Unravelling the genetic mechanisms for target spot resistance in soybean

by

Sejal Jinesh Patel

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 10, 2022

Keywords: Target spot, Soybean, *Corynespora cassiicola*, Resistance, RNA-Seq, Differential gene expression, GWAS

Approved by

Jenny Koebernick, Chair, Associate Professor of Crop, Soil and Environmental Science
Charles Chen, Professor of Crop, Soil and Environmental Science
Kira L. Bowen, Professor of Entomology and Plant Pathology
Alex Harkess, Faculty Investigator, Hudson Alpha Institute for Biotechnology, Huntsville

Abstract

Soybean production has grown over the years and has become one of the main crops in the world. Biotic stress is an obstacle to soybean production, and developing resistant varieties is one of the major focuses of soybean breeders. Target spot of soybean is caused by foliar pathogen *Corynespora cassiicola*, a devastating disease in tropical and subtropical areas. The development of highly humid conditions and warm temperatures in the Mid-south and southeastern United States facilitated the spread of disease and reduced the potential crop yield. The overall objective of this dissertation was to identify germplasm resistant to the target spot and understand the underlying genetic mechanism of resistance. The first chapter reviewed the literature on *C. cassiicola* infection in soybean, pathogen characteristics, target spot management, soybean genomics, and the use of available resources and techniques. Chapter two focused on using a consistent and reproducible procedure for screening soybean varieties against multiple *C. cassiicola* isolates to identify horizontal resistance to the target spot. Herein, we also investigated the genetic diversity of *C. cassiicola* isolates by performing phylogenetic analysis based on four loci and cassiicolin-encoding genes. The third chapter indicated the first molecular mechanisms of soybean resistance to *C. cassiicola* infection by conducting comparative RNA sequencing (RNA-Seq) study. This study revealed several genes encoding leucine-rich repeat (LRR) domain, dirigent proteins (DIRs), and Cysteine (C)-rich receptor-like kinases (CRKs), flavonoids, jasmonic, salicylic, and brassinosteroids acid were upregulated in resistant genotypes after *C. cassiicola* infection. The fourth chapter revealed the first Genome-wide Association Study (GWAS) for target spot resistance in soybean. The GWAS and RNA-Seq study helped narrow down candidate target spot defense-associated genes. Additionally, we identified genomic regions that might be co-localized for resistance to several biotic and abiotic stress.

These findings provided germplasm and insight into the complex genetic architecture of target spot resistance in soybean and would promote marker-assisted selection in soybean breeding.

Acknowledgments

I want to express my gratitude and appreciation to my major advisor, Dr. Jenny Koebernick, for providing excellent opportunity, meticulous guidance, encouragement, advice, and support throughout my Ph.D. research at Auburn University. I am thankful to my committee members Dr. Kira Bowen, Dr. Charles Chen, and Dr. Alex Harkess, for their time, help, and suggestions. I appreciate Dr. Sushan Ru for contributing her time as an outside reader for my dissertation. I want to thank my fellow graduate students, research techs, and student workers, who have always helped in the greenhouse and laboratory task. I warmly thank the friends I met in Auburn for being part of this journey.

Last but most importantly, I would like to thank my husband, Jinesh Patel, for his constant motivation, unwavering support, and companionship. I am equally grateful to my beloved family Jivrajbhai, Dahyabhai, Ranjanben, Manjulaben, Nicky, Nirav, Jaydeep, Jyoti, and Krishav for their unconditional love, support, sacrifices, and prayers during this journey. I want to dedicate this dissertation to my family.

I am grateful to God and appreciate all of you for being part of my journey!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

°C	Celsius
DNA	deoxyribonucleic acid
µl	microliter
PCR	polymerase chain reaction
W	Watt
TxU	University of Texas at Austin
UU	University of Utah
WSU	Washington State University
DEGs	Differentially expressed genes
PTI	pathogen triggered immunity
ETI	Effectors-triggered immunity
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genomes
JA	Jasmonic acid
SA	Salicylic acid
BR	Brassinosteroid
LRR	Leucine-rich repeat
NBS-LRR	Nucleotide binding site-leucine rich repeat
PR protein	Pathogenesis-related protein
qRT-PCR	Real-time quantitative PCR
ROS	Reactive oxygen species
SAR	Systemic acquired resistance

TF	Transcription factor
FLS2	Flagellin sensing 2
bHLH	Basic helixloop-helix
ERF	Ethylene-responsive factor
MYB	Myeloblastosis viral oncogene homolog

Chapter 1. Literature review

1.1 History and wild relatives of soybean

Soybean (*Glycine max* (L.) Merr.), a native legume seed crop of China, is internationally important for food and feed due to its high protein and oil content. Around 6000 to 9000 years ago in East Asia, humans started reshaping the wild progenitor of soybean (*Glycine soja* Sieb. & Zucc) through subconscious artificial selection to derive the current domesticated soybean species (*Glycine max* L.) (FUKUDA 1933). Multiple genomic studies have supported the hypothesis that domesticated soybean has a single origin and is derived from one cluster of *G. soja* (ZHOU *et al.* 2015a; HAN *et al.* 2016). Cultivated and wild soybean have contrasting morphological characteristics. Wild soybean accessions (*G. soja*) have weed-like characteristics: vining stems, small black seeds in pods, and low yields with low oil content. Cultivated accessions (*G. max*) have erect stems with less branching, are high yielding with white to yellow seed coats, and higher oil content in seeds. Genome compositions of G. *soja* and *G. max* are similar and are the two species are sexually compatible (SINGH AND HYMOWITZ 1999). The F1 hybrid between *G. soja* and *G. max* show characteristics of the semi-wild soybean, *Glycine gracilis,* an intermediate species during the evolutionary process of domestication (HAN *et al.* 2016). The tertiary or quaternary gene pool for soybean consists of 16 wild perennial *Glycine* species native to Australia which are more distantly related and highly incompatible with *G. max* (CHUNG AND SINGH 2008; STUPAR AND SPECHT 2013).

1.2 Soybean production, importance and uses

Globally, soybean is planted on 128 million ha with a total production of 364.06 million metric tons (FAOSTAT 2020). The U.S., Brazil, and Argentina produce ~82% of the world's soybean crop with the U.S. accounting for more than 33% (FAOSTAT 2020). In 2020, soybean

was planted on 33.6 million hectares and yielded on average 112.54 bu ha^{-1} (USDA 2020). Thirty-one states grow soybean with Illinois, Iowa and Minnesota being the top three, in terms of area planted and total production. Alabama planted around 113 thousand ha of soybean with average production of 101.31 bu ha^{-1} (USDA 2020-2021).

1.3 Soybean seed composition

Soybean seed are composed of protein and oil making them important for human and livestock consumption. The main products of soybean are soymeal and oil. Soy meal is used to feed poultry and livestock. The oil is used for cooking, biodiesel, and other industrial uses like paints, plastic, and cleaners. Humans consume soybean in the form of soymilk, soy flour, tofu etc., and it is frequently used as a meat substitute which helps to fulfil daily protein requirements. Additionally, soybean also has bioactive components that improve human health by reducing cholesterol, lowering the risks of heart diseases, playing a preventive role in chronic renal disease and contributing to the reduction of risk for many types of cancer such as breast, liver, bladder, and stomach (MESSINA AND BARNES 1991; BARNES 1995; FRANKE *et al.* 1995; ZHOU *et al.* 1999; ISANGA AND ZHANG 2008; REBHOLZ *et al.* 2013).

The main components of soybean seeds are protein (36%), insoluble carbohydrate (19%), soluble carbohydrate (9%), oil (19%), water (9%), and ash (5%) (LIU 1997; REDONDO-CUENCA *et al.* 2007). The seed also contains phytic acid, vitamins B1, B2, B3, B5, and B6, K, dietary minerals, and bioactive compounds like isoflavones (CARRAO-PANIZZIL AND KITAMURA 1995; LIU 1997; CHARRON *et al.* 2005). The seeds contain a large portion of energy in the form of protein. Soy protein is actually a combination of different proteins classified into four categories according to their sedimentation during centrifugation and are classified as 2S (conglycinin), 7S (β-conglycinin), 11S (glycinin), and 15S (globulins) (SAIO *et al.* 1969; NISHINARI *et al.* 2014;

LAMAMING *et al.* 2021). The 2S and 15S components consist of 20% total protein, while 7S and 11S consist of 80% protein. The presence of glycinin (11S), β-conglycinin (7S) and their ratio have been associated to tofu yield, quality, and tofu gel firmness (TAIRA 1990; MURPHY 1997; MUJOO *et al.* 2003). Generally, soybean oil consists of different fatty acids: 11% palmitic (C16:0), 4% stearic (C18:0), 26% oleic (C18:1), 52% linoleic (C18:2), and 7 % linolenic (C18:3) acids (GOYARY *et al.* 2015). Palmitic and stearic are unsaturated fatty acids, while oleic, linoleic, and linolenic are saturated fatty acids. High levels of linoleic and linolenic acids in soybean oil reduce oxidative stability and food storage time (WARNER AND FEHR 2008). Additionally, they also reduce the frying stability of soybean oil and overall flavor. On the other hand, high oleic acid content reduces the trans-fat, which reduces risks of health problems and improves oxidative stability and shelf life. Recent breeding efforts have successfully improved the composition of soybean oil by increasing oleic acid percentage and reducing linoleic and linolenic percentages.

1.4 Growth and development of soybean plant

Depending on the maturity group, a soybean plant takes 4 to 5 months from sowing to full maturity. Adequate soil content provides oxygen, temperature, and water required for successful germination. The seed swells as soon as it absorbs water by hydrating the seed coat, embryo axis, and cotyledon; this initiates the process of germination and emergence by enzyme activity and respiration (HILLEL AND KOZLOWSKI 2012). Dynamic changes in the chemical composition include reduction in oil content, iodine content, sugar, and unsaturated fatty acids while increases in total saturated fatty acid concentration, non-protein nitrogen, and amino acid content occurs (MOSTAFA *et al.* 1987). The emergence process occurs three to four days after planting; the seed coat is removed completely, and the cotyledon moves longitudinally, pulling itself out of the soil (WRIGHT AND LENSSEN 2013). Soybean development is divided into vegetative and reproductive

stages. The vegetative stages consist of emergence (VE), cotyledon stage (VC), and first node (V1) to nth node (Vn). The reproductive stages are beginning bloom, full bloom, beginning pod, full pod, beginning seed, full seed, beginning maturity, and full maturity, respectively R1 to R8 (FEHR AND CAVINESS 1977).

1.5 Target spot impact on soybean

Target spot is a disease caused by the fungus, *Corynespora cassiicola*. In 1868, the fungus was first identified as *Helminthosporium cassiicola* by C. T. WEI (1950), that now label as *Corynespora cassiicola* (Berk. & M. A. Curtis). The initial occurrence of target spot on soybean was reported in the U.S. by OLIVE *et al.* (1945). A research group from Mississippi observed 18-32% yield losses associated with soybean target spot (HARTWIG 1959). Later, cases of soybean infected by *C. cassiicola* were reported in several countries, namely Brazil, Argentina, Canada, Japan, and Sri Lanka (SEAMAN *et al.* 1965; ONESIROSAN *et al.* 1974; PLOPER AND RAMALLO 1988; SILVA *et al.* 2000; ALMEIDA *et al.* 2001; SHIMOMOTO *et al.* 2011). Conditions of high humidity in the southeastern states have facilitated the spread of the pathogen which has become a great concern for farmers (KOENNING *et al.* 2006; FASKE 2016). The infection of leaves with *C. cassiicola* occurs when relative humidity is greater than 80% during warm temperatures (20 to 30 ◦C). Prolonged humidity and free moisture on the leaf surface are the most important factors for disease development (GODOY 2015). The phytopathogenic fungus infects leaves, stems, petioles, and pods. On leaves, the symptoms are brown-red, somewhat circular (10-15 mm), lesions with an alternating concentric ring pattern, frequently surrounded by a yellow halo (SEAMAN *et al.* 1965; GODOY 2015). Soybean plant stems and roots can be infected during the seedling stage, with roots showing symptoms of reddish–brown lesions about three days after emergence when the soil temperature is 15 to 18°C; lesions will turn dark violet-brown when *C.*

cassiicola sporulates (GODOY 2015; HARTMAN *et al.* 2015). The petioles and stems exhibit dark brown lesions with different shapes. Premature defoliation occurs with a severe infection under optimum conditions. The spread of the symptoms on leaves starts from the lower canopy and moves upward through the canopy later in the season.

1.6 Taxonomy of *Corynespora cassiicola*

Kingdom: Fungi

Phylum: Ascomycota

Family: *Corynesporascaceae*

Genus: *Corynespora*

Species: *cassiicola*

1.7 Life cycle and colony characteristics of *C. cassiicola*

The fungus *Corynespora cassiicola* infects more than 500 plant species from 380 genera (SHRESTHA *et al.* 2017) and is known as necrotrophic fungus which destroys infected host tissues to utilize nutrients (ALMEIDA *et al.* 2001). In general, *C. cassiicola* survives winter months associated with plant debris and it is likely that conidia (i.e., primary inoculum) are produced when temperature and moisture conditions are suitable. The survival of *C. cassiicola* on debris of plants other than soybeans provides a higher level of adaptability and allows survival for a more extended period. It has been reported that chlamydospore formation by *C. cassiicola* survived for more than two years on host debris or on soil in unfavorable conditions. Target spot is a polycyclic disease, where *C. cassiicola* finishes several generations in a single growing season (MACKENZIE *et al.* 2018).

Corynespora cassiicola has colonies that are grey to whitish with sparse hair-like growth. The conidiophores are brown, branched, erect, and swollen at the basal cell, having 1-20 septate long, which measure 44-350 × 4-11 µm. The conidia germinate on mostly single conidiophores or chains

of two to six. They are straight, cylindrical, broad at the base with slight tapering towards the tip, smooth and brownish. The size of conidia ranges from 39-520 × 7-22 µm (HARTMAN *et al.* 2015).

1.8 Diversity and host specialization of *C. cassiicola*

The genetic diversity of *C. cassiicola* isolates has been studied to better understand it's evolution evaluation and variability among different isolates, which helps to develop management strategies (DIXON *et al.* 2009; DÉON *et al.* 2014; SHRESTHA *et al.* 2017; SUMABAT *et al.* 2018). Although *C. cassiicola* has been documented to have a broad host range, specific isolates may infect one or few hosts. A few studies have described the phylogenetic diversity and host specialization of *C. cassiicola. DIXON et al. (2009)* genotyped 143 *C. cassiicola* isolates with four loci (*act*, *caa5*, *ga4*, and ITS) which separated them into six distinct phylogenetic lineages (PLs). The analysis determined that the pathogenicity of *C. cassiicola* isolates is limited to host plant. For example, papaya isolates were pathogenic only to papaya, revealing host specificity. Moreover, SUMABAT *et al.* (2018) screened and genotyped 53 *C. cassiicola* isolates from eight crops with the four loci described by (DIXON *et al.* 2009). The pathogenicity experiment determined that soybean and cotton isolates clustered in one PL group, but soybean isolates were more diverse than those from cotton and might not infect cotton as severely, indicating host specialization.

1.9 Target spot management and use of fungicides

Management strategies to control target spot disease include fungicide application, rotation to non-host crops (rice, corn), and managing crop residues by tillage. Moreover, avoiding narrow row spacing, which allows good airflow circulation during canopy closure, reduces favorable conditions for developing diseases. However, growers continue to rely primarily on a weekly spray of a fungicide when highly favorable weather conditions are present; this is not an

economical nor eco-friendly control strategy and presents the possibility of *C. cassiicola* developing fungicide resistance. *Corynespora cassiicola* is listed as being at high risk for developing fungicide resistance (FRAC 2019). It is recommended that the use of combinations of fungicides in different FRAC groups might reduce incidents and prevent the development of fungicide resistance in *C. cassiicola.*

In soybean, the fungicide groups consisting of methyl benzimidazole carbamate (MBCs), succinate dehydrogenase inhibitors (SDHIs), demethylation inhibitors (DMIs) and quinone outside inhibitors (QoIs) are mostly used to control foliar fungal diseases such as brown spot, Cercospora leaf blight (CLB), frogeye leaf spot, and soybean rust (SCHERM *et al.* 2009; SWOBODA AND PEDERSEN 2009). Multiple studies have been conducted to evaluate the efficacy of different fungicides and their combinations for controlling *C. cassiicola* (EDWARDS MOLINA *et al.* 2019; REZNIKOV *et al.* 2019; NEVES AND BRADLEY 2020). A meta-analysis of data from 56 field trials over five growing seasons in Brazil using four different combinations of six fungicides identified the combinations of fluxapyroxad + pyraclostrobin and epoxiconazole + fluxapyroxad + pyraclostrobin as having high efficacy for controlling target spot disease (EDWARDS MOLINA *et al.* 2019). However, studies have indicated that fungicide application might not be an efficient way to control target spot due to mutation conferring target-site resistance and declining effects of fungicide after application (DUAN *et al.* 2019; RONDON AND LAWRENCE 2019). With the use of the colorimetric microtiter method, it was found that several *C. cassiicola* isolates collected from the Paraná (PR) and Mato Grosso (MT) states in Brazil had developed resistance against fungicide of group MBC and QoI (XAVIER *et al.* 2013; XAVIER *et al.* 2021). In 2016, two separate trials conducted in Starkville, MS, to check the efficacy of nine combinations of

fungicides against different fungal pathogens found that none of the combinations effectively reduced target spot severity over time. This suggests that multiple fungicide applications will be required to control the disease, which is not economical (ALLEN AND IRBY 2017). A better alternative to target spot management will be to develop resistant genotypes in soybean which can be used with Integrated Pest Management strategies to reduce the effect of target spot. This would also help reduce the environmental and economic impact of fungicide applications.

