-1189
- Rommens JM, Iannuzzi MC, Kerem B-s, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N** (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* **245**: 1059-1065
- Sandhu D, Schallock K, Rivera-Velez N, Lundeen P, Cianzio S, Bhattacharyya M** (2005) Soybean *Phytophthora resistance* gene *Rps8* maps closely to the *Rps3* region. *Journal of Heredity* **96**: 536-541
- Schmitthenner A** (1999) *Phytophthora* rot of soybean. *Compendium of Soybean Diseases*. 4th edn., The American Phytopathological Society Press, St. Paul, Minnesota: 39-42
- Schmitthenner AF** (1985) Problems and progress in control of *Phytophthora* root rot of soybean. *Plant Disease* **69**: 362-368
- Schmitthenner AF** (2001) *Phytophthora* rot of soybean. *In*, <http://www.plantmanagementnetwork.org/pub/php/management/phytophthora/>

- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA** (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178-183
- Selvaraj N, Ramadass A, Amalraj R, Palaniyandi M, Rasappa V** (2014) Molecular profiling of systemic acquired resistance (SAR)-responsive transcripts in sugarcane challenged with *Colletotrichum falcatum*. *Applied Biochemistry and Biotechnology* **174**: 2839-2850
- Shimada N, Nakatsuka T, Nishihara M, Yamamura S, Ayabe S-i, Aoki T** (2006) Isolation and characterization of a cDNA encoding polyketide reductase in *Lotus japonicus*. *Plant Biotechnology* **23**: 509-513
- Sparkes IA, Runions J, Kearns A, Hawes C** (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols* **1**: 2019-2025
- Strom BL, Schinnar R, Ziegler EE, et al.** (2001) Exposure to soy-based formula in infancy and endocrinological and reproductive outcomes in young adulthood. *The Journal of the American Medical Association* **286**: 807-814

- Subramanian S, Graham MY, Yu O, Graham TL** (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
- Sun J, Guo N, Lei J, Li L, Hu G, Xing H** (2014) Association mapping for partial resistance to *Phytophthora sojae* in soybean (*Glycine max* (L.) Merr.). *Journal of Genetics* **93**: 355-363
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S** (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**: 2725-2729
- Thomas R, Fang X, Ranathunge K, Anderson TR, Peterson CA, Bernards MA** (2007) Soybean root suberin: anatomical distribution, chemical composition, and relationship to partial resistance to *Phytophthora sojae*. *Plant Physiology* **144**: 299-311
- Tooley P, Grau C** (1984) Field characterization of rate-reducing resistance to *Phytophthora megasperma* f. sp. *glycinea* in soybean. *Phytopathology* **74**: 1201-1208
- Touzet P, Winkler R, Helentjaris T** (1995) Combined genetic and physiological analysis of a locus contributing to quantitative variation. *Theoretical and Applied Genetics* **91**: 200-205
- Tucker DM, Saghai Maroof MA, Mideros S, Skoneczka JA, Nabati DA, Buss GR, Hoeschele I, Tyler BM, St. Martin SK, Dorrance AE** (2010) Mapping quantitative trait loci for partial resistance to in a soybean interspecific cross. *Crop Science* **50**: 628-635

- Tyler BM** (2007) *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. *Molecular Plant Pathology* **8**: 1-8
- Wang H, St Martin SK, Dorrance AE** (2012) Comparison of phenotypic methods and yield contributions of quantitative trait loci for partial resistance to in soybean. *Crop Science* **52**: 609-622
- Wang H, Waller L, Tripathy S, St Martin SK, Zhou L, Krampis K, Tucker DM, Mao Y, Hoeschele I, Maroof S** (2010) Analysis of genes underlying soybean quantitative trait loci conferring partial resistance to *Phytophthora sojae*. *The Plant Genome* **3**: 23-40
- Wang H, Wijeratne A, Wijeratne S, Lee S, Taylor CG, St Martin SK, McHale L, Dorrance AE** (2012) Dissection of two soybean QTL conferring partial resistance to *Phytophthora sojae* through sequence and gene expression analysis. *BMC Genomics* **13**: 428
- Ward E, Lazarovits G, Unwin C, Buzzell R** (1979) Hypocotyl reactions and glyceollin in soybeans inoculated with zoospores of *Phytophthora megasperma* var. *sojae*. *Phytopathology*
- Weng C, Yu K, Anderson T, Poysa V** (2007) A quantitative trait locus influencing tolerance to *Phytophthora* root rot in the soybean cultivar 'Conrad'. *Euphytica* **158**: 81-86
- Wink M** (1988) Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. *Theoretical and Applied Genetics* **75**: 225-233