1.10 Soybean genomics

Cultivated soybean (*G. max*) is a diploid species which experienced at least two rounds of genome duplications (SHOEMAKER *et al.* 1996; GRANT *et al.* 2000). It consists of 20 chromosomes with a genome size of about 1.12 Gb (ARUMUGANATHAN AND EARLE 1991). In 2010, the first *Glycine max* genome, 'Williams 82', was sequenced through whole-genome shotgun sequencing covering 950 megabases (Mb) assembled in 397 scaffolds arranged on 20 linkage groups of soybean genome (SCHMUTZ *et al.* 2010). Two other soybean accessions, Zhonghuang 13 (ZH13) and wild soybean (W05), have been sequenced and are available as reference genomes (SHEN *et al.* 2019; XIE *et al.* 2019). A genome resequencing using 31 soybean accession sheds light on the genetic diversity between 17 wild and 14 cultivated species. The study suggested that cultivated species were less diverse when compared to wild species and emphasized that the high linkage disequilibrium nature of the soybean genome would facilitate marker-assisted breeding over gene cloning (LAM *et al.* 2010). Another resequencing study consists of 302 soybean accession (62 wild soybeans, 130 landraces, and 110 improved cultivars) in which more than 9 million SNPs (single-nucleotide polymorphisms) and 876799 indels were identified, providing an excellent source for developing a molecular marker platform (ZHOU *et al.* 2015b).

Resequencing studies require mapping the short read sequences to the reference genome. Short reads generated from accessions highly diverged from the reference genome will only capture a limited amount of variation, as many of the reads will fail to map the reference genome. Additionally, variations like copy number variation and present-absent variation are generally missed in resequencing studies (LI *et al.* 2010; LIU *et al.* 2020). Such missed variation might be impacting crucial agronomic traits. The high degree of genetic variation between different accessions of the same plant species suggests the need for multiple reference genomes (BAYER *et al.* 2020).

A pan-genome consists of all possible gene sets within a plant species. It can be divided into the core genome, which is shared among all the investigating accessions of the study, and the dispensable genome which is present in the one or few accessions of the study. The first pan-genome was developed in bacteria, while the first plant pan-genome study was conducted using seven accessions of wild soybean species (*Glycine soja*) (Tettelin et al. 2005; Li et al. 2014). A recent pan-genome study was conducted in soybean using 26 representatives out of 2898 deeply sequenced soybean accessions (LIU *et al.* 2020). Each of these accessions was de novo assembled, which was used to identify the structural variation of the 2,898 soybean accessions. All genes in 27 genomes (26 representatives + ZH13 reference genome) were distributed into 57,492 families out of which 28,746 families were considered a core or soft-core gene set (found in >90% of individuals), 28,679 families were considered dispensable genes, and 27 families were only identified in single individuals of the collection. The dispensable and private genomes were enriched by biotic and abiotic stress genes indicating that dispensable genome will play a major role in determining the response to pathogen attack.

1.11 Gene families involved in plant pathogen response

Several gene families have been identified to be involved in mechanisms related to plant response to pathogen attacks. For instance, *GmDIR22* encodes dirigent proteins that play an important role in structural fortification by participating in the biosynthetic process of lignin-like molecules, thus providing resistance to biotic and abiotic stress. Transgenic soybean containing *GmDIR22* had higher lignin production and reduced *Phytophthora sojae* hyphal growth, thus providing resistance against the stem and root rot (LI *et al.* 2017). In wheat, expression of a dirigent protein encoding gene, *HfrDrd*, increased by 20-fold during Hessian fly attack (SUBRAMANYAM *et al.* 2013). Another study found that the *ScDir* gene was associated with production of a 187 amino acid-long dirigent protein that was highly expressed when sugarcane seedlings were exposed to abiotic stress like salt, drought, and H_2O_2 stress (JIN-LONG *et al.* 2012).

Chitin, an amide derivative of glucose, is a long-chain polymer of N-acetylglucosamine, found as a cell wall component of pathogens like fungi (PUSZTAHELYI 2018). The chitinase family participates in the degradation of chitin and plays a role in restricting the growth of fungal pathogens such as *Trichosanthes dioica*, *Aspergillus niger Alternaria solani*, *Fusarium sp*., *Rhizoctonia solani*, and *Verticillium dahlia* (SCHLUMBAUM *et al.* 1986; KABIR *et al.* 2016; KUMAR *et al.* 2018; TOUFIQ *et al.* 2018). A study found that overexpression of CHI with β-1,3-glucanase (GLU) or ribosome-inactivating protein (RIP) in transgenic tobacco plants increases resistance against the soil-borne fungal pathogen, *Rhizoctonia solani* (JACH *et al.* 1995). Protein extracted from a *Fusarium solani*-infected bean plant, which was rich in CHI and GLU, inhibited the growth of 15 pathogenic fungi *in vitro* by hydrolysis of the fungal cell wall (MAUCH *et al.* 1988).

Phenylalanine ammonia laLyase (PAL I & PALII) encode enzymes for the first step of the general phenylpropanoid pathway that catalyzes the reaction of converting L-phenylalanine to ammonia and trans-cinnamic acid (CAMM AND TOWERS 1973). Phenylalanine ammonia-lyase (PAL) and tyrosine ammonia-lyase (TAL) expression increases in the presence of chitin (a fungal cell wall component) which increases phenolic content in soybean leaves (Khan et al. 2003). Arabidopsis mutant plants that had all copies of PAL gene knockout showed a significant reduction in salicylic acid and higher susceptibility to the bacterial pathogen *Pseudomonas syringae* (HUANG *et al.* 2010). In rice, a phenylalanine ammonia-lyase gene (*OsPAL4*) provides resistance to broad range of diseases such as bacterial blight, sheath blight, and rice blast (TONNESSEN *et al.* 2015).

Transcription factors such as *AP2*, *ERF*s, *WRKY*, *MYB*, and *bHLH* play an important role in plant-pathogen interaction. Expression of WRKY TFS, e.g., WRKY29 and WRKY41, were found at elevated levels which conferred Fusarium head blight resistance in wheat (SAROWAR *et al.* 2019) and *Pseudomonas* resistance in Arabidopsis (HIGASHI *et al.* 2008), respectively. *CsERF004* is a gene in cucumber that belongs to subfamily of AP2/ERF-like transcription factors (NAKANO *et al.* 2006). Such transcription factors help in the transcription of genes in the presence of ethylene; they are also known to be expressed in response to pathogen attack and abiotic stress (ONATE-SANCHEZ AND SINGH 2002). Gene expression and functional analysis of *CsERF004* during infection by *C. cassiicola* of cucumber suggest that it confers resistance against target spot (LIU *et al.* 2017).

One of the most common R-genes contains a nucleotide-binding site (NBS) and a leucine-rich repeat (LRR), also known as NBS-LRR family. These families might have toll and interleukin 1 receptors (TIR) domain or coiled-coil (CC) (McHale et al. 2006). The nucleotide-binding (NB)

domains are found in human Apaf1 (Apoptotic protease-activating factor-1), plant R-genes and *C*ed4 (Caenorhabditis elegans death-4) of *Caenorhabditis elegans*, so they are also known as NB-ARC protein (VAUX 1997; VAN DER BIEZEN AND JONES 1998). A *Csa6m375730* is a NB-ARC type resistance gene that was identified on a fine map region of a target spot resistance gene, cca-3in cucumber (WEN *et al.* 2015). NB-ARC contains three subunits-- NB (nucleotide-binding), ARC1, and ARC2-- and activation of this protein results in cell death in plants and animals (VAN DER BIEZEN AND JONES 1998; VAN OOIJEN *et al.* 2008). Such proteins are involved in pathogen recognition and immune response.

An RNA-Seq study will not only aid in understanding if these genes are involved in resistance pathways in soybean against target spot disease but also provide a comprehensive picture of changes in gene networks and biological pathways.

1.12 RNA-Sequencing

Many techniques have been developed to quantify RNA transcripts, including ETS sequencing, microarrays, and RNA-sequencing. Microarray uses a hybridization-based approach which is constrained by needing prior knowledge of a gene sequence, availability of probes for known genes, low abundance transcripts, and an ability to distinguish different isoforms and allelic expressions (JAKSIK *et al.* 2015; RAI *et al.* 2018; RAO *et al.* 2018). While an ETS-based approach is sequenced-based, which directly determines cDNA sequencing or ETS libraries, it is limited by low throughput, expense, the need for a high amount of RNA, and it is generally not quantitative (BOGUSKI *et al.* 1994; GERHARD *et al.* 2004)

Due to the rapid revolution of next generation sequencing techniques, it has become cheaper and faster to conduct genome and transcriptome sequencing projects (PARK AND KIM 2016; LEVY AND BOONE 2019). RNA-Seq technology uses next-generation sequencing, directly

identifying and quantifying RNA transcripts for the whole genome. It has numerous advantages for examining transcriptome structure, such as detecting novel transcripts, splice junctions, and allele-specific expression, and gene expression levels over time or in different treatments. In addition, it could also identify other types of transcripts, including mRNAs and non-coding RNAs (miRNA, small RNA) (WANG *et al.* 2009). Thus, RNA-Seq is a widely applied technique for identifying differently expressed genes (DEGs) and identifying alternative splicing (AS).

Several RNA-Seq studies in response to plant pathogen attacks have been reported in soybean. MCCABE *et al*. (2018) investigated a cluster of receptors-like proteins that were involved as candidates for the RbS3 resistance gene by conducting a RNA-seq study in response to *Phialophora gregata* infection. DONG *et al.* (2018) investigated the differential expression of sixteen WRKY transcription factors (TF) in response to *Peronospora manshurica* infection to soybean downy mildew resistant and susceptible lines; their results illustrate that WRKY31 TF might control SAGT1 gene expression by activating a salicylic acid-associated immune response. KIM *et al.* (2011) conducted an RNA-seq study to investigate plant molecular mechanisms against *Xanthomonas axonopodis* inoculation. The results determined that the association of PAMP (pathogen-associated molecular patterns) and accumulation of defense-related genes contributes to bacterial leaf pustule resistance.

1.13 Genetic map and Quantitative Trait Loci (QTL) studies in soybean

Developing genetic maps and identifying Quantitative trait loci (QTL) are standard practices to analyze genomic regions for polygene traits and to characterize genetic interactions. Several genetic maps have been constructed using different mapping populations and molecular markers like RFLP (restriction fragment length polymorphisms), SSR (Simple sequence repeats), EST (expressed sequence tag), and SNP (single-nucleotide polymorphism) markers. An early study in

soybean performed with RFLP developed 26 linkage groups and mapped several reproductive and morphological traits (KEIM *et al.* 1990). These molecular makers have also been used to identify numerous QTLs in soybean for traits like seed quality, seed composition, seed yield, flowering time, pod dehiscence, and reactions to biotic and abiotic stresses. QTLs have been identified for a total of 124 phenotypic traits and deposited in the soybase (GRANT *et al.* 2010).

The availability of genome sequences, high-density molecular markers, and functional verification tools have helped fine map the QTLs, identify candidate genes, and validate their function. These resources have accelerated identifying genomic regions and candidate genes that confer resistance to different pathogens. High-density SNP markers and RIL populations segregating for root-knot nematode resistance helped to identify two candidate genes (Glyma10g02150 and Glyma10g02160) in the vicinity of a major QTL for root-knot nematode resistance (XU *et al.* 2013). A significant QTL for soybean cyst nematode on GM08, *Rhg4*, was fine mapped and candidate gene mining identified a gene encoding a serine hydroxymethyltransferase (SHMT) to provide resistance (LIU *et al.* 2012). White mold caused by the fungus *Sclerotinia sclerotiorum* is a major concern for soybean production in the U.S. Using a RIL population and SNP makers, the QTL region providing resistance to white mold was narrowed down to a few candidate genes (ZHAO *et al.* 2015). The function of such genes can be validated through different transformation techniques, and markers in such regions can be used for early resistance screening to white mold. A RIL population was developed by crossing 'Taekwangkong' and 'Danbaekkong', which are susceptible and resistant, respectively, to *X. axonopodis* pv. *glycines*, the cause of bacterial leaf pustule. Screening this population with molecular markers helped to fine map the QTL, linking the resistance to the 33 kb region, which has two putative genes (KIM *et al.* 2010).

A cutting-edge technique known as QTL-seq is in demand, which rapidly maps quantitative trait loci through genome resequencing and Bulked-segregant analysis (BSA). The first QTL-seq study was reported in rice with two bulked populations, each consisting of about 50 individuals and representing the highest or lowest bulk of a given phenotypic characteristic. This study successfully identified QTL regions associated with seedling vigor and rice blast disease resistance, thus demonstrating the power of BSA and next generation sequencing (NGS) over traditional QTL mapping, which is tedious and time consuming (TAKAGI *et al.* 2013). QTL-seq studies have been performed in soybean to identify QTL regions for agronomically important traits such as plant height (ZHANG *et al.* 2018), pod length and width (XIE *et al.* 2021), protein content (WANG *et al.* 2021), *Phytophthora sojae* resistance (ZHONG *et al.* 2018), and charcoal rot resistance (DA SILVA *et al.* 2020). In conclusion, these studies demonstrate that the availability of a reference genome sequence, rapid development in sequencing technology, and large sets of molecular markers have drastically improved soybean research in a cost-effective manner.

1.14 Genome wide association mapping study (GWAS)

Complex traits are influenced by many large and small effect genomic regions (QTLs), and their interaction with each other and the environment. The bi-parental population has been used extensively to develop linkage groups and conduct QTL studies. However, there are several limitations to such populations. These populations have few possible recombination events, which are time consuming and tedious. Genome-wide association mapping is an alternative technique to linkage mapping of bi-parental populations, which leverages natural variation among the diverse population and provides a high mapping resolution for trait variation (ZHU *et al.* 2008). This approach has benefited from the development of significantly increased numbers of markers due to the origin of next-generation sequencing techniques and increased sample sizes by building

biological repositories. GWAS' are now routinely conducted in crop species, including soybean, wheat, barley, maize, and rice (WANG *et al.* 2012; FANG *et al.* 2017; JABBARI *et al.* 2018; KIM AND REINKE 2019; SINGH *et al.* 2020; YUAN *et al.* 2020). The availability of increased marker data and the development of new statistical methods provide critical knowledge of QTLs and candidate genes associated with the trait of interest. With the accessibility of marker data (SoySNP50K iSelect Bead Chip) in the soybase (USDA soybean germplasm collection) (SONG *et al.* 2013), several studies have mapped significant major effects of quantitative trait loci (QTL) using GWAS analysis in response to biotic stress, such as infection with *Sclerotinia sclerotiorum* (WEN *et al.* 2018), *Heterodera glycines* (VUONG *et al.* 2015; RAVELOMBOLA *et al.* 2020), *Macrophomina phaseolina* (COSER *et al.* 2017), *Pythium sylvaticum* (LIN *et al.* 2020), and *Aphis glycines* (CHANG *et al.* 2016a). A comprehensive GWAS study on 13 soybean pathogens including nematodes, bacteria, fungi, and viruses, identified 74 genomic regions that were associated with resistance. The study highlighted enrichment of leucine-rich repeat (LRR) receptor-like kinases and nucleotide-binding site-LRR candidate resistance genes in these regions (CHANG *et al.* 2016b). A meta-analysis involving around 17,500 accessions used in 73 different GWAS studies identify 571 SNPs associated with soybean diseases. Additionally, they identified 33 candidate genes in the region, and future experiments will reveal their role in conferring disease resistance. Such GWAS studies not only pivot the road to developing genome prediction models but also to develop models to understand how different QTLs and genes interacts (SHOOK *et al.* 2021).

In conclusion, soybean is one of the crop species for which there are abundant genomic resources available. Several studies show the importance of using techniques like RNA-Seq and GWAS to identify genomic regions and candidate genes that bestow resistance to biotic and abiotic stresses. Target spot is an important disease in tropical and subtropical regions including the southern U.S.

growing region; however, no efforts have been made to identify genomic regions that contribute to resistance of this disease. Thus, here our research objectives are to:

1) Develop a reproducible methodology to screen soybean genotypes against *C. cassiicola*;

2) Determine genomic regions that are involved in resistance against target spot; and

3) Identify candidate genes and gene networks activated in response to *C. cassiicola* attack in resistant and susceptible genotypes.

References

Allen, T., and T. Irby, 2017 Managing target spot of soybean: 2016 Fungicide Trial Efficacy Results, pp.

Almeida, Á. M. R., O. F. Saraiva, J. R. B. Farias, C. A. Gaudêncio and E. Torres, 2001 Survival of pathogens on soybean debris under no-tillage and conventional tillage systems. Pesquisa Agropecuária Brasileira 36**:** 1231-1238.

Arumuganathan, K., and E. D. Earle, 1991 Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9**:** 208-218.

Barnes, S., 1995 Effect of genistein on in vitro and in vivo models of cancer. The Journal of Nutrition 125**:** 777S-783S.

Bayer, P. E., A. A. Golicz, A. Scheben, J. Batley and D. Edwards, 2020 Plant pan-genomes are the new reference. Nature Plants 6**:** 914-920.

Boguski, M. S., C. M. Tolstoshev and D. E. Bassett Jr, 1994 Gene discovery in dbEST. Science 265**:** 1993-1994.

Camm, E. L., and G. H. N. Towers, 1973 Phenylalanine Ammonia-Lyase. Phytochemistry 12**:** 961-973.

Carrao-Panizzil, M. C., and K. Kitamura, 1995 Isoflavone content in Brazilian soybean cultivars. Japanese Journal of Breeding 45**:** 295-300.

Chang, H.-X., P. J. Brown, A. E. Lipka, L. L. Domier and G. L. Hartman, 2016a Genome-wide association and genomic prediction identifies associated loci and predicts the sensitivity of Tobacco ringspot virus in soybean plant introductions. BMC Genomics 17**:** 1-11.

Chang, H.-X., A. E. Lipka, L. L. Domier and G. L. Hartman, 2016b Characterization of Disease Resistance Loci in the USDA Soybean Germplasm Collection Using Genome-Wide Association Studies. Phytopathology® 106**:** 1139-1151.

Charron, C. S., F. L. Allen, R. D. Johnson, V. R. Pantalone and C. E. Sams, 2005 Correlations of oil and protein with isoflavone concentration in soybean [*Glycine max* (L.) Merr.]. Journal of Agricultural and Food Chemistry 53**:** 7128-7135.

Chung, G., and R. J. Singh, 2008 Broadening the genetic base of soybean: a multidisciplinary approach. Critical Reviews in Plant Sciences 27**:** 295-341.

Coser, S. M., R. V. Chowda Reddy, J. Zhang, D. S. Mueller, A. Mengistu *et al.*, 2017 Genetic architecture of charcoal rot (*macrophomina phaseolina*) resistance in soybean revealed using a diverse panel. Frontiers in Plant Science 8.

Da Silva, M. P., A. Z. Zaccaron, B. H. Bluhm, J. C. Rupe, L. Wood *et al.*, 2020 Bulked segregant analysis using next-generation sequencing for identification of genetic loci for charcoal rot resistance in soybean. Physiological and Molecular Plant Pathology 109**:** 101440.

Déon, M., B. Fumanal, S. Gimenez, D. Bieysse, R. R. Oliveira *et al.*, 2014 Diversity of the cassiicolin gene in *Corynespora cassiicola* and relation with the pathogenicity in *Hevea brasiliensis*. Fungal Biology 118**:** 32-47.

Dixon, L. J., R. L. Schlub, K. Pernezny and L. Datnoff, 2009 Host specialization and phylogenetic diversity of *Corynespora cassiicola*. Phytopathology 99**:** 1015-1027.

Dong, H., S. Shi, C. Zhang, S. Zhu, M. Li *et al.*, 2018 Transcriptomic analysis of genes in soybean in response to *Peronospora manshurica* infection. BMC Genomics 19**:** 1-13.

Duan, Y., W. Xin, F. Lu, T. Li, M. Li *et al.*, 2019 Benzimidazole-and QoI-resistance in *Corynespora cassiicola* populations from greenhouse-cultivated cucumber: An emerging problem in China. Pesticide Biochemistry and Physiology 153**:** 95-105.

Edwards Molina, J. P., P. A. Paul, L. Amorim, L. H. C. P. da Silva, F. V. Siqueri *et al.*, 2019 Meta-analysis of fungicide efficacy on soybean target spot and cost–benefit assessment. Plant Pathology 68**:** 94-106.

Fang, C., Y. Ma, S. Wu, Z. Liu, Z. Wang *et al.*, 2017 Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. Genome Biology 18**:** 1-14.

Faostat, 2020 Faostat, Agriculture Organization of the United Nations, pp.

Faske, T., 2016 Arkansas Soybeans: Target Spot – What Do We Know?, pp.

Fehr, W. R., and C. E. Caviness, 1977 Stages of soybean development. Iowa Agricultural and Home Economics Experiment Station Special Report**:** 3-11.

FRAC, 2019 Pathogen Risk List, pp. in *Fungicide Resistance Action Committee*, edited by B. CropLife International, Belgium.

Franke, A. A., L. J. Custer, C. M. Cerna and K. Narala, 1995 Rapid HPLC analysis of dietary phytoestrogens from legumes and from human urine. Proceedings of the Society for Experimental Biology and Medicine 208**:** 18-26.

Fukuda, Y., 1933 Cytogenetical studies on the wild and cultivated Manchurian soybeans (Glycine L.). Jpn. J. Bot. 6**:** 489-506.

Gerhard, D. S., L. Wagner, E. A. Feingold, C. M. Shenmen, L. H. Grouse *et al.*, 2004 The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Research 14**:** 2121-2127.

Godoy, C. V., 2015 Target spot, pp. 62-63 in *Compendium of soybean diseases and pests*, edited by G. L. Hartman, J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis *et al.* Am Phytopath Society.

Goyary, J., A. Kumar and P. K. Nayak, 2015 Changes of quality parameters of soybean oil in deep fat frying: A review. J. Basic and App. Engg. Res 2**:** 951-954.

Grant, D., P. Cregan and R. C. Shoemaker, 2000 Genome organization in dicots: genome duplication in *Arabidopsis* and synteny between soybean and *Arabidopsis*. Proc Natl Acad Sci U S A 97**:** 4168-4173.

Grant, D., R. T. Nelson, S. B. Cannon and R. C. Shoemaker, 2010 SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res 38**:** D843-846.

Han, Y., X. Zhao, D. Liu, Y. Li, D. A. Lightfoot *et al.*, 2016 Domestication footprints anchor genomic regions of agronomic importance in soybeans. New Phytologist 209**:** 871-884.

Hartman, G. L., J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis *et al.*, 2015 *Compendium of soybean diseases and pests*. Am Phytopath Society.

Hartwig, E. E., 1959 Effect of target spot on yield of soybeans. Plant Dis. Rep 43**:** 504-505.

Higashi, K., Y. Ishiga, Y. Inagaki, K. Toyoda, T. Shiraishi *et al.*, 2008 Modulation of defense signal transduction by flagellin-induced WRKY41 transcription factor in *Arabidopsis* thaliana. Molecular Genetics and Genomics 279**:** 303-312.

Hillel, D., and T. T. Kozlowski, 2012 Soil moisture and seed germination. Water deficits and plant growth 3**:** 65-89.

Huang, J., M. Gu, Z. Lai, B. Fan, K. Shi *et al.*, 2010 Functional analysis of the arabidopsis pal gene family in plant growth, development, and response to environmental stress Plant Physiology 153**:** 1526-1538.

Isanga, J., and G.-N. Zhang, 2008 Soybean bioactive components and their implications to health—a review. Food Reviews International 24**:** 252-276.

Jabbari, M., B. A. Fakheri, R. Aghnoum, N. Mahdi Nezhad and R. Ataei, 2018 GWAS analysis in spring barley (*Hordeum vulgare L.)* for morphological traits exposed to drought. PloS one 13**:** e0204952.

Jach, G., B. Gornhardt, J. Mundy, J. Logemann, E. Pinsdorf *et al.*, 1995 Enhanced quantitative resistance against fungal disease by combinatorial expression of different barley antifungal proteins in transgenic tobacco. Plant J 8**:** 97-109.

Jaksik, R., M. Iwanaszko, J. Rzeszowska-Wolny and M. Kimmel, 2015 Microarray experiments and factors which affect their reliability. Biol Direct 10**:** 46.

Jin-Long, G., X. Li-Ping, F. Jing-Ping, S. Ya-Chun, F. Hua-Ying *et al.*, 2012 A novel dirigent protein gene with highly stem-specific expression from sugarcane, response to drought, salt and oxidative stresses. Plant Cell Rep 31**:** 1801-1812.

Kabir, S. R., M. M. Rahman, S. Tasnim, M. R. Karim, N. Khatun *et al.*, 2016 Purification and characterization of a novel chitinase from Trichosanthes dioica seed with antifungal activity. International Journal Of Biological Macromolecules 84**:** 62-68.

Keim, P., B. W. Diers, T. C. Olson and R. C. Shoemaker, 1990 RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics 126**:** 735-742.

Khan, W., B. Prithiviraj and D. L. Smith, 2003 Chitosan and chitin oligomers increase *phenylalanine ammonia-lyase* and tyrosine *ammonia-lyase* activities in soybean leaves. J Plant Physiol 160**:** 859-863.

Kim, D. H., K. H. Kim, K. Van, M. Y. Kim and S.-H. Lee, 2010 Fine mapping of a resistance gene to bacterial leaf pustule in soybean. Theoretical And Applied Genetics 120**:** 1443-1450.

Kim, K. H., Y. J. Kang, D. H. Kim, M. Y. Yoon, J.-K. Moon *et al.*, 2011 RNA-Seq analysis of a soybean near-isogenic line carrying bacterial leaf pustule-resistant and-susceptible alleles. DNA Research 18**:** 483-497.

Kim, S.-M., and R. F. Reinke, 2019 A novel resistance gene for bacterial blight in rice, *Xa43* (t) identified by GWAS, confirmed by QTL mapping using a bi-parental population. PloS one 14**:** e0211775.

Koenning, S., T. Creswell, E. Dunphy, E. Sikora and J. Mueller, 2006 Increased occurrence of target spot of soybean caused by *Corynespora cassiicola* in the Southeastern United States. Plant Disease 90**:** 974-974.

Kumar, M., A. Brar, M. Yadav, A. Chawade, V. Vivekanand *et al.*, 2018 Chitinases-potential candidates for enhanced plant resistance towards fungal pathogens. Agriculture 8**:** 88.

Lam, H.-M., X. Xu, X. Liu, W. Chen, G. Yang *et al.*, 2010 Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nature Genetics 42**:** 1053.

Lamaming, S. Z., J. Lamaming, N. F. M. Rawi, R. Hashim, M. H. M. Kassim *et al.*, 2021 Improvements and limitation of soy protein-based adhesive: A review. Polymer Engineering & Science 61**:** 2393-2405.

Levy, S. E., and B. E. Boone, 2019 Next-generation sequencing strategies. Cold Spring Harbor Perspectives in Medicine 9**:** a025791.

Li, N., M. Zhao, T. Liu, L. Dong, Q. Cheng *et al.*, 2017 A novel soybean dirigent gene gmdir22 contributes to promotion of lignan biosynthesis and enhances resistance to phytophthora sojae. Front Plant Sci 8**:** 1185.

Li, R., Y. Li, H. Zheng, R. Luo, H. Zhu *et al.*, 2010 Building the sequence map of the human pan-genome. Nature Biotechnology 28**:** 57-63.

Li, Y.-h., G. Zhou, J. Ma, W. Jiang, L.-g. Jin *et al.*, 2014 De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nature Biotechnology 32**:** 1045-1052.

Lin, F., S. H. Wani, P. J. Collins, Z. Wen, W. Li *et al.*, 2020 QTL mapping and GWAS for identification of loci conferring partial resistance to *Pythium sylvaticum* in soybean (Glycine max (L.) Merr). Molecular Breeding 40**:** 1-11.

Liu, D., M. Xin, X. Zhou, C. Wang, Y. Zhang *et al.*, 2017 Expression and functional analysis of the transcription factor-encoding Gene *CsERF004* in cucumber during *Pseudoperonospora cubensis* and *Corynespora cassiicola* infection. BMC Plant Biology 17**:** 96.

Liu, K., 1997 Chemistry and nutritional value of soybean components, pp. 25-113 in Soybeans. Springer.

Liu, S., P. K. Kandoth, S. D. Warren, G. Yeckel, R. Heinz *et al.*, 2012 A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. Nature 492**:** 256.

Liu, Y., H. Du, P. Li, Y. Shen, H. Peng *et al.*, 2020 Pan-genome of wild and cultivated soybeans. Cell 182**:** 162-176. e113.

MacKenzie, K. J., L. G. Sumabat, K. V. Xavier and G. E. Vallad, 2018 A review of *Corynespora cassiicola* and its increasing relevance to tomato in Florida. Plant Health Progress 19**:** 303-309.

Mauch, F., B. Mauch-Mani and T. Boller, 1988 Antifungal Hydrolases in Pea Tissue : II. inhibition of fungal growth by combinations of Chitinase and beta-1,3-glucanase. Plant Physiol 88**:** 936-942.

McCabe, C. E., S. R. Cianzio, J. A. O'Rourke and M. A. Graham, 2018 Leveraging Rna-seq to characterize resistance to brown stem rot and the rbs3 locus in soybean. Molecular Plant-Microbe Interactions 31**:** 1083-1094.

McHale, L., X. Tan, P. Koehl and R. W. Michelmore, 2006 Plant NBS-LRR proteins: adaptable guards. Genome Biology 7**:** 1-11.

Messina, M., and S. Barnes, 1991 The role of soy products in reducing risk of cancer. JNCI: Journal of the National Cancer Institute 83**:** 541-546.

Mostafa, M. M., E. H. Rahma and A. H. Rady, 1987 chemical and nutritional changes in soybean during germination. Food Chemistry 23**:** 257-275.

Mujoo, R., D. T. Trinh and P. K. W. Ng, 2003 Characterization of storage proteins in different soybean varieties and their relationship to tofu yield and texture. Food Chemistry 82**:** 265-273.

Murphy, P. A., 1997 Soybean storage protein composition and tofu quality. Food Technol. 51**:** 86-88.

Nakano, T., K. Suzuki, T. Fujimura and H. Shinshi, 2006 Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol 140**:** 411-432.

Neves, D. L., and C. A. Bradley, 2020 Baseline sensitivity of *Cercospora sojina* and *Corynespora cassiicola* to pydiflumetofen. Crop Protection**:** 105461.

Nishinari, K., Y. Fang, S. Guo and G. O. Phillips, 2014 Soy proteins: A review on composition, aggregation and emulsification. Food Hydrocolloids 39**:** 301-318.

Olive, L. S., D. Bain and C. Lefebvre, 1945 A leaf spot of cowpea and soybean caused by an undescribed species of *Helminthosporium*. Phytopathology 35**:** 822-831.

Onate-Sanchez, L., and K. B. Singh, 2002 Identification of Arabidopsis ethylene-responsive element binding factors with distinct induction kinetics after pathogen infection. Plant Physiol 128**:** 1313-1322.

Onesirosan, P. T., D. C. Arny and R. D. Durbin, 1974 Host specificity of nigerian and north-american isolates of *Corynespora cassiicola*. Phytopathology 64**:** 1364-1367.

Park, S. T., and J. Kim, 2016 trends in next-generation sequencing and a new era for whole genome sequencing. Int Neurourol J 20**:** S76-83.

Ploper, L., and N. d. Ramallo, 1988 Las enfermedades de soja en tucumán: evolución, situación actual y recomendaciones para su control. Revista de la Asociación Argentina de la Soja 8**:** 29-39.

Pusztahelyi, T., 2018 Chitin and chitin-related compounds in plant–fungal interactions. Mycology 9**:** 189-201.