- Wu X, Zhou B, Zhao J, Guo N, Zhang B, Yang F, Chen S, Gai J, Xing H** (2011) Identification of quantitative trait loci for partial resistance to *Phytophthora sojae* in soybean. *Plant Breeding* **130**: 144-149
- Xu R-Y, Nan P, Pan H, Zhou T, Chen J** (2012) Molecular cloning, characterization and expression of a chalcone reductase gene from *Astragalus membranaceus* Bge. var. *mongholicus* (Bge.) Hsiao. *Molecular Biology Reports* **39**: 2275-2283
- Yi J, Derynck MR, Chen L, Dhaubhadel S** (2010) Differential expression of *CHS7* and *CHS8* genes in soybean. *Planta* **231**: 741-753
- Yoshikawa M, Fujita H, Matoba N, Takenaka Y, Yamamoto T, Yamauchi R, Tsuruki H, Takahata K** (2000) Bioactive peptides derived from food proteins preventing lifestyle-related diseases. *Biofactors* **12**: 143-146
- Zhang C-H, Wu J-Y** (2003) Ethylene inhibitors enhance elicitor-induced paclitaxel production in suspension cultures of *Taxus* spp. cells. *Enzyme and Microbial Technology* **32**: 71-77
- Zhang C, Ghabrial SA** (2006) Development of bean pod mottle virus-based vectors for stable protein expression and sequence-specific virus-induced gene silencing in soybean. *Virology* **344**: 401-411

Appendices

Appendix A. Complete list of GmCHRs, genetic characteristics and reason of elimination

Name Gene	Locus Name	Alias Name (Wm81 a.1 v.1)	Locus Range	Splice variants	Coding sequence (nt)	Protein Molecular Weight (kDa)	Predicted Sub-localization
<i>GmCHR2A</i>	Glyma.02G307300	Glyma02g47750	Chr02: 48163443..48164865	1	948	35.59	Cytoplasm
<i>GmCHR2B</i>	Glyma.02G182500	Glyma02g31440	Chr02: 31249609..31253723	1	978	36.37	Cytoplasm
<i>GmCHR7</i>	Glyma.07G167700	Glyma07g24635	Chr07: 27067671..27075093	1	390	14.62	Cytoplasm
<i>GmCHR9A</i>	Glyma.09G169700	Glyma09g29991	Chr09:3 9437065..9439064	1	846	31.86	Cytoplasm
<i>GmCHR9B</i>	Glyma.09G169800	--	Chr09: 39439452..39441906	1	924	34.51	Cytoplasm
<i>GmCHR9C</i>	Glyma.09G170000	Glyma09g30010	Chr09: 39453516..39456085	1	957	35.48	Cytoplasm
<i>GmCHR9D</i>	Glyma.09G229500	Glyma09g36391	Chr09: 45341437..45343797	1	969	36.49	Cytoplasm
<i>GmCHR9E</i>	Glyma.09G169900	--	Chr09: 39441495..39441821	1	327	11.81	Cytoplasm
<i>GmCHR12</i>	Glyma.12G006700.	Glyma12g00940	Chr12: 508878..511003	1	948	35.85	Chloroplast
<i>GmCHR14</i>	Glyma.14G005700	Glyma14g00870	Chr14: 463009..464543	1	981	36.77	Nucleus
<i>GmCHR15</i>	Glyma.15G189300	Glyma15g21740	Chr15: 20075936..20082802	1	948	35.43	Cytoplasm
<i>GmCHR16A</i>	Glyma.16G219400	Glyma16g34560	Chr16: 37672779..37675776	3	963	36.15	Cytoplasm
<i>GmCHR16B</i>	Glyma.16G219500	Glyma16g34570	Chr16: 37677371..37670638	1	969	36.04	Cytoplasm
<i>GmCHR16C</i>	Glyma.16G219600	Glyma16g34580	Chr16: 37675532..37677780	1	843	30.58	Cytoplasm
<i>GmCHR18</i>	Glyma.18G285800	Glyma18g52250	Chr18: 56611460..56612978	2	948	35.14	Cytoplasm
<i>GmCHR20</i>	Glyma.20G031100	Glyma20g03900	Chr20: 3790362...3793599	1	966	35.91	Cytoplasm