Rai, M. F., E. D. Tycksen, L. J. Sandell and R. H. Brophy, 2018 Advantages of RNA-seq compared to RNA microarrays for transcriptome profiling of anterior cruciate ligament tears. Journal of Orthopaedic Research® 36**:** 484-497.

Rao, M. S., T. R. Van Vleet, R. Ciurlionis, W. R. Buck, S. W. Mittelstadt *et al.*, 2018 Comparison of RNA-Seq and microarray gene expression platforms for the toxicogenomic evaluation of liver from short-term rat toxicity studies. Front Genet 9**:** 636.

Ravelombola, W. S., J. Qin, A. Shi, L. Nice, Y. Bao *et al.*, 2020 Genome-wide association study and genomic selection for tolerance of soybean biomass to soybean cyst nematode infestation. PLoS One 15**:** e0235089.

Rebholz, C., K. Reynolds, M. Wofford, J. Chen, T. Kelly *et al.*, 2013 Effect of soybean protein on novel cardiovascular disease risk factors: a randomized controlled trial. European Journal Of Clinical Nutrition 67**:** 58-63.

Redondo-Cuenca, A., M. Villanueva-Suárez, M. Rodríguez-Sevilla and I. Mateos-Aparicio, 2007 Chemical composition and dietary fibre of yellow and green commercial soybeans (*Glycine max*). Food Chemistry 101**:** 1216-1222.

Reznikov, S., V. De Lisi, P. Claps, V. Gonzalez, M. R. Devani *et al.*, 2019 Evaluation of the efficacy and application timing of different fungicides for management of soybean foliar diseases in northwestern Argentina. Crop Protection 124**:** 104844.

Rondon, M. N., and K. S. Lawrence, 2019 *Corynespora cassiicola* isolates from soybean in Alabama detected with G143A mutation in the cytochrome b gene. Plant Health Progress 20**:** 247-249.

Saio, K., M. Kamiya and T. Watanabe, 1969 Food processing characteristics of soybean 11s and 7s proteins: Part i. Effect of difference of protein components among soybean varieties on formation of tofu-gel. Agricultural and Biological Chemistry 33**:** 1301-1308.

Sarowar, S., S. T. Alam, R. Makandar, H. Lee, H. N. Trick *et al.*, 2019 Targeting the pattern-triggered immunity pathway to enhance resistance to *Fusarium graminearum*. Mol Plant Pathol 20**:** 626-640.

Scherm, H., R. Christiano, P. Esker, E. Del Ponte and C. Godoy, 2009 Quantitative review of fungicide efficacy trials for managing soybean rust in Brazil. Crop Protection 28**:** 774-782.

Schlumbaum, A., F. Mauch, U. Vögeli and T. Boller, 1986 Plant chitinases are potent inhibitors of fungal growth. Nature 324**:** 365-367.

Schmutz, J., S. B. Cannon, J. Schlueter, J. Ma, T. Mitros *et al.*, 2010 Genome sequence of the palaeopolyploid soybean. Nature 463**:** 178.

Seaman, W. L., R. A. Shoemaker and E. A. Peterson, 1965 Pathogenicity of *Corynespora cassiicola* on soybean. Canadian Journal of Botany 43**:** 1461-1469.

Shen, Y., H. Du, Y. Liu, L. Ni, Z. Wang *et al.*, 2019 Update soybean Zhonghuang 13 genome to a golden reference. Sci. China Life Sci 62**:** 1257-1260.

Shimomoto, Y., T. Sato, H. Hojo, Y. Morita, S. Takeuchi *et al.*, 2011 Pathogenic and genetic variation among isolates of *Corynespora cassiicola* in Japan. Plant Pathology 60**:** 253-260.

Shoemaker, R. C., K. Polzin, J. Labate, J. Specht, E. C. Brummer *et al.*, 1996 Genome duplication in soybean (*Glycine* subgenus *soja*). Genetics 144**:** 329-338.

Shook, J. M., J. Zhang, S. E. Jones, A. Singh, B. W. Diers *et al.*, 2021 Meta-GWAS for quantitative trait loci identification in soybean. G3 Genes|Genomes|Genetics 11**:** jkab117.

Shrestha, S. K., K. Lamour and H. Young-Kelly, 2017 Genome sequences and SNP analyses of *Corynespora cassiicola* from cotton and soybean in the southeastern United States reveal limited diversity. PloSone 12**:** e0184908.

Silva, W. P. K., R. L. C. Wijesundera, E. H. Karunanayake, C. K. Jayasinghe and U. M. S. Priyanka, 2000 New hosts of *Corynespora cassiicola* in Sri Lanka. Plant Dis 84**:** 202.

Singh, R. J., and T. Hymowitz, 1999 Soybean genetic resources and crop improvement. Genome 42**:** 605-616.

Singh, S., D. Sehgal, S. Kumar, M. Arif, P. Vikram *et al.*, 2020 GWAS revealed a novel resistance locus on chromosome 4D for the quarantine disease Karnal bunt in diverse wheat pre-breeding germplasm. Scientific Reports 10**:** 1-11.

Song, Q., D. L. Hyten, G. Jia, C. V. Quigley, E. W. Fickus *et al.*, 2013 Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PloSone 8**:** e54985.

Stupar, R. M., and J. E. Specht, 2013 Insights from the soybean (*Glycine max* and *Glycine soja*) genome: past, present, and future. Advances in Agronomy 118**:** 177-204.

Subramanyam, S., C. Zheng, J. T. Shukle and C. E. Williams, 2013 Hessian fly larval attack triggers elevated expression of disease resistance dirigent-like protein-encoding gene, HfrDrd, in resistant wheat. Arthropod-Plant Interactions 7**:** 389-402.

Sumabat, L. G., R. C. Kemerait Jr and M. T. Brewer, 2018 Phylogenetic diversity and host specialization of *Corynespora cassiicola* responsible for emerging target spot disease of cotton and other crops in the southeastern United States. Phytopathology 108**:** 892-901.

Swoboda, C., and P. Pedersen, 2009 Effect of fungicide on soybean growth and yield. Agronomy Journal 101**:** 352-356.

Taira, H., 1990 Quality of soybeans for processed foods in Japan. JARQ, Japan Agricultural Research Quarterly 24**:** 224-230.

Takagi, H., A. Abe, K. Yoshida, S. Kosugi, S. Natsume *et al.*, 2013 QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. The Plant Journal 74**:** 174-183.

Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini *et al.*, 2005 Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proceedings of the National Academy of Sciences 102**:** 13950-13955.

Tonnessen, B. W., P. Manosalva, J. M. Lang, M. Baraoidan, A. Bordeos *et al.*, 2015 Rice *phenylalanine ammonia-lyase* gene OsPAL4 is associated with broad spectrum disease resistance. Plant Mol Biol 87**:** 273-286.

Toufiq, N., B. Tabassum, M. U. Bhatti, A. Khan, M. Tariq *et al.*, 2018 Improved antifungal activity of barley derived chitinase I gene that overexpress a 32 kDa recombinant chitinase in Escherichia coli host. Brazilian Journal Of Microbiology 49**:** 414-421.

USDA, 2020 World agricultural production, pp. in *Circular Series WAP 9-20*. Global market analysis, foreign agricultural service/USDA.

USDA, N. S. R., 2020-2021 Alabama county estimates, soybean, pp., USDA/NASS cooperating with the Alabama Department Of Agriculture.

van der Biezen, E. A., and J. D. Jones, 1998 The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol 8**:** R226-227.

van Ooijen, G., G. Mayr, M. M. Kasiem, M. Albrecht, B. J. Cornelissen *et al.*, 2008 Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. J Exp Bot 59**:** 1383-1397.

Vaux, D. L., 1997 CED-4—the third horseman of apoptosis. Cell 90**:** 389-390.

Vuong, T., H. Sonah, C. Meinhardt, R. Deshmukh, S. Kadam *et al.*, 2015 Genetic architecture of cyst nematode resistance revealed by genome-wide association study in soybean. BMC Genomics 16**:** 1-13.

Wang, J., L. Mao, Z. Zeng, X. Yu, J. Lian *et al.*, 2021 Genetic mapping high protein content QTL from soybean 'Nanxiadou 25'and candidate gene analysis. BMC Plant Biology 21**:** 1-13.

Wang, M., J. Yan, J. Zhao, W. Song, X. Zhang *et al.*, 2012 Genome-wide association study (GWAS) of resistance to head smut in maize. Plant Science 196**:** 125-131.

Wang, Z., M. Gerstein and M. Snyder, 2009 RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10**:** 57-63.

Warner, K., and W. Fehr, 2008 Mid-oleic/ultra low linolenic acid soybean oil: A healthful new alternative to hydrogenated oil for frying. Journal of the American Oil Chemists' Society 85**:** 945.

Wei, C., 1950 Notes on *Corynespora*. Mycological Papers 34.

Wen, C., A. Mao, C. Dong, H. Liu, S. Yu *et al.*, 2015 Fine genetic mapping of target leaf spot resistance gene *cca-3* in cucumber, *Cucumis sativus L*. Theoretical and Applied Genetics 128**:** 2495-2506.

Wen, Z., R. Tan, S. Zhang, P. J. Collins, J. Yuan *et al.*, 2018 Integrating GWAS and gene expression data for functional characterization of resistance to white mould in soya bean. Plant Biotechnology Journal 16**:** 1825-1835.

Wright, D., and A. W. Lenssen, 2013 Staging soybean development.

Xavier, S. A., M. G. Canteri, D. Barros and C. V. Godoy, 2013 Sensitivity of *Corynespora cassiicola* from soybean to carbendazim and prothioconazole. Tropical Plant Pathology 38**:** 431-435.

Xavier, S. A., F. Elis de Mello, H. Prudente da Silva, M. G. Canteri, L. J. Koga *et al.*, 2021 Microtiter method to monitor *Corynespora cassiicola* and sensitivity of the pathogen to carbendazim, prothioconazole and pyraclostrobin. Crop Protection 144**:** 105554.

Xie, J., Q. Wang, Z. Zhang, X. Xiong, M. Yang *et al.*, 2021 QTL-seq identified QTLs and candidate genes for two-seed pod length and width in soybean (*Glycine max*). Plant Breeding 140**:** 453-463.

Xie, M., C. Y.-L. Chung, M.-W. Li, F.-L. Wong, X. Wang *et al.*, 2019 A reference-grade wild soybean genome. Nature Communications 10**:** 1-12.

Xu, X., L. Zeng, Y. Tao, T. Vuong, J. Wan *et al.*, 2013 Pinpointing genes underlying the quantitative trait loci for root-knot nematode resistance in palaeopolyploid soybean by whole genome resequencing. Proceedings of the National Academy of Sciences 110**:** 13469-13474.

Yuan, J., X. Wang, Y. Zhao, N. U. Khan, Z. Zhao *et al.*, 2020 Genetic basis and identification of candidate genes for salt tolerance in rice by GWAS. Scientific Reports 10**:** 1-9.

Zhang, X., W. Wang, N. Guo, Y. Zhang, Y. Bu *et al.*, 2018 Combining QTL-seq and linkage mapping to fine map a wild soybean allele characteristic of greater plant height. BMC Genomics 19**:** 1-12.

Zhao, X., Y. Han, Y. Li, D. Liu, M. Sun *et al.*, 2015 Loci and candidate gene identification for resistance to *Sclerotinia sclerotiorum* in soybean (*Glycine max* L. Merr.) via association and linkage maps. The Plant Journal 82**:** 245-255.

Zhong, C., S. Sun, Y. Li, C. Duan and Z. Zhu, 2018 Next-generation sequencing to identify candidate genes and develop diagnostic markers for a novel *phytophthora* resistance gene, *RpsHC18*, in soybean. Theoretical and Applied Genetics 131**:** 525-538.

Zhou, J.-R., E. T. Gugger, T. Tanaka, Y. Guo, G. L. Blackburn *et al.*, 1999 Soybean phytochemicals inhibit the growth of transplantable human prostate carcinoma and tumor angiogenesis in mice. The Journal of nutrition 129**:** 1628-1635.

Zhou, Z., Y. Jiang, Z. Wang, Z. Gou, J. Lyu *et al.*, 2015a Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. Nature Biotechnology 33**:** 408-414.

Zhou, Z., Y. Jiang, Z. Wang, Z. Gou, J. Lyu *et al.*, 2015b Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. Nat Biotechnol 33**:** 408-414.

Zhu, C., M. Gore, E. S. Buckler and J. Yu, 2008 Status and prospects of association mapping in plants. The Plant Genome 1.

Chapter 2. Evaluating Target spot (*Corynespora cassiicola*) resistance in soybean (*Glycine max* (L.) Merrill) in a controlled environment

Abstract

Target spot, caused by the fungus *Corynespora cassiicola*, has been an increasing problem for soybean growers in the Mid-south and southeastern United States. With favorable environmental conditions, the disease causes premature defoliation and substantial yield losses; therefore, identifying resistant germplasm is critical. Herein, a rapid screening technique for target spot resistance was used to evaluate varieties for possible use in a breeding program. Fifteen soybean varieties, along with a known susceptible variety, were evaluated against isolate *C. cassiicola* LIM01. Two of the sixteen varieties were identified as resistant and nine were moderately susceptible. A subset of these was evaluated with three additional *C. cassiicola* isolates; these isolates differed in combinations of cassiicolin-encoding genes. The soybean varieties Bedford and Council were observed to be resistant to each of the *C. cassiicola* isolates tested in this study. Plant populations developed using these parent varieties could provide a foundation for QTL mapping of regions conferring resistance to target spot and identifying disease resistance mechanisms that can be transferred into elite soybean varieties through marker-assisted breeding.

2.1 Introduction

Soybean (*Glycine max* (L.) Merrill) is an essential oilseed crop and is globally planted to approximately 128 million ha of land with a total production of 364 million metric tons (USDA, 2021). Soybean yield and seed quality can be detrimentally impacted by various fungal diseases (Hartman et al., 2015). Target spot, a foliar disease caused by the fungus *Corynespora cassiicola* (Berk. & M.A. Curtis) C.T. Wei, has emerged as problematic in the Mid-South and

southeastern United States. Target spot was first reported in the U.S. on soybean in 1945 (Olive et al., 1945), and caused between 11 and 32% yield loss during the 1950's (Hartwig, 1959). Until recently, *C. cassiicola* had been considered a minor pathogen (Godoy, 2015). However, damaging target spot outbreaks observed in the U.S., Brazil, Sri Lanka, Canada, and Mexico (Koenning et al., 2006; Edwards Molina et al., 2019; Sacon et al., 2021), have renewed interest in this disease. With yield losses reported at 18 to 32% (Godoy, 2015; Faske, 2016), the U.S. Southern Soybean Disease Workers group started separating target spot loss estimates from other diseases in 2018, indicating the need for additional resources to focus on addressing this disease (Bradley et al., 2021).

Target spot occurrence varies from year to year and site to site, but the disease is more common in regions with warm temperatures and humid conditions (Dixon et al., 2009). Typically, target spot lesions first appear on soybean leaves in the lower and mid-canopy at canopy closure, and with favorable environmental conditions, often quickly spread upward through the plant canopy. Lesions are generally 10 to 15 mm in diameter, with alternating concentric rings of light and dark brown bands and are often surrounded by a yellow halo. Leaves with multiple lesions prematurely senesce (Seaman et al., 1965; Koenning et al., 2006; Godoy, 2015). Stems, seeds, pods, and hypocotyls have also been reported to present symptoms associated with target spot.

The current management strategy for target spot on soybean is the use of fungicides, usually as a preventative application or at the first symptoms of the disease. However, fungicide application might not be an efficient way to control target spot due to the possibility of fungicide resistance development (Xavier et al., 2021) and reduced efficacy of fungicide products (Duan et al., 2019; Rondon and Lawrence, 2019). A better strategy is to develop resistant varieties of

soybean, which can be incorporated into integrated pest management strategies to reduce the potential negative impact of target spot.

Identifying germplasm with resistance to phylogenetically different *C. cassiicola* isolates is necessary. Phylogenetic diversity studies through molecular markers have previously reported that *C. cassiicola* isolates collected from soybean were generally more diverse than isolates collected from other hosts (Sumabat et al., 2018a). Pathogenicity experiments have suggested that isolates are host-specific, but some could cause disease in different hosts. For example, isolates collected from soybean fields could generate a similar level of infection in cotton and tomato (Sumabat et al., 2018a). Cassiicolin is a host-specific toxin produced by *C. cassiicola* (Breton et al., 2000; Déon et al., 2014) and diversity in cassiicolin-encoding (*Cas*) genes in different isolates have been identified as encoding six different protein isoforms (Déon et al., 2014). Different *Cas* genes and their combinations have differential impacts on the host plants. For example, in pathogenicity tests, isolates with *Cas*1 were most aggressive on *Hevea brasiliensis* than were isolates with other *Cas* forms (Déon et al., 2014). Isolates collected from soybean fields have been observed with cassiicolin-encoding genes in combinations of *Cas2*, *Cas6,* or *Cas2+6* and no detectable *Cas* gene (*Cas0*) (Rondon, 2020).

From the 1950's through the 1980's, the soybean breeding program led by Dr. E.E. Hartwig maintained some resistance to *C. cassiicola* in most variety releases (Shurtleff and Aoyagi, 2018). However, it appears that currently grown commercial varieties have little or no such resistance. Teramoto et al. (2013) evaluated twelve soybean cultivars for their response to each of six isolates of *C. cassiicola* and noted that target spot developed on all cultivars, with each of the isolates. Numerous soybean varieties are regularly evaluated in field trials for target spot reactions, and often there are no differences among varieties for susceptibility (Solomon et al.

2018, 2019). Fortunato et al. (2017) have shown that production of phenylpropanoid compounds in soybean may be related to resistance to *C. cassiicola*. Phenylpropanoid compounds have also been associated with tolerance to chlorosis due to iron deficiency (Waters et al., 2018). This is of interest for soybean producers in the western and central parts of Alabama which have high concentrations of thermic *Rendollic Eutrudepts* soils. This soil type has high alkalinity contributing to a high portion of iron deficiency in soybean.

It is hypothesized that the apparent absence of resistance in commercial soybean varieties to *C. cassiicola* may be due to variability in the pathogen population or limited selection for resistance in breeding programs. Therefore, the objective herein is to use a consistent and reproducible procedure for screening soybean varieties against multiple *C. cassiicola* isolates in order to identify horizontal resistance to target spot.

2.2 Material and methods

2.2.1 Isolates

Five *C. cassiicola* isolates were used in this study. Isolates SAR09 and SSTA were obtained from the Brewer lab, Univ. Georgia (Sumabat et al., 2018b). Isolates LIM01, PBU4, and PBU11 were collected from soybean fields in Alabama. Field-collected infected leaf tissue was rinsed in 70% ethanol, washed in 2% NaOCl then rinsed with sterile distilled water before being placed on plates of potato dextrose agar (PDA; 24 gm PDA powder, CRITERION Hardy Diagnostics, Santa Maria, CA, and 12 gm of agar powder in 1,000 ml of deionized water) with kanamycin (50 mg/L). Tissue was maintained at 28°C to obtain colonies (Rondon and Lawrence, 2019). Single colonies of each isolate were transferred to slant tubes containing V8 juice agar (200 ml of V8 juice, 2 gm of $CaCO_3$, 15 gm of agar in 800 ml of deionized water). Inoculated slant tubes were incubated at 28°C for two weeks and then stored in a refrigerator at 4°C.

2.2.2 Morphology: Cultural characteristics and conidial production

To characterize the growth and conidial production of *C. cassiicola,* isolates were grown on each of three culture media: PDA, quarter-strength PDA (Q-PDA; 6 gm of PDA and 12 gm of agar in 1,000 ml of deionized water), and V8 agar medium. A 5-mm piece of 10-day old mycelium of each isolate was transferred to each of 8 plates of each of the three media. Plates were placed in an incubator at 28°C with 12/12 hours of light/dark for 10 to 12 days. Colony color, texture, and conidia production were evaluated after the incubation period. Twelve-day old cultures were flooded with 20 ml of sterile distilled water to determine conidial production. The colony surface was rubbed gently with a sterile glass rod and the spore suspension was filtered twice through a double-layered sterile cheesecloth to remove mycelia. Spores were counted under a microscope using a hemocytometer. The experiment was conducted twice.

To examine the optimal light/dark conditions for spore production, the isolate LIM01 was grown on V8 media under three light regimes: 12 h light/12 h dark, 24 h dark, or 24 h light. For each light regime, five Petri dishes were inoculated and placed in an incubator at 28°C with or without 40 W fluorescent lamps. Plates were grown for 12 days before examining for spore production. Spore counts were done as described above. The experiment was conducted twice.

2.2.3 DNA extraction and Polymerase chain reaction (PCR amplification)

For molecular identification and phylogenetic study, five isolates (LIM01, PBU11, PBU4, SAR09, and SSTA) were grown on V8 agar plates at 28°C for 10 days. Approximately 100 mg fresh mycelium was removed from a culture plate and transferred to 2 ml tubes. Genomic D.N.A. was extracted using E.Z.N.A. Fungal D.N.A. Mini Kit (Omega Bio-tek, Norcross, Georgia) according to the manufacturer's instructions. D.N.A. concentration was determined using a Nanodrop 2000 spectrophotometer (ThermoFisher Sci., Waltham, MA). Samples were diluted to

20 ng/µl and stored at -20°C. Primer pairs amplifying six cassiicolin isoforms (*Cas1*, *Cas2*, *Cas3*, *Cas4*, *Cas5*, and *Cas6*) for understanding variation in pathogenicity and four loci for phylogenetic studies were used for PCR reactions (Table 2.1). A total of 25µl of the PCR reaction mixture was prepared to contain 1.5µl of 20 ng/µl of genomic D.N.A. of each isolate, 0.75 µl of each forward and reverse primers (with a concentration of 10 µM), 2.5 µl of 1X standard reaction buffer, 0.5 µl of dNTP (200 µM), 0.13 µl of D.N.A. Taq polymerase (New England Biolabs, Ipswich, MA) and deionized water to bring the volume to 25 µl. The cycling program consisted of an initial activation step of 95°C for 30 s, followed by 35 cycles of 95°C for 20 s; 1 min annealing at 56°C, and extension of 1 min at 68°C, followed by a final extension step of 5 min at 68°C. An agarose gel containing 1.5% ethidium bromide solution was run to visualize the PCR product. The PCR product was purified using E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek) and sent to Eurofins M.W.G. Operon L.L.C. for sequencing (Louisville, KY).

2.2.4 Phylogenetic analysis

For phylogenetic analysis, four loci (*act*, *caa5*, *ga4*, and ITS) were sequenced from each of the five isolates of *C. cassiicola* used in the current study; sequences of an additional 32 isolates for the same four loci were downloaded from NCBI data (Table 2.2). The final phylogenetic tree was obtained from combined data of the four loci. The CLUSTAL W program on default settings was used to align the sequences of 37 isolates (Thompson et al., 1994). The alignment data were converted into a mega file and loaded to Mega X to perform phylogenetic analyses. Initial analysis identified Hasegawa-Kishino-Yano with the lowest BIC scores (Bayesian Information Criterion) as the best model for the analysis.

The evolutionary history was evaluated using the Maximum Likelihood method and the Hasegawa-Kishino-Yano model with Neighbor-Join and BioNJ algorithms to generate an initial

tree for a heuristic search (Nei and Kumar, 2000). A total of 1,000 replications of this method were performed using the bootstrap method, and the final tree was selected with the highest log likelihood value. The final tree was developed using the TEN1 isolate as an outgroup and rooting the tree (Kumar et al., 2018).

2.2.5 Characterizing resistant varieties:

For this experiment, the material used is referred to as a variety due to the age of the release and no longer available for commercial production, whereas the term cultivar is considered available for commercial production. Thirteen soybean varieties, identified as resistant to target spot in the National Germplasm Database (GRIN) (Table 2.3), were selected. Two additional varieties, Council and Pembina, with tolerance to chlorosis due to iron deficiency (Helms and Halvorson, 1996; Helms et al., 2006) were selected to explore the possible connection between tolerance to iron deficiency chlorosis and target spot resistance. 'Henderson' (PI 665225), susceptible to target spot, was also included as a control (Dr. Koebernick, personal comm.).

Greenhouse trials were done at the Plant Science Research Center, Auburn University. Three seeds of each of sixteen varieties were sown into five separate 11.5 × 11.5 cm^2 pots filled with potting mix (PRO-MIX BX, Premier Tech Horticulture, Quakertown, PA). Five days after emergence, seedlings were thinned to one per pot. Plants were grown for twenty-five (25) days in the greenhouse with a day length of 14 hours per day; 1000-watt halide bulbs provided supplemental light, and ambient temperatures ranged from 24°C to 34°C. Plants were watered as needed prior to inoculation at growth stage V3-V4 (Onesirosan et al., 1974) with isolate LIM01 which was arbitrarily selected for use in this study. This experiment was done three times with five replications in each trial.

Mycelium of isolate LIM01, stored on a V8 agar slant, was transferred to V8 agar plates between twelve and fourteen days before inoculation. The plates were placed in an incubator at 28°C with 12/12 hours of light/dark. Each plate was flooded with sterile distilled water, the colony surface was disturbed by rubbing with a glass rod to dislodge conidia, and the conidial suspension was filtered with double-layered sterile cheesecloth to remove mycelium. Conidia were counted using a hemocytometer and adjusted to 50,000 per ml (Onesirosan et al., 1974; Fortunato et al., 2015).

Prior to inoculation, plants were sprayed with 0.05% of a Tween 20 solution and allowed to dry for approximately 5 minutes. The conidial suspension was then sprayed onto axial and abaxial leaf sides using a fine mist professional spray bottle (Spray Pro) until run-off. The plants were transferred into a plastic mist chamber inside the greenhouse with a non-inoculated control of each variety (sprayed with distilled water). The mist chamber was constructed of a 2.54-centimeter PVP pipe (4 m × 4 m × 8 m) and covered with a transparent plastic sheet. Plants were arranged in a completely randomized design (CRD) in the mist chamber; mist ran for two seconds every 10 minutes for three days to maintain high humidity (> 90%). The temperature was maintained at 25 to 30°C in the mist chamber. The plants were kept in the mist chamber for 14 days and were watered as needed.

2.2.6 Evaluating isolate differences

Eight soybean varieties that differed in their apparent response (four susceptible, two moderately susceptible, and two apparently resistant) to LIM01 were selected from the 16 initially evaluated. These varieties were inoculated with each of the three *C. cassiicola* isolates (PBU11, SSTA, and SAR09) to evaluate possible variability in disease response due to isolate.

Methods for plant growth and maintenance, as well as inoculation of plants, were conducted as described above. The experiment was conducted three times, with four replications in each trial.

2.2.7 Disease assessment

Plants were assessed for target spot severity 14 days after inoculation. The severity rating was based on the percentage of damaged leaf area according to a visual diagram on the scale of 0 to 90% damage (Vincelli and Hershman, 2011). A variety was considered susceptible if disease severity exceeded 50%, between 25% to 50% disease severity were considered moderately susceptible to the *C. cassiicola*, and resistant varieties had disease severity below 25%.

2.2.8 Data analysis

All experiments were analyzed with PROC GLIMMIX (Generalized linear mixed model) in SAS (version 9.4; SAS Institute Inc., Cary, NC). Trial effect was a fixed effect in preliminary analyses of each study and was not significant at $P > 0.53$. Therefore, for further analyses, data from all trials of each study were combined. For conidia production of *C. cassiicola* on three different media; medium, isolate, and medium × isolate was treated as fixed, with replication as a random effect. In the light effect study, different light regimes were treated as fixed with replication as a random effect. For the greenhouse experiment, with isolate LIM01 and 16 soybean varieties; variety was kept as a fixed effect, replication was a random effect. The study with three isolates and eight varieties; the fixed effects were variety, isolate, and variety × isolate with replication as random. In all studies, differences between fixed effects were considered significant based on Fisher's protected least significant difference at $P \leq 0.05$.

2.3 Results

2.3.1 Cultural attributes and isolate diversity

Differences in mycelial growth and morphology were noted between *C. cassiicola* isolates grown on PDA, Q-PDA, and V8 culture media. In general, isolates on V8 had grey to whitish-grey with green or orange pigmentation on the upper surface, and dark orange bottom surfaces (Figure 2.1). The upper surfaces of isolates on PDA were light to dark grey with abundant aerial mycelia, while the bottom surfaces were pale white to black. The growth of all isolates on Q-PDA was reduced with less abundant aerial mycelia than on PDA; upper surfaces were whitish to grey while bottom surfaces were light cream to pale yellow or brown (Figure 2.1).

The greatest numbers of conidia were produced on V8 agar as compared to other media (P <0.0001) (Figure 2.2 A). On V8 media, the isolate SSTA produced more than twice the numbers of spores compared to SAR09, which produced substantially more spores than LIM01 and PBU11. In contrast, PBU4 grew slowly on all media and produced a negligible number of spores on V8 media (Figure 2.2 B). Greatest spore production was noted with 12 h light/12 h dark (P <0.0001) than 24 h dark or 24 h light (Figure 2.3).

2.3.2 Phylogenetic analysis

The BLAST search for amplified sequences of the Internal Transcribed Spacer (ITS) region for different isolates showed 99% similarity to ITS sequences of *C. cassiicola* in the NCBI database. The isolates SAR09 and LIM01 contained only the *Cas2* gene, while PBU4, PBU11, and SSTA contained *Cas2+Cas6* genes (Table 2.2). The *Cas6* only gene was not present in any of the five isolates. Diversity among the isolates was also assessed by sequencing four loci (ITS, *caa5*, *ga4*, and *act1*) and conducting phylogenetic analysis using the four previously sequenced isolates (Déon et al., 2014; Rondon, 2020). This analysis showed that 37 isolates generated five sister groups with isolates tested in this experiment distributed in three of the sister groups (Table 2.2).

Isolates PBU4 and PBU11(*Cas2+6)* clustered with ELM04 (*Cas2+6*) and ELM07 (*Cas6*) whereas SSTA, carries *Cas2+6*, was in the sister group, with isolates only having *Cas2+6* (Figure 2.4). SAR09 and LIM01 (*Cas2* only), clustered with isolates having either *Cas2* or *Cas0*. In addition, none of these soybean isolates clustered with cotton isolates (Figure 2.4).

2.3.3 Identifying target spot resistance soybean varieties

Disease severity differed significantly ($P < 0.0001$) among varieties inoculated with isolate LIM01. Severity ranged from 7% on Council to 64.4% on Henderson (Table 2.3, Figure 2.5). Based on the overall mean of the three trials, six varieties were rated as susceptible, eight were moderately susceptible, and two (Bedford and Council) resistant to *C. cassiicola* (Table 2.3).

Eight varieties screened against additional *C. cassiicola* isolates (PBU11, SSTA, and SAR09) showed that there were no significant differences due to isolate ($P = 0.06$) or the soybean variety x isolate interaction ($P = 0.50$). Council followed by Bedford had significantly less disease when compared to other soybean varieties and were categorized as resistant to each tested *C. cassiicola* isolate. Sumter had significantly greater disease severity than Henderson with SSTA but was not significantly different from Henderson when inoculated with PBU11 or SAR09 (Table 2.4). Also, McNair 600 was similar to Pembina when inoculated with LIM01 and PBU11 but had lower disease when inoculated with SAR09 (Table 2.3, Table 2.4).

2.4 Discussion

The current study focuses on optimizing fungal development and subsequent spore production along with inoculation conditions needed to develop a rapid, simple, and reliable protocol to screen soybean germplasm for resistance to target spot in a controlled environment at the seedling stage. *C. cassiicola* grown on V-8 agar with 12 hours light and 12 hours dark had the

best growth and produced the greatest number of conidia which is consistent with previous observations (Rondon, 2020). Modifications to an existing protocol to evaluate *C. cassiicola* on cotton (Moore et al., 2021) were successful at identifying soybean genotypic responses. The primary differences between the work on cotton and soybean involved increasing the misting time (from one to two seconds every 10 mins) and priming plants with Tween 20 solution prior to inoculation. In previous studies (Moore et al., 2021), inoculum suspensions did not adequately adhere to cotton leaves. The surfactant Tween 20 prolongs the adherence of water on leaves and increases the chances of fungal infection by preventing immediate runoff of the spore suspension. In addition, the inoculum concentration was increased from 40,000 spores/ml to 50,000 spores/ml. The greater conidial concentration used herein produced distinct disease symptoms and allowed different resistance levels to be distinguished among varieties which is consistent with work by Onesirosan et al. (1974) and Fortunato et al. (2015).