CHR chalcone reductase; nt nucleotide; kDa kilodalton, -- no alias name

Appendix B. Complete Alignment of GmCHRs, CHRs from other plant species and ARKs

2,5-DRG	2,5-DRG
XVR	XVR
ALDH	ALDH
ALR	ALR
GmCHR9D	GmCHR9D
GmCHR12	GmCHR12
GmCHR15	GmCHR15
GmCHR20	GmCHR20
GmCHR18	GmCHR18
CHR P. montana	CHR P. montana
GmCHR24	GmCHR24
GmCHR14	GmCHR14
CHR M. sativa	CHR M. sativa
CHR M. sativa	CHR M. sativa
CHR M. sativa	CHR M. sativa
CHR M. sativa	CHR M. sativa
CHR M. japonicus	CHR M. japonicus
POR G. glabra	POR G. glabra
POR G. glabra	POR G. glabra
GmCHR7	GmCHR7
COR1	COR1P
GmCHR9C	COR2E
GmCHR16C	GmCHR9C
CHR S. rostrata	GmCHR16C
GmCHR	GmCHR9E
GmCHR9B	GmCHR9B
GmCHR16B	GmCHR16B
GmCHR9A	GmCHR9A
GmCHR16A	GmCHR16A

2,5-DRG	2,5-DRG
XVR	XVR
ALDH	ALDH
ALR	ALR
GmCHR9D	GmCHR9D
GmCHR12	GmCHR12
GmCHR15	GmCHR15
GmCHR20	GmCHR20
GmCHR18	GmCHR18
GmCHR24	GmCHR24
GmCHR14	GmCHR14
CHR M. sativa	CHR M. sativa
CHR M. sativa	CHR M. sativa
CHR M. sativa	CHR M. sativa
CHR M. sativa	CHR M. sativa
CHR M. japonicus	CHR M. japonicus
POR G. glabra	POR G. glabra
POR G. glabra	POR G. glabra
GmCHR7	GmCHR7
COR1	COR1
GmCHR9C	GmCHR9C
GmCHR16C	GmCHR16C
CHR S. rostrata	CHR S. rostrata
GmCHR	GmCHR9E
GmCHR9B	GmCHR9B
GmCHR16B	GmCHR16B
GmCHR9A	GmCHR9A
GmCHR16A	GmCHR16A