Cassiicolin is a glycoprotein which contributes to isolate pathogenicity (Breton et al., 2000; Déon et al., 2012; Déon et al., 2014). Six isoforms of this protein have been identified in different *C. cassiicola* isolates capable of infecting a wide range of plant species, but only two of the currently identified isoforms, *Cas2* and *Cas6*, were present in isolates collected from soybean (Déon et al., 2014; Rondon, 2020). The five isolates used herein either had *Cas2* (LIM01 and SAR09) or *Cas2+6* (SSTA, PBU4, and PBU11) cassiicolin-encoding genes. Additionally, the five isolates used herein were distributed in three of five possible sister groups containing soybean isolates in a phylogenetic tree previously developed from 44 different isolates collected from cotton and soybean fields in Georgia, Tennessee, Alabama, and Brazil (Rondon, 2020). It may be that the pathogen population has changed in recent years, as have resistance reactions, suggesting that candidate germplasm should be screened against multiple *C. cassiicola* isolates

prior to integration into a breeding program as a disease resistant line. Preparing inoculum suspensions containing a combination of several isolates of *C. cassiicola* to infect soybean plants may expedite the screening process. Such an approach has been adapted to identify resistance in oilseed crops against *Pseudocercosporella capsellae* (Eshraghi et al., 2007).

Two varieties, Council and Bedford, had consistently lower disease than other varieties with each of the four isolates used in the current study. These four *C. cassiicola* isolates clustered into three different phylogenetic groups based on four loci (ITS, *caa5*, *ga4*, and *act1*), and differed in *Cas* genes (two isolates had *Cas2*, another two had *Cas2+6*). Consistency of the low disease reactions with Council and Bedford suggests that these varieties are sources for resistance that may be used in a breeding program to develop target spot resistant soybean germplasm.

In soybean, the phenylpropanoid pathway may play a major role in providing resistance to target spot (Fortunato et al., 2017). Studies conducted on iron chlorosis have also suggested the involvement of phenylpropanoid compounds in providing tolerance to iron chlorosis (Waters et al., 2018). However, of the two varieties we looked at with tolerance to iron-deficiency chlorosis, one (Council) showed resistance to *C. cassiicola* herein, while the other (Pembina) was susceptible to the pathogen. Further research is required to investigate the link between these traits.

A search for soybean target spot resistance in the GRIN database identified 13 varieties but had no supporting information as to how the varieties were scored or evaluated. When inoculated with isolate LIM01, 12 out of 13 varieties were scored as susceptible or moderately susceptible. It may be that the apparent resistance noted in the GRIN database was due to field evaluations in less conducive prevailing conditions then was used herein. *C. cassiicola* has been reported in tropical and subtropical regions with warm temperatures and excess humidity (Farr and

Rossman, 2017; Sumabat et al., 2018a). Field studies can sometimes be misleading when pathogen pressure is low and suboptimal environmental conditions prevail, thus making it more challenging to distinguish between a susceptible and a resistant variety. For example, a field study at Milan, TN, which claims to have high disease pressure for target spot, reported disease severity less than 25% on any variety (e.g., Zuchelli et al. 2016a, 2016b). Increased aggressiveness of a pathogen isolate could cause higher disease levels, or apparently lower resistance of a specific cultivar (Cowger and Mundt, 2002). These 12 varieties were registered before 1996; therefore, it is possible that *C. cassiicola* may have not had its current level of aggressiveness on soybean at that time (Sumabat et al., 2018a). The one resistant variety noted in this study, Bedford, was developed by soybean breeder, Edgar Hartwig (Hartwig and Epps, 1978), who maintained some level of disease resistance to *C. cassiicola* in all of his variety releases (Hartwig, 1976; Shurtleff and Aoyagi, 2018).

To date, no study has been conducted to identify the genomic region or genetic mechanism involved in soybean resistance to *C. cassiicola*. Molecular studies using Council and/or Bedford as one of the parents will identify genomic regions associated with *C. cassiicola* resistance and markers to integrate a resistance trait into an elite line. Additionally, soybean varieties in the current study show different levels of resistance indicating the possibility of different regions and mechanisms involved in providing such resistance. If needed, such varieties can be used to conduct QTL or gene pyramiding for improving resistance against target spot.

2.5 Conclusion

A modified rapid method specific for soybean successfully identified varieties with resistance to *C. cassiicola* in a controlled environment. Two varieties, Bedford and Council, were identified as resistant due to low disease severity when inoculated with phylogenetically

different *C. cassiicola* isolates. Future field experiments under *C. cassiicola* pressure must be conducted to ensure that resistance to *C. cassiicola* is persistent in these varieties before using them in a breeding program or conducting QTL studies to identify genomic regions and resistance mechanisms to target spot. Further omics research is required to evaluate the possible role of the phenylpropanoid pathway on iron chlorosis tolerance and target spot resistance and explore possibilities of any shared genetic network and mechanism for these traits.

References

Bradley, C.A., Allen, T.W., Sisson, A.J., Bergstrom, G.C., Bissonette, K.M., Bond, J., Byamukama, E., Chilvers, M.I., Collins, A.A., Damicone, J.P., Dorrance, A.E., Dufault, N.S., Esker, P.D., Faske, T.R., Fiorellino, N.M., Giesler, L.J., Hartman, G.L., Hollier, C.A., Isakeit, T., Jackson-Ziems, T.A., Jardine, D. J., Kelly, H.M., Kemerait, R.C, Kleczewski, N.M., Koehler, A.M., Koenning, S.R., Kratochvil, R.J., Kurle, J.E., Malvick, D.K., Markell, S.G., Mathew, F.M., Mehl, H.L., Mehl, K.M., Mueller, D.S., Mueller, J.D., Nelson, B.D., Overstreet, C., Padgett, G.B., Price, P.P., Sikora, E.J., Small, I., Smith, D.L., Spurlock T.N., Tande, C.A., Telenko, D.E.P.,Tenuta, A.U., Thiessen, L.D., Warner, R., Wiebold, W.J., Wise, K.A., 2021. Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2015 to 2019. Plant Health Progress. Doi.org/10.1094/PHP-01-21-0013-RS.

Breton, F., Sanier, C., D'Auzac, J., 2000. Role of cassiicolin, a host-selective toxin, in pathogenicity of *Corynespora cassiicola*, causal agent of a leaf fall disease of Hevea. J. Rubber Research 3, 115-128.

Carbone, I, Kohn M, L, 1999. A method for designing primer sets for speciation studies in filamentous ascomycetes. Mycologia 91, 553–556.

Cowger, C., Mundt, C.C. 2002. Aggressiveness of *Mycosphaerella graminicola* isolates from susceptible and partially resistant wheat cultivars. Phytopathology 92, 624-630.

Déon, M., Bourré, Y., Gimenez, S., Berger, A., Bieysse, D., De Lamotte, F., Poncet, J., Roussel, V., Bonnot, F., Oliver, G., 2012. Characterization of a cassiicolin-encoding gene from *Corynespora cassiicola*, pathogen of rubber tree (*Hevea brasiliensis*). Plant Science 185, 227-237.

Déon, M., Fumanal, B., Gimenez, S., Bieysse, D., Oliveira, R.R., Shuib, S.S., Breton, F., Elumalai, S., Vida, J.B., Seguin, M., 2014. Diversity of the cassiicolin gene in *Corynespora cassiicola* and relation with the pathogenicity in *Hevea brasiliensis*. Fungal Biology 118, 32-47.

Dixon, L.J., Schlub, R.L., Pernezny, K., Datnoff, L., 2009. Host specialization and phylogenetic diversity of *Corynespora cassiicola*. Phytopathology 99, 1015-1027.

Duan, Y., Xin, W., Lu, F., Li, T., Li, M., Wu, J., Wang, J., Zhou, M., 2019. Benzimidazole-and QoI-resistance in *Corynespora cassiicola* populations from greenhouse-cultivated cucumber: An emerging problem in China. Pesticide Biochemistry and Physiology 153, 95-105.

Edwards Molina, J., Paul, P., Amorim, L., Da Silva, L., Siqueri, F., Borges, E., Campos, H., Venancio, W., Meyer, M., Martins, M., 2019. Effect of target spot on soybean yield and factors affecting this relationship. Plant Pathology 68, 107-115.

Eshraghi, L., Barbetti, M., Li, H., Danehloueipour, N., Sivasithamparam, K., 2007. Resistance in oilseed rape (*Brassica napus*) and Indian mustard (*Brassica juncea*) to a mixture *of*

Pseudocercosporella capsellae isolates from Western Australia. Field Crops Research 101, 37-43.

Farr, D.F., Rossman, A.Y., 2017. Fungal Databases, U.S. National Fungus Collections, ARS, USDA. Online publication. United States Department of AgricultureAgricultural Research Service. https://nt.ars-grin.gov/fungaldatabases/

Faske, T., 2016. Target spot of soybean: What Do We Know? Arkansas Row Crop.https://agfax.com/2016/11/02/arkansas-soybeans-target-spot-what-do-we-know/

Fortunato, A.A., Araujo, L., Rodrigues, F.Á., 2017. Association of the production of phenylpropanoid compounds at the infection sites of *Corynespora cassiicola* with soybean resistance against target spot. Journal of Phytopathology 165, 131-142.

Fortunato, A.A., Debona, D., Bernardeli, A.M., Rodrigues, F.A., 2015. Defence-related enzymes in soybean resistance to target spot. Journal of Phytopathology 163, 731-742.

Godoy, C.V., 2015. Target spot. Pp. 62-63 In: Hartman, G.L., Rupe, J.C., Sikora, E.J., Domier, L.L., Davis, J.A., Steffey, K.L. (Eds.), Compendium of Soybean Diseases and Pests. APS Press, St. Paul, MN.

Hartman, G.L., Rupe, J.C., Sikora, E.J., Domier, L.L., Davis, J.A., Steffey, K.L., 2015. Compendium of Soybean Diseases and Pests. APS Press, St. Paul, MN.

Hartwig, E.E., 1959. Effect of target spot on yield of soybeans. Plant Dis. Reptr. 43:504-505.

Hartwig, E.E., 1976. Breeding soybeans resistant to diseases. International Workshop on Grain Legumes, ICRISAT. http://oar.icrisat.org/645/1/RA_00019.pdf

Hartwig, E., Epps, J., 1978. Registration of Bedford Soybeans (Reg. No. 118). Crop Science 18, 915.

Helms, T., Halvorson, M., 1996. Registration of 'Council' soybean. Crop Science 36:206.

Helms, T., Nelson, B., Goos, R., 2006. Registration of 'Pembina' soybean. Crop Science 46,469-471.

Koenning, S., Creswell, T., Dunphy, E., Sikora, E., Mueller, J., 2006. Increased occurrence of target spot of soybean caused by *Corynespora cassiicola* in the Southeastern United States. Plant Dis. 90, 974-974.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35, 1547-1549.

Moore, C., Sharma, N., Bowen, K.L., Koebernick, J., 2021. Developing a greenhouse protocol for evaluating resistance to *Corynespora cassiicola* in cotton. Plant Dis. 105:2351-2355.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, NY.

Olive, L.S., Bain, D., Lefebvre, C., 1945. A leaf spot of cowpea and soybean caused by an undescribed species of Helminthosporium. Phytopathology 35, 822-831.

Onesirosan, P.T., Arny, D.C., Durbin, R., 1974. Host specificity of Nigerian and North American isolates of *Corynespora cassiicola*. Phytopathology 64, 1364-1367.

Rondon, M.N., 2020. Genetic diversity of the fungal pathogen *Corynespora cassiicola* and its fungicide resistance. Dissertation, Auburn Univ., AL https://etd.auburn.edu/handle/10415/7505.

Rondon, M.N., Lawrence, K.S., 2019. *Corynespora cassiicola* isolates from soybean in Alabama detected with G143A mutation in the cytochrome b gene. Plant Health Progress 20, 247-249.

Sacon, D., Bottan, A.J., Zanella, L.C., Netto, A., Borges, L.L., 2021. *Corynespora cassiicola* affecting soybean. Communications in Plant Sci. 11, 039-045.

Seaman, W., Shoemaker, R., Peterson, E., 1965. Pathogenicity of *Corynespora cassiicola* on soybean. Canadian Journal of Botany 43, 1461-1469.

Shurtleff, W., Aoyagi, A., 2018. Biography of Edgar E. Hartwig, Soybean Breeder for the U.S. South (1913-2018): Extensively annotated bibliography and sourcebook. Soyinfo Ctr., Lafayette, CA.

Solomon, W.L., Allen, T.W., Burgess, B.A., 2019. Field evaluation of maturity group IV late soybean cultivars to Cercospora leaf blight, frogeye leaf spot, Septoria brown spot, southern blight, and target spot in east MS, 2018. Plant Dis. Manag. Repts. 13:CF097.

Solomon, W.L., Allen, T.W., Burgess, B.A., 2018. Field evaluation of maturity group V early soybean cultivars to Cercospora leaf blight, frogeye leaf spot, Septoria brown spot, target spot, and sudden death syndrome in west MS, 2017. Plant Dis. Manag. Repts. 12:CF174.

Sumabat, L.G., Kemerait Jr, R.C., Brewer, M.T., 2018a. Phylogenetic diversity and host specialization of *Corynespora cassiicola* responsible for emerging target spot disease of cotton and other crops in the southeastern United States. Phytopathology 108, 892-901.

Sumabat, L.G., Kemerait Jr, R.C., Kim, D.K., Mehta, Y.R., Brewer, M.T., 2018b. Clonality and geographic structure of host-specialized populations of *Corynespora cassiicola* causing emerging target spot epidemics in the southeastern United States. Plos one 13, e0205849.

Teramoto, A., Machado, T.A., Santos, L.M.d., Volf, M.R., Meyer, M.C., Cunha, M.G.d., 2013. Reação de cultivares de soja à *Corynespora cassiicola*. Tropical Plant Pathology 38, 68-71.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680.

USDA, 2021. World agricultural production:Foreign Agricultural Service/USDA. Global Market Analysis.Circular Series WAP July 2021. https://usda.library.cornell.edu/concern/publications/5q47rn72z?locale=en

Vincelli, P., Hershman, D.E., 2011. Assessing foliar diseases of corn, soybeans, and wheat. Cooperative Extension Service, Univ. Kentucky, Publ. PPFS-MISC-06.

Waters, B.M., Amundsen, K., Graef, G., 2018. Gene expression profiling of iron deficiency chhlorosis sensitive and tolerant soybean indicates key roles for phenylpropanoids under alkalinity stress. Front. Plant Sci. doi:10.3389/fpls.2018.00010.

White, T., et al., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: A guide to Methods and Applications 18, 315–322.

Xavier, S.A., Canteri, M.G., Barros, D.C.M., Godoy, C.V., 2021. Sensitivity of *Corynespora cassiicola* from soybean to carbendazim and prothioconazole. Trop. Plant Pathol. 38:431-435.
Zuchelli, E., Crower, W., Kelly, H.M., 2016a. Field evaluation of maturity group III soybean cultivars to frogeye leaf spot, Septoria brown spot, target spot and fungicide in Tennessee, 2016. Plant Dis. Manag. Repts. 12:CF096.

Zuchelli, E., Crower, W., Kelly, H.M., 2016b. Field evaluation of maturity group IV early soybean cultivars to frogeye leaf spot, Septoria brown spot, target spot and fungicide in Tennessee, 2016. Plant Dis. Manag. Repts. 12:CF095.

Figure 2.1. Colony morphology of five *Corynespora casiicola* isolates (LIM01, SAR09, SSTA, PBU11, PBU4) growing on each of three media: (A) V8 agar, (B) potato dextrose agar, (C) quarter-strength potato dextrose agar, following 10 days incubation.