Continued.....

```
2.5-DKG 173 IRRFQDEGRTESSPTLRRSS----SFRTECHLRRIVYVEFTVQ---
XVR 202 IREERAKQAKGTAAIIEPOSFVEMNGRALINPTPFAHDILKAKYKARRAB---
ALDH 195 IREERAKQAKGTAAIIEPOSFVEMNGRALINPTPFAHDILKAKYKARRAB---
ALR 191 IREERAKQAKGTAAIIEPOSFVEMNGRALINPTPFAHDILKAKYKARRAB---
GmCHR9D 191 IREERAKQAKGTAAIIEPOSFVEMNGRALINPTPFAHDILKAKYKARRAB---
GmCHR12 191 IREERAKQAKGTAAIIEPOSFVEMNGRALINPTPFAHDILKAKYKARRAB---
GmCHR15 191 IREERAKQAKGTAAIIEPOSFVEMNGRALINPTPFAHDILKAKYKARRAB---
GmCHR18 191 IREERAKQAKGTAAIIEPOSFVEMNGRALINPTPFAHDILKAKYKARRAB---
GmCHR20 201 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
CHR P.montana 200 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
GmCHR22 200 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
GmCHR14 186 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
CHR M.sativa 199 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
CHR M.sativa 199 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
CHR M.sativa 199 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
CHR M.sativa 199 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
POR L.japonicus 200 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
POR G.glabra 201 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
POR G.glabra 9  NRRFQEOGEGHSAIPLGAGKSR---LST--AMSPKLTIAIAIGKSAOV---
COR1 198 NRRFQEOGEGHSAIPLGAGKSR---LST--AMSPKLTIAIAIGKSAOV---
COR2 198 NRRFQEOGEGHSAIPLGAGKSR---LST--AMSPKLTIAIAIGKSAOV---
GmCHR9C 198 NRRFQEOGEGHSAIPLGAGKSR---LST--AMSPKLTIAIAIGKSAOV---
GmCHR16C 173 NRRFQEOGEGHSAIPLGAGKSR---LST--AMSPKLTIAIAIGKSAOV---
CHR S.rostrata 199 NRRFQEOGEGHSAIPLGAGKSR---LST--AMSPKLTIAIAIGKSAOV---
GmCHR9E 49  KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
GmCHR9B 199 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
GmCHR16B 188 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
GmCHR9A 200 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---
GmCHR16A 200 KIRFEKQNGIHTAASPFGKESKRL---LQILNIGPKLDAIGKTAQVV---

2.5-DKG 218 IRRFHTOLGEP---P-----KSSID
XVR 260 IRRFHTOLGEP---P-----KSSID
ALDH 249 IRRFHTOLGEP---P-----KSSID
ALR 242 IRRFHTOLGEP---P-----KSSID
GmCHR9D 242 IRRFHTOLGEP---P-----KSSID
GmCHR12 244 IRRFHTOLGEP---P-----KSSID
GmCHR2B 244 KGFNQOOAFFIICRCDLGYEGGFAA---KSSIN
GmCHR15 250 IRRFHTOLGEP---P-----KSSIN
GmCHR20 255 IRRFHTOLGEP---P-----KSSIN
GmCHR18 251 IRRFHTOLGEP---P-----KSSIN
CHR P.montana 249 IRRFHTOLGEP---P-----KSSID
GmCHR2A 250 IRRFHTOLGEP---P-----KSSID
GmCHR14 236 IRRFHTOLGEP---P-----KSSID
CHR2M.sativa 249 IRRFHTOLGEP---P-----KSSID
CHR4M.sativa 249 IRRFHTOLGEP---P-----KSSID
CHR1M.sativa 249 IRRFHTOLGEP---P-----KSSID
CHR3M.sativa 249 IRRFHTOLGEP---P-----KSSID
POR L.japonicus 250 IRRFHTOLGEP---P-----KSSID
POR G.glabra 251 IRRFHTOLGEP---P-----KSSID
POR G.glabra 60  LPKVELIKSKTIVARTASQWGYIRIMAGTFIGLGFV---
COR1 249 IRRFHTOLGEP---P-----KSSIN
COR2 249 IRRFHTOLGEP---P-----KSSIN
GmCHR9C 249 IRRFHTOLGEP---P-----KSSIN
GmCHR16C 224 IRRFHTOLGEP---P-----KSSIN
CHR S.rostrata 250 IRRFHTOLGEP---P-----KSSIN
GmCHR9E 200 IRRFHTOLGEP---P-----KSSIN
GmCHR9B 150 IRRFHTOLGEP---P-----KSSIN
GmCHR16B 150 IRRFHTOLGEP---P-----KSSIN
GmCHR9A 239 IRRFHTOLGEP---P-----KSSIN
GmCHR16A 239 IRRFHTOLGEP---P-----KSSIN
```

```
2.5-DKG 236 PRRVQRADLGGRA---LQVLRVSGRERGWD-----
XVR 278 PRRVQRADLGGRA---LQVLRVSGRERGWD-----
ALDH 267 PRRVQRADLGGRA---LQVLRVSGRERGWD-----
ALR 260 PRRVQRADLGGRA---LQVLRVSGRERGWD-----
GmCHR9D 260 PRRVQRADLGGRA---LQVLRVSGRERGWD-----
GmCHR2B 281 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR15 286 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR20 273 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR18 268 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
CHR P.montana 267 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR22 254 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR14 254 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
CHR M.sativa 267 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
CHR M.sativa 267 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
CHR M.sativa 267 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
POR L.japonicus 268 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
POR G.glabra 268 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
POR G.glabra 101 MRVLRVQELMLKPKPKRERKFRDRA---
GmCHR7 267 EGRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
COR1 267 EGRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
COR2 267 EGRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR9C 242 RR-----ARFIEELGPKRSPD---
GmCHR16C 242 RR-----ARFIEELGPKRSPD---
CHR S.rostrata 268 KRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR9E 118 RRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR9B 268 RRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR16B 267 RRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR9A 269 RRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---
GmCHR16A 269 RRRMQRVQEDRMDLTKTKVLRKPKRMR---GFVVVEITSPVATQ---

2.5-DKG 269 ---GPDTHREEM
XVR 313 IRRVRFIYF---
ALDH 325 IRRVRFIYF---
ALR 311 IRRVRFIYF---
GmCHR9D 311 IRRVRFIYF---
GmCHR12 311 IRRVRFIYF---
GmCHR2B 311 IRRVRFIYF---
GmCHR15 311 IRRVRFIYF---
GmCHR20 312 IRRVRFIYF---
GmCHR18 312 IRRVRFIYF---
CHR P.montana 310 IRRVRFIYF---
GmCHR2A 311 IRRVRFIYF---
GmCHR14 297 IRRVRFIYF---
CHR M.sativa 310 IRRVRFIYF---
CHR M.sativa 310 IRRVRFIYF---
CHR M.sativa 310 IRRVRFIYF---
POR L.japonicus 311 IRRVRFIYF---
POR G.glabra 312 IRRVRFIYF---
POR G.glabra 311 IRRVRFIYF---
GmCHR7 317 IRRVRFIYF---
COR1 317 IRRVRFIYF---
COR2 317 IRRVRFIYF---
GmCHR9C 276 IRRVRFIYF---
GmCHR16C 318 IRRVRFIYF---
CHR S.rostrata 168 IRRVRFIYF---
GmCHR9E 318 IRRVRFIYF---
GmCHR9B 307 IRRVRFIYF---
GmCHR16B 319 IRRVRFIYF---
GmCHR9A 319 IRRVRFIYF---
GmCHR16A 319 IRRVRFIYF---
```



Entrance of the catalytic site



AKR catalytic tetrad



CHR active site with AKR catalytic tetrad



Unique amino acids from COR

Appendix C. Number of reads from RNA-seq data collected from <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29163>.

Gene name	Root	Stem	Leaf	Flower Buds	Seedlings	Globular	Heart	Cotyledon	Early Maturation	Dry Seeds
GmCHR 2A	643	26	13	11	358	0	0	0	101	7
GmCHR 9C	0	0	0	0	0	89	217	66	5	0
GmCHR 12	2	0	0	0	3	0	0	0	1	0
GmCHR 14	3235	1136	17	109	2497	107	14	57	42	55
GmCHR 15	0	0	0	1	3	83	5	1	0	0
GmCHR 16A	492	137	1172	500	1024	355	313	473	381	391
GmCHR 16B	552	1505	9304	9408	4664	5971	3542	3842	5167	5
GmCHR 18	2991	116	185	44	1899	12	1	0	229	29
GmCHR 20	428	167	117	275	256	293	191	187	283	175

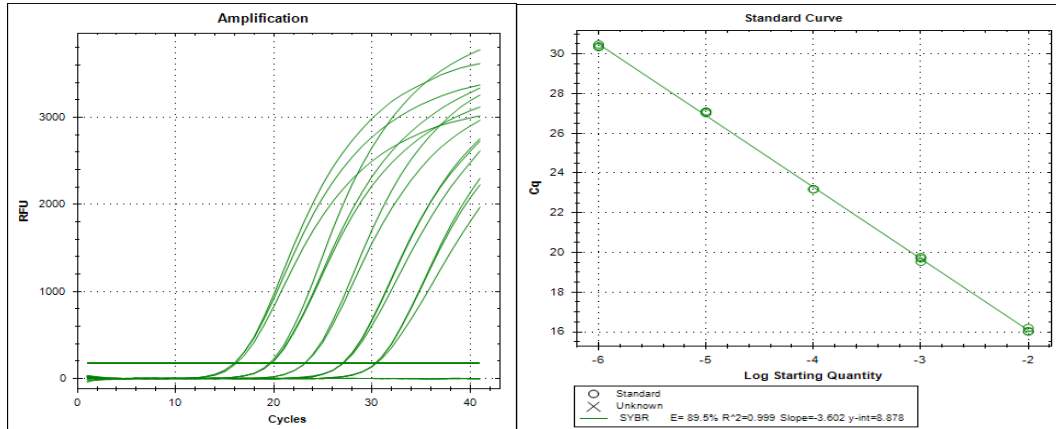
Appendix D. qPCR primer efficiencies for root-specific gene expression for AgNO₃ treatment and silencing