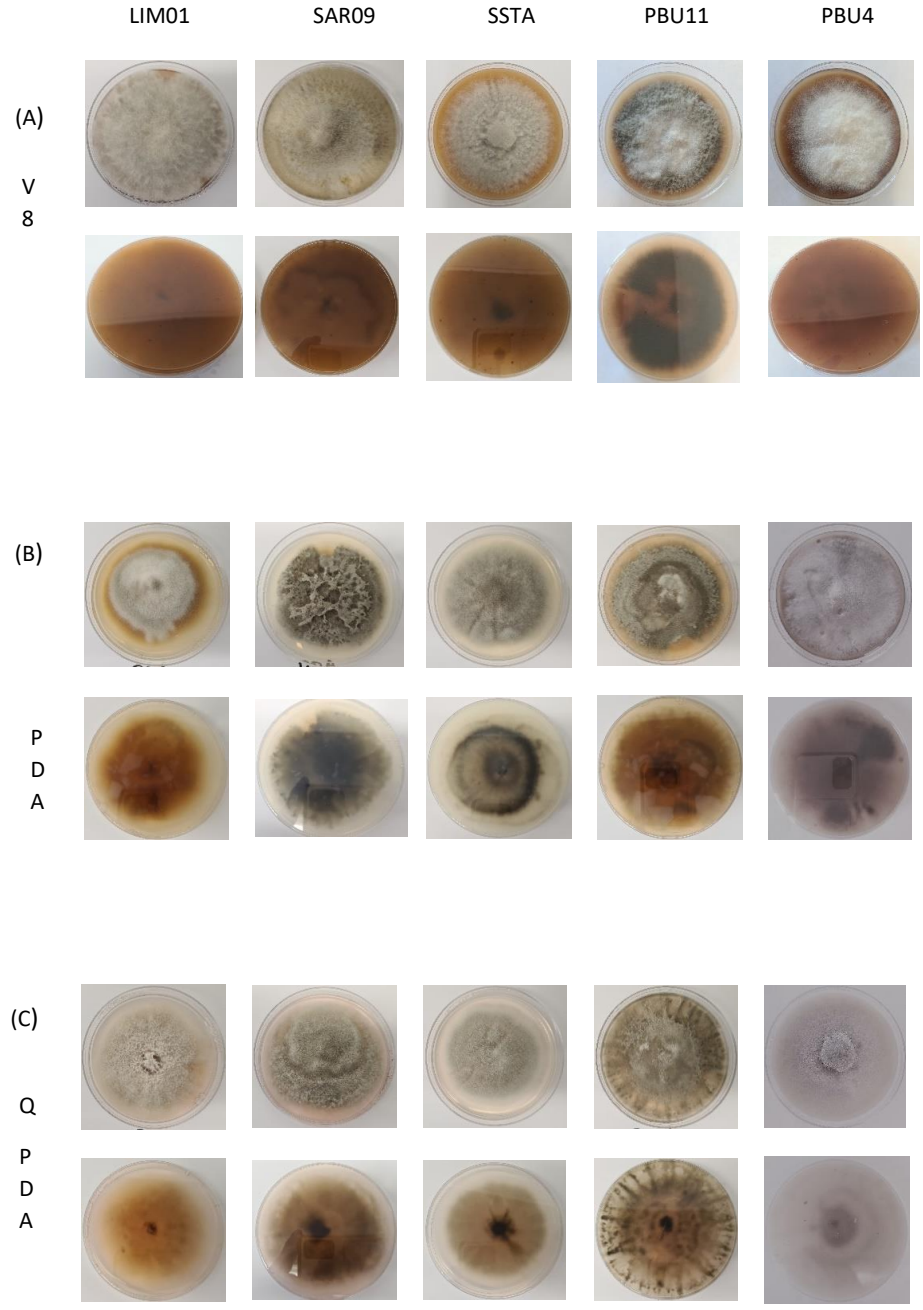


Figure 2.2 Conidia production of *Corynespora casiicola* on (A) different media: potato dextrose agar (PDA), quarter-strength PDA (QPDA), and V8 agar, and (B) on V8 agar for each of five *C. cassiicola* isolates.

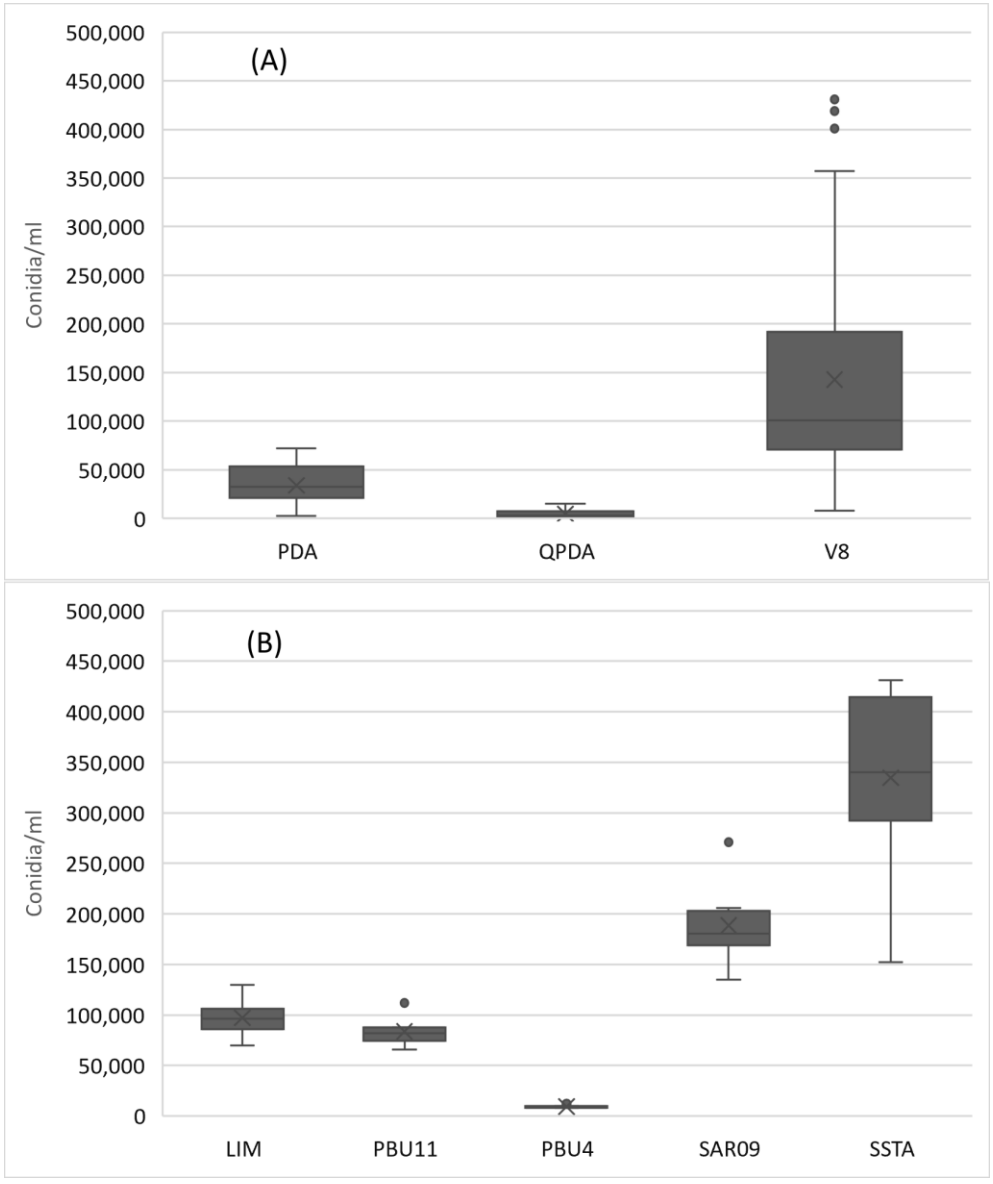


Figure 2.3 Conidia production by LIM01 *Corynespora cassiicola* isolate on V8 media under different light regimes. 'X' inside the boxes is the mean value and horizontal line is median value.

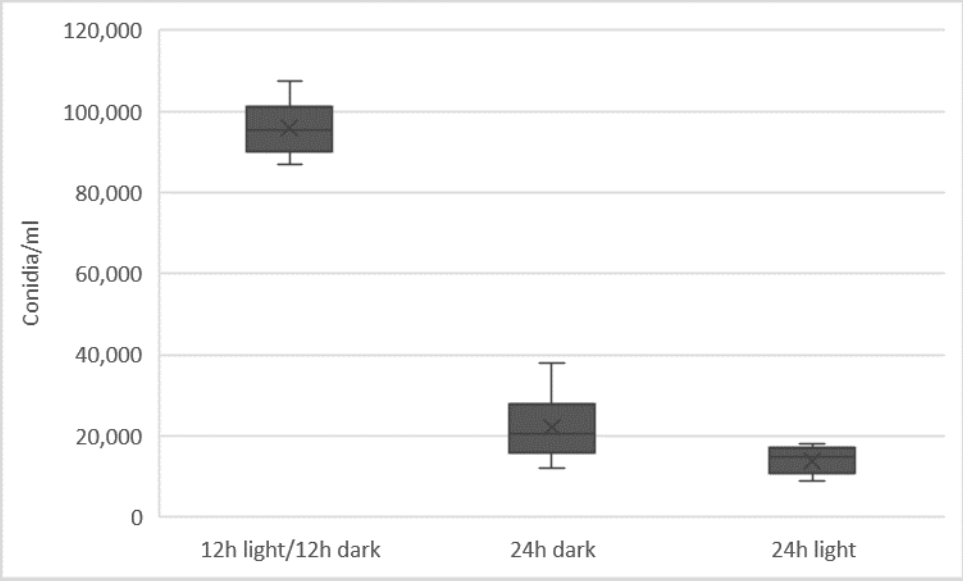


Figure 2.4. Phylogenetic analysis of *Corynespora cassiicola* isolates. The tree was created from the combined data of four different loci, *act1*, *caa5*, ITS, and *ga4*, using maximum likelihood method in MEGA X. Symbol * represents isolates from soybean and ¥ symbol represents isolates from cotton. The isolates outlined in blue boxes were used in current studies.

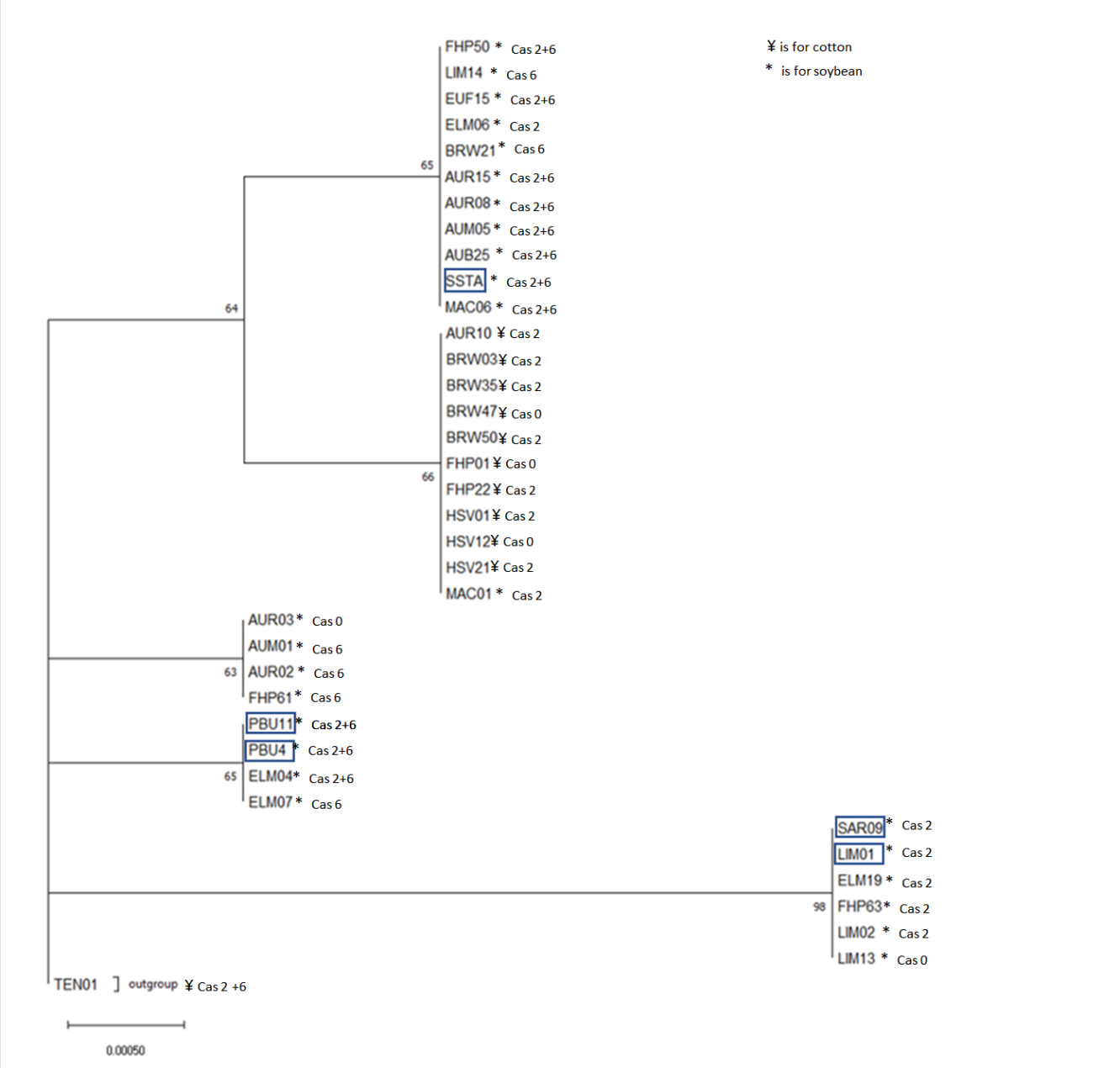


Figure 2.5. Disease ratings of LIM01 *Corynespora cassiicola* isolate on soybean genotypes, 14 days after inoculation. Symptoms of target spot on: (A) Resistant genotype, (B) Susceptible genotype leaf and stem

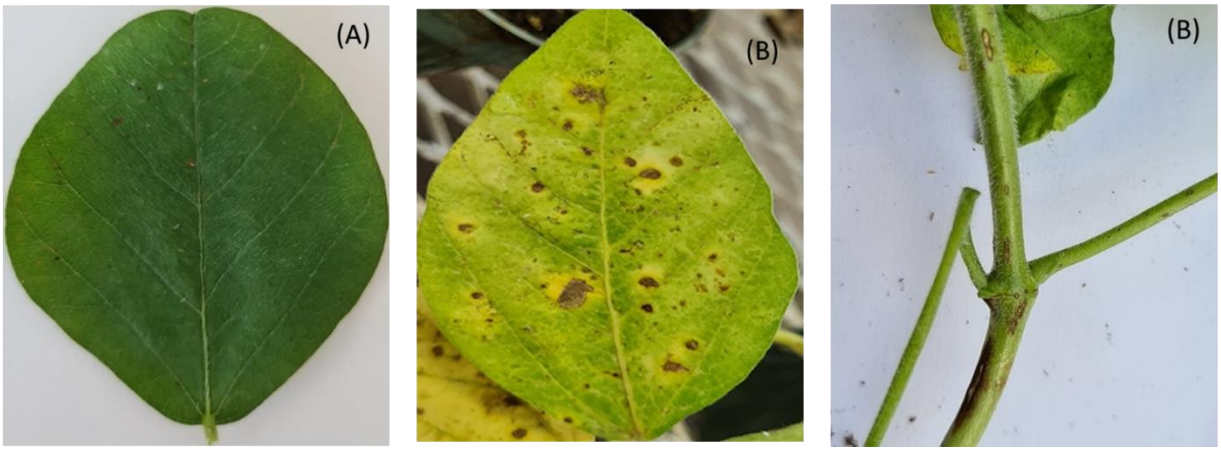


Table 2.1. List of primers used for PCR

Genes	Forward sequence	Reverse sequence	Reference
ITS	TCCGTAGGTGAACCTGCGG	TCCTCCGCTTATTGATATGC	White et al., 1990
caa5	GTCCACAAGTGGAACCTCGT	CCTCGTCTGCCAGTTCTTCT	Dixon et al., 2009
cas2	GGATTTGCCTGAGATCCTA	CAAACAATGCTAACCAAACAAAC	Déon et al., 2014
act1	ATGTGCAAGGCCGGTTTCGC	TACGAGTCCTTCTGGCCCAT	Carbone and Kohn, 1999
gaa4	CCTGCTCCGACTTTGTTGAG	GTCTGGGAGCAGCAAAGACT	Dixon et al., 2009

Table 2.2 List of *Corynespora cassiicola* isolates used for phylogenetic analysis.