GmCHR2A

qsiCHR2AF: CGG GGA GGT GCT TCT TGG TTA TAG

qsiCHR2AR: CTC CTT AGT GTC TTT CTT GCA

Primer Efficiency: 89.5%

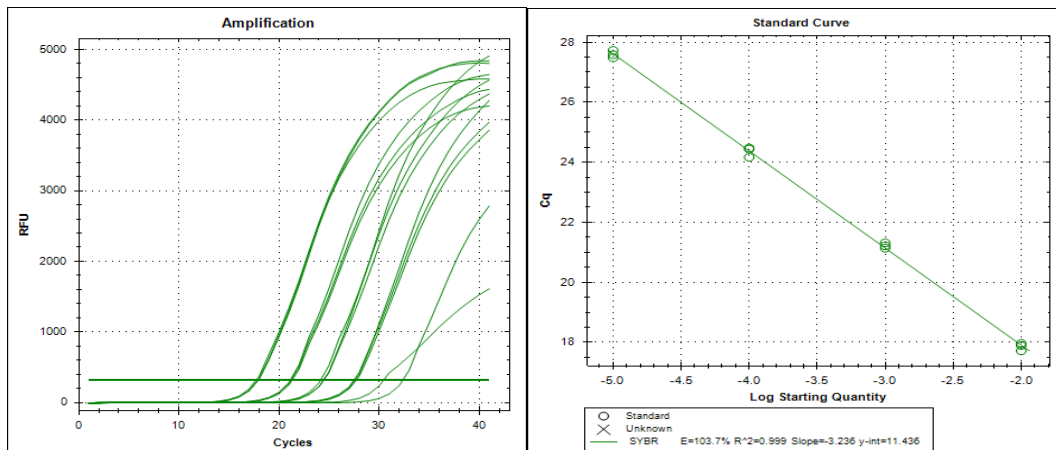


GmCHR14

qiCHR14F: CCA CCT ACC TCC AAT TGC TGA C

qiCHR14R: GGG ATT TCA ATA GCA GCA GCC

Primer Efficiency: 103.7%

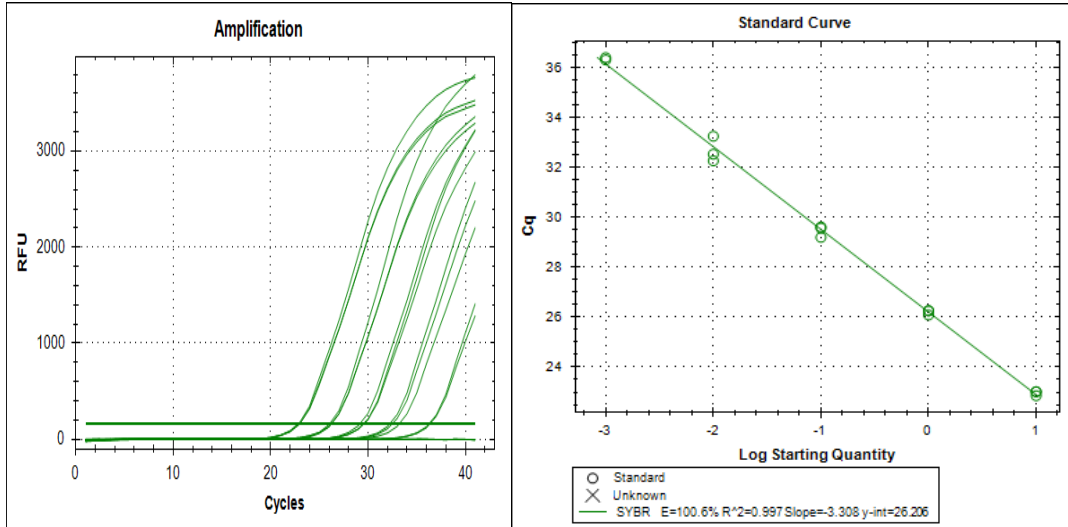


GmCHR18A

qiCHR18F: ATC CCA CCT ACC CAC ATC GT

qiCHR18R: GGA GGA GGG GAG AGT GAC TT

Primer Efficiency: 100.6%

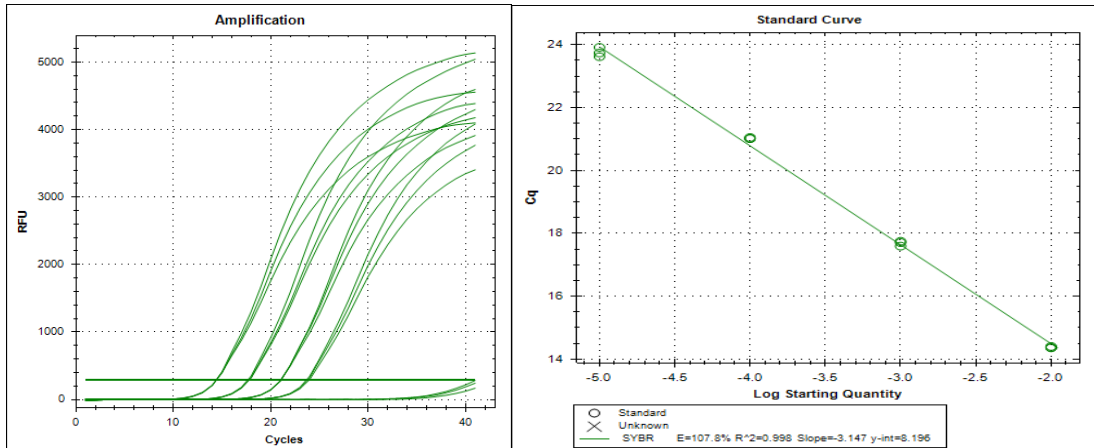


GmCHR20

qsiCHR20F: CAC TCG TAT CTG TGA CAC CGT GT

qsiCHR20R: GTT GTA ACA GCA CTT TGG GGA CGT

Primer Efficiency: 107.8%



Appendix E. Complete list of QTLs found in the SoyBase and Soybean Breeder's Toolbox

QTL	Chromosome	Parents	Reference
Phytoph 1-1	13	Parent 1: Conrad Parent 2: Harosoy	(Tucker et al., 2010)
Phytoph 1-2	2	Parent 1: Conrad Parent 2: Harosoy	(Tucker et al., 2010)
Phytoph 10-1	13	Parent 1: Conrad Parent 2: OX760-6-1	(Han et al., 2008)
Phytoph 10-2	13	Parent 1: Conrad Parent 2: OX760-6-1	(Han et al., 2008)
Phytoph 10-3	2	Parent 1: Conrad Parent 2: OX760-6-1	(Han et al., 2008)
Phytoph 11-19	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-2	8	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-20	17	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-21	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 11-22	14	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 12-1	13	Parent 1: S99-2281 Parent 2: PI 408105A	(Nguyen et al., 2012)
Phytoph 12-2	17	Parent 1: S99-2281 Parent 2: PI 408105A	(Nguyen et al., 2012)
Phytoph 13-1	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 13-2	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 13-3	1	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 13-4	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 13-5	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Phytoph 14-1	1	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-10	20	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-2	13	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-3	18	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)

Phytoph 14-4	2	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-5	3	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-6	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-7	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-8	7	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 14-9	15	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Phytoph 2-1	13	Parent 1: Conrad Parent 2: Sloan	(Burnham et al., 2003)
Phytoph 2-2	2	Parent 1: Conrad Parent 2: Sloan	(Burnham et al., 2003)
Phytoph 3-1	13	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Phytoph 3-2	2	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Phytoph 4-1	13	Parent 1: Conrad Parent 2: Williams	(Burnham et al., 2003)
Phytoph 4-2	2	Parent 1: Conrad Parent 2: Williams	(Burnham et al., 2003)
Phytoph 5-1	6	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Phytoph 5-2	15	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Phytoph 5-3	10	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Phytoph 6-1	13	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-2	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-3	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-4	8	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-5	11	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Phytoph 6-6	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)

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

Appendix F. Complete list of QTL markers found in the literature search

QTL Marker	Chromosome	Parents	Reference
Sat_414	1	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt266	2	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Satt252	13	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Satt579	2	Parent 1 : Conrad Parent 2: Williams	(Burnham et al., 2003)
Satt579	2	Parent 1 : Conrad Parent 2: Williams	(Burnham et al., 2003)
Satt252	13	Parent 1 : Conrad Parent 2: Williams	(Burnham et al., 2003)
Satt252	13	Parent 1: Conrad Parent 2: Sloan	(Burnham et al., 2003)
Satt252	13	Parent 1: Conrad Parent 2: Harosoy	(Burnham et al., 2003)
Satt274	2	Parent 1:Conrad Parent 2:OX760-6-1	(Han et al., 2008)
Satt030	13	Parent 1:Conrad Parent 2:OX760-6-1	(Han et al., 2008)
Satt343	13	Parent 1:Conrad Parent 2:OX760-6-1	(Han et al., 2008)
Satt439	1	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt351	2	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt009	3	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
A078_1	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt578	4	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt463	7	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Sat_133	13	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Satt384	15	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
OPAD08	18	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)
Sat_267	20	Parent 1: OX20-8 Parent 2: PI 398841	(Lee et al., 2013)

Sat_089	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt600	2	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt100	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt307	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt365	6	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt437	8	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt453	11	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Satt343	13	Parent 1: Conrad Parent 2: Hefeng 25	(Li et al., 2010)
Sct_033	13	Parent 1: S99-2281 Parent 2: PI 408105A	(Nguyen et al., 2012)
Satt543	17	Parent 1: S99-2281 Parent 2: PI 408105A	(Nguyen et al., 2012)
Satt634	2	Multiple accession were used	(Sun et al., 2014)
Satt542	2	Multiple accession were used	(Sun et al., 2014)
Satt266	2	Multiple accession were used	(Sun et al., 2014)
Sat_423	2	Multiple accession were used	(Sun et al., 2014)
Satt_222	17	Multiple accession were used	(Sun et al., 2014)
Satt226	17	Multiple accession were used	(Sun et al., 2014)
Satt300	17	Multiple accession were used	(Sun et al., 2014)
Satt574	17	Multiple accession were used	(Sun et al., 2014)
Satt543	17	Multiple accession were used	(Sun et al., 2014)
Satt615	17	Multiple accession were used	(Sun et al., 2014)
Satt301	17	Multiple accession were used	(Sun et al., 2014)
K644_1	13	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
Satt414	16	Parent 1: V71-370	(Tucker et al.,

		Parent 2: PI407162	2010)
Satt678	18	Parent 1: V71-370 Parent 2: PI407162	(Tucker et al., 2010)
A036_1	12	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010 Geno. 2010, 3(1):23-40
Satt160	13	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010 Geno. 2010, 3(1):23-40
Sct_033	13	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010 Geno. 2010, 3(1):23-40
Satt304	14	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010 Geno. 2010, 3(1):23-40
Satt574	17	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010 Geno. 2010, 3(1):23-40
Satt527	19	Parent 1: Conrad Parent 2: Sloan	Wang et al. 2010 Geno. 2010, 3(1):23-40
A671_1	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt472	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARC-021321-04035	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt527	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
GMA2_OSU19	8	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt632	8	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
GMH_OSU31	12	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
F424_294	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt160	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)

Sct_033	13	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt304	14	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt574	17	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARC-039397-07314	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_18_1707	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_18_1777	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1393	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
GM19_OSU10	18	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARC-03997-07624	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARC-064609-18739	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_18_1793	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1243	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1473	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1494	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
BARCSOYSSR_19_1532	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
GML_OSU42	19	Parent 1: Conrad Parent 2: Sloan	(Wang et al., 2012)
Satt596	16	Parent 1: Conrad Parent 2: OX760-6-1	(Weng et al., 2007)
Satt520	6	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Sat_274	10	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
Satt651	15	Parent 1: Su88-M21(S) Parent 2: Xinyixiaoheidou (X)	(Wu et al., 2011)
AW734043	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)

Satt322	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt520	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt557	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt100	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt079	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt307	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt316	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt376	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Staga001	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_251	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_246	6	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt358	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_321	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt445	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt241	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt345	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt094	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt550	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt576	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt479	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt188	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt420	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_274	10	Parent 1 : Su88-M21	(Wu et al., 2011)

		Parent 2 : Xinyixiaoheidou	
Sat_242	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt592	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_196	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_273	10	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt369	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt204	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt491	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt267	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Sat_380	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt598	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt651	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt384	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt720	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)
Satt403	15	Parent 1 : Su88-M21 Parent 2 : Xinyixiaoheidou	(Wu et al., 2011)

Curriculum Vitae

Caroline Sepiol

EDUCATION

2013-present	M.Sc candidate Biology	The University of Western Ontario London, Ontario, Canada
2009-2013	B.Sc (Honors Specialization) Genetics and Biochemistry	The University of Western Ontario London, Ontario, Canada

AWARDS AND SCHOLARSHIPS

2013-present	Western Graduate Research Scholarship
2009	James R. Hoffa Memorial Scholarship
2009	Western Scholarship of Distinction

WORK EXPERIENCE

2013 – Present	Graduate Research Assistant, The University of Western Ontario
2013-2015	Graduate Teaching Assistant The University of Western Ontario
2011-2012	Undergraduate Volunteer Research Assistant The University of Western Ontario
2011	Volunteer Worker Kidney Foundation of Canada

CONFERENCES

Sepiol, C., and Dhaubhadel, S. (2015). Root-specific chalcone reductases hold the secret to *Phytophthora sojae* resistance in soybean. Canadian Plant Genomics Workshop, Victoria, British Columbia

Sepiol, C., Yu, J. and Dhaubhadel, S. (2014). Molecular and biochemical analysis of chalcone reductase genes in soybean. Canadian Society of Plant Biologists, Eastern Regional Meeting, Guelph, Ontario

Sepiol, C., and Dhaubhadel, S. (2014). Identification and characterization of chalcone reductase genes in soybean, Biology Graduates Research Forum (BGRF 2014), London, Ontario.