No.	Isolate	Class	Geographic Source	Host plant	GenBank accession number			
					act1	cca5	ga4	rDNA ITS
1	SSTA	Cas2+6	Tift County, GA	*Gly. max*	OM105585	OM033040	OM033035	OM033467
2	SAR09	Cas2	Tift County, GA	*Gly. max*	OM105584	OM033039	OM033034	OM033466
3	LIM01	Cas2	Limestone County, AL	*Gly. max*	OM105581	OM033036	OM033031	OM033463
4	PBU4	Cas2+6	Elmore County, AL	*Gly. max*	OM105582	OM033037	OM033032	OM033464
5	PBU11	Cas2+6	Elmore County, AL	*Gly. max*	OM105583	OM033038	OM033033	OM033465
6	LIM02[Z]	Cas2	Limestone County, AL	Gly. max	MT820829	MT820859	MT820891	MT822673
7	LIM13[Z]	Cas0	Limestone County, AL	*Gly. max*	MT820830	MT820860	MT820892	MT822674
8	LIM14[Z]	Cas6	Limestone County, AL	*Gly. max*	MT820831	MT820861	MT820893	MT822675
9	ELM04[Z]	Cas2+6	Elmore County, AL	*Gly. max*	MT820821	MT820846	MT820878	MT822660
10	ELM06[Z]	Cas2	Elmore County, AL	*Gly. max*	MT820824	MT820847	MT820879	MT822661
11	ELM07[Z]	Cas6	Elmore County, AL	*Gly. max*	MT820822	MT820848	MT820880	MT822662
12	MAC01[Z]	Cas2	Macon County, AL	*Gly. max*	MT820832	MT820862	MT820894	MT822676
13	AUB25[Z]	Cas2+6	Auburn, AL	*Gly. max*	MT820801	MT820833	MT820865	MT822647
14	AUM01[Z]	Cas6	Auburn, AL	*Gly. max*	MT820802	MT820834	MT820866	MT822648
15	AUM05[Z]	Cas2+6	Auburn, AL	*Gly. max*	MT820803	MT820835	MT820867	MT822649
16	AUR02[Z]	Cas6	Auburn, AL	*Gly. max*	MT820804	MT820836	MT820868	MT822650
17	AUR03[Z]	Cas0	Auburn, AL	*Gly. max*	MT820805	MT820837	MT820869	MT822651
18	AUR08[Z]	Cas2+6	Auburn, AL	*Gly. max*	MT820806	MT820838	MT820870	MT822652
19	EUF15[Z]	Cas2+6	Eufaula, AL	*Gly. max*	MT820814	MT820850	MT820882	MT822664
20	BRW21[Z]	Cas2+6	Brewton, AL	*Gly. max*	MT820809	MT820842	MT820874	MT822656
21	FHP50[Z]	Cas2+6	Fairhope, AL	*Gly. max*	MT820815	MT820853	MT820885	MT822667
22	FHP61[Z]	Cas6	Fairhope, AL	*Gly. max*	MT820816	MT820854	MT820886	MT822668
23	FHP63[Z]	Cas2	Fairhope, AL	*Gly. max*	MT820817	MT820855	MT820887	MT822669
24	MAC06[Z]	Cas2+6	Macon County, AL	*Gly. max*	MT820819	MT820863	MT820895	MT822677
25	ELM19[Z]	Cas2	Elmore County, AL	*Gly. max*	MT820813	MT820849	MT820881	MT822663
26	AUR15[Z]	Cas2+6	Auburn, AL	*Gly. max*	MT820808	MT820840	MT820872	MT822654
27	FHP01[Z]	Cas0	Fairhope, AL	*Gos. hirsutum*	MT820825	MT820851	MT820883	MT822665
28	FHP22[Z]	Cas2	Fairhope, AL	*Gos. hirsutum*	MT820826	MT820852	MT820884	MT822666
29	BRW03[Z]	Cas2	Brewton, AL	*Gos. hirsutum*	MT820823	MT820841	MT820873	MT822655

30	HSV01[Z]	Cas2	Huntsville, AL	*Gos. hirsutum*	MT820827	MT820856	MT820888	MT822670
31	HSV12[Z]	Cas0	Huntsville, AL	*Gos. hirsutum*	MT820828	MT820857	MT820889	MT822671
32	TEN01[Z]	Cas2+6	Brewton, AL	*Gos. hirsutum*	MT820820	MT820864	MT820896	MT822678
33	BRW35[Z]	Cas2	Fairhope, AL	*Gos. hirsutum*	MT820810	MT820843	MT820875	MT822657
34	BRW47[Z]	Cas0	Brewton, AL	*Gos. hirsutum*	MT820811	MT820844	MT820876	MT822658
35	BRW50[Z]	Cas2	Brewton, AL	*Gos. hirsutum*	MT820812	MT820845	MT820877	MT822659
36	HSV21[Z]	Cas2	Huntsville, AL	*Gos. hirsutum*	MT820818	MT820858	MT820890	MT822672
37	AUR10[Z]	Cas2	Auburn, AL	*Gos. hirsutum*	MT820807	MT820839	MT820871	MT822653

[Z] The reference sequences were collected from NCBI

Table 2.3 Plant introduction numbers, year of release (YOR), and mean target spot severity ratings of sixteen soybean varieties inoculated with *Corynespora cassiicola* isolate LIM01.

Source of variance			***P* value**
Trial			0.75
Variety			<0.0001
Variety × Trial			0.99
Variety	**Accession** [x]	**YOR**	**Lsmeans** [y, z]
Henderson	PI 665225	2012	64.44 a
Pembina	PI 638510	2005	57.73 ab
McNair 600	PI 556480	1974	56.87 b
Sumter	PI 556722	1983	55.55 b
A7372	PI 556686	1983	51.22 bc
Coker 338	PI 556515	1976	49.88 c
McNair 800	PI 556481	1974	48.55 c
Davis	PI 553039	1967	48.10 cd
RA502	PI 556704	1983	43.11 de
FFR560	PI 556711	1983	41.88 ef
RA 451	PI 556759	1984	40.55 eg
RA606	PI 556703	1983	36.66 eh
Shiloh	PI 556664	1983	36.22 gh
RA580	PI 556700	1983	33.11 h
Bedford	PI 548974	1991	20.00 i
Council	PI 587091	1995	6.99 j

[z] Means followed by the same letter are not significantly different according to Fisher's least significant differences test (P <0.05).

[y] Lsmeans are from three trials and five replications within each trial (two-way interaction) of genotype, 14 days after inoculation.

[x] Germplasm resources information network (GRIN) database provided plant accession numbers used in this study.

Table 2.4 Mean target spot severity ratings with three isolates on eight soybean varieties.

Source of variance			***P* value**	
Variety			< 0.0001	
Isolate			0.06	
Variety × Isolate			0.50	
		Isolates [y]		
Variety [x]	**PBU11 (cas2+6)**	**SAR09 (cas2)**	**SSTA (cas2+6)**	**Across Isolates** [z]
Henderson	62.50	59.16	55.00	58.88 a-d
Pembina	52.50	53.33	50.83	52.22 c-g
Sumter	59.16	58.33	60.83	59.44 a-c
McNair 600	57.50	51.66	54.16	54.44 a-d
McNair 800	46.66	49.16	46.66	47.49 e
Shiloh	38.33	36.66	35.83	36.94 f
Bedford	20.83	17.50	14.16	17.49 g-h
Council	7.50	5.83	5.00	6.11 i
Across Variety [z]	43.12 a	41.45 ab	40.31 b	

[z] Means followed by the same letter are not significantly different according to Fisher's least significant differences test ($P < 0.05$).

Chapter 3. Comparative transcriptome profiling unfolds a complex defense and secondary metabolite networks imparting *Corynespora cassiicola* resistance in soybean

Abstract

Target spot caused by *Corynespora cassiicola* is an overwhelming disease in soybean production areas that are hot and humid. Resistant soybean genotypes have been identified however, the molecular mechanisms governing resistance to infection are unknown. A comparative transcriptomic profiling using two known resistant genotypes and two susceptible genotypes was performed under infected and controlled conditions to understand the regulatory network operating between soybean and *C. cassiicola*. RNA-Seq analysis identified a total of 2,571 differentially expressed genes (DEGs) which were shared by all four genotypes. These DEGs are related to secondary metabolites, immune response, defense response, phenylpropanoid, and flavonoid/isoflavonoid pathways in all four genotypes after *C. cassiicola* infection. In the two resistant genotypes, additional upregulated DEGs were identified affiliated with the defense network: flavonoids, jasmonic, salicylic, and brassinosteroids acid. Further analysis led to the identification of differentially expressed transcription factors, immune receptor, and defense genes with a leucine-rich repeat domain, dirigent proteins, and Cysteine (C)-rich receptor-like kinases. These results will provide insight into molecular mechanisms of soybean resistance to *C. cassiicola* infection and valuable resources to potentially pyramid quantitative resistance loci for improving soybean germplasm.

3.1 Introduction

Soybean (*Glycine max* (L.) Merrill) is widely consumed in many forms, contributing 59% of the world's edible oil production and suppling 31- 40% in high-quality protein, making it a good staple food source for human and animal consumption. In 2021, 128 million ha of soybean

was produced with yields of 364 million metric tons (USDA 2021). Soybean production typically does not achieve its full yield potential due to biotic stresses (HARTMAN *et al.* 2015). One disease, target spot, caused by the plant pathogenic fungus, *Corynespora cassiicola*, threatens the Southern US soybean industry due to the fungus' affinity for warm temperatures and high moisture conditions with an estimated 18-32% in yield losses (GODOY 2015; FASKE 2016). Symptoms of infection by *C. cassiicola* are necrotic spots with alternating concentric rings of light and dark brown bands, usually encircled by a yellow halo on the foliage. In addition, lesions on stems, pods, and seeds, and premature leaf senescence could occur in severe cases (KOENNING *et al.* 2006; GODOY 2015). The present method for controlling target spot relies on fungicide applications. The immense use of fungicide leads to the development of resistance in *C. cassiicola* isolates and reduces fungicide effectiveness (XAVIER *et al.* 2013; DUAN *et al.* 2019; DE MELLO *et al.* 2021). Another eco-friendlier and sustainable approach is to develop a resistant cultivar.

Corynespora cassiicola resistance exists in soybean genotypes (PATEL *et al.* 2022); however, the genes and mechanism of resistance are unknown. Investigating disease resistance mechanisms relies on the identification of genomic regions, genes, and gene networks associated with defense responses triggered by the host upon infection with pathogens (AMARAL *et al.* 2008). RNA sequencing (RNA-Seq) is a well-advanced and effective technology used for studying gene expression at the whole genome level which can detect novel transcripts, splice junctions, facilitates DEG analysis, and allow functional gene mining (WANG *et al.* 2009; VAN VERK *et al.* 2013). The enhancement of high-throughput sequencing (HTS) technology and the availability of comprehensive soybean genome sequences have allowed the full-scale examination of transcriptomic response to disease (SCHMUTZ *et al.* 2010). Soybean RNA-Seq

studies have provided an opportunity to gain in-depth knowledge of plant-pathogen interactions by identifying responsive genes and pathways for disease resistance such as to soybean cyst nematode (MIRAEIZ *et al.* 2020; KOFSKY *et al.* 2021), Phytophthora root and stem rot (LIN *et al.* 2014), Fusarium root rot (LANUBILE *et al.* 2015), bacterial leaf pustule (KIM *et al.* 2011), *Soybean mosaic virus* (DEMERS *et al.* 2020), downy mildew (DONG *et al.* 2018), and brown stem rot (MCCABE *et al.* 2018).

Corynespora cassiicola is a devastating pathogen in crops such as rubber tree (*Hevea brasiliensis*), cotton (*Gossypium hirsutum* L.), and cucumber (*Cucumis sativus L.*) (BLAZQUEZ 1967; CHEE 1987; CONNER et al. 2013) RNA-Seq studies in cucumber and rubber, after infection by C. cassiicola, have found differential gene expression associated with Ca2+ signaling pathways, pathways targeting salicylic acid (SA), ethylene (ET), and phenylpropanoid biosynthesis (LIU et al. 2017; WANG et al. 2018; ROY et al. 2019; RIBEIRO et al. 2021). Such studies have aided in identifying miRNAs, genes, and gene variations that are critical for understanding the genetic basis and marker development for disease resistance. In soybean, histochemical characterization and biochemical assays *in* control and infected tissues of the plant suggest a major role of total soluble phenolics (TSP) and lignin-thioglycolic acid (LTGA) derivatives in controlling the spread of *C. cassiicola* (FORTUNATO et al. 2015; FORTUNATO et al. 2017). However, to date, there has been no transcriptome studies for *C. cassiicola* interaction in soybean to validate such claims. Two genotypes, Bedford and Council, have been identified as having low disease severity when inoculated with *C. cassiicola* and are considered resistant, while two genotypes, Henderson and Pembina, are documented as susceptible (PATEL *et al.* 2022) Therefore the objective of this study is to conduct a comparative transcriptomic analysis to determine the differentially expressed genes (DEGs) between non-inoculated control and post *C.*

cassiicola infection at 24- and 48-hours post-infection. These DEGs will be further evaluated for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. This will provide insight into genes and molecular mechanisms that underlie soybean resistance against *C. cassiicola* infection.

3.2 Materials and methods

3.2.1 Plant materials and inoculation

Two soybean resistant genotypes 'Council' (PI 587091) and 'Bedford' (PI 548974), and two susceptible 'Henderson' (PI 665225), and 'Pembina' (PI 638510) were sown in the greenhouse in twelve separate 11.5 × 11.5 cm^2 pots filled with PRO-MIX BX (Premier Tech Horticulture, Quakertown, PA). Plants were watered as needed under a photoperiod of 14/10 (light/dark); and a thermocycle of 24°C /12°C for 25 days (growth stage V3-V4) before fungal infection. A LIM01 *C. cassiicola* isolate was grown on V8 agar plates for twelve days at 28°C with 12/12 h of light/dark in an incubator and then conidia scraped into sterile distilled water and filtered with cheesecloth to make conidial suspension with a concentration of 50,000 per ml. (PATEL *et al.* 2022). Plants were sprayed with 0.05% of a Tween 20 solution and allowed to dry for 5 minutes before inoculation. A conidial suspension of was sprayed onto axial and abaxial leaf sides using a fine mist professional spray bottle (Spray Pro) until run-off. Afterward, inoculated plants were transferred into a plastic mist chamber inside the greenhouse with a non-inoculated control of each genotype (sprayed with distilled water). Plants were arranged in a completely randomized design (CRD) in the mist chamber; mist ran for 2 s every 10 min for three days to maintain high humidity (PATEL *et al.* 2022).

3.2.2 RNA extraction, library preparation, and Illumina sequencing

Three biological replicates from all four genotypes at 24- and 48-hours post-inoculation (hpi), and their respective controls, were selected for RNA isolation. RNA was extracted from 100-150 mg tissue samples using the Direct-zol RNA Mini-Prep Kit (Zymo Research, Irvine, CA) and concentration estimated using a Nanodrop 2000 spectrophotometer (ThermoFisher Sci., Waltham, MA). The RNA degradation was evaluated on an agarose gel electrophoresis and intact RNA with high concentrations was sent for RNA-Sequencing (RNA-Seq). Novo gene Bioinformatics Technology Co., Ltd., CA, US, performed cDNA library construction and 150-bp paired-end sequencing run on an Illumina NovaSeq 6000 instrument. Data was released after a read quality check for a percentage of reads containing N >10% (N represents the base that cannot be determined) and low-quality reads (Q score ≤ 5).

3.2.3 RNA-Seq data analysis

An additional quality check of raw reads was performed using fastp software v0.23.1 (CHEN *et al.* 2018) to remove adapters, Poly A sequences, low-quality reads (Q < 30), and reads < 15bp in length after trimming. These cleaned reads were mapped to the soybean (*Glycine max* (L.) Merr.) reference genome (SCHMUTZ *et al.* 2010), and transcript quantification was performed using Salmon ver.0.9.1 (PATRO *et al.* 2017). A total of eight comparisons, which include 24 hpi vs control and 48 hpi vs control for all the four genotypes, were performed to identify differentially expressed genes (DEGs) using R-package DESeq2 (Version 1.37.6) (LOVE *et al.* 2014). The genes with an absolute value of log2 fold change ≥ 2 (upregulated genes) or ≤ -2 (downregulated genes) and a false discovery rate (FDR) < 0.01 were considered significant DEGs.

3.2.4 GO Enrichment and KEGG pathway analysis

Gene ontology functional enrichment (GO) analysis for the DEGs was conducted using ShinyGo v 0.76 (GE *et al.* 2020) with *Glycine max* as background. Gene ontology provided annotations at cellular, molecular, and biological levels. The function categories of enriched GO terms were considered significant with FDR -adjusted $P \leq 0.05$.

KEGG pathway analysis of DEGs was carried out using KOBAS v 3.0 software (BU *et al.* 2021) with *Glycine max* sequences as background, a hypergeometric test, and Bonferroni FDR correction (FDR $P \leq 0.05$). This analysis tests the statistical enrichment of DEGs in KEGG pathways.

3.2.5 Identification of transcription factors (TFs)

PlantTFDB (http://planttfdb.gao-lab.org/) was used to identify transcription factors involved in regulating soybean response after *C. cassiicola* infection, containing 58 plant transcription factor (TF) families from 165 plant (JIN *et al.* 2016). The analysis was conducted using the TF enrichment tool and soybean (*G. max*) transcription factor database, which contains 6,150 TFs (3,747 loci) distributed into 57 families. TFs were searched in differentially expressed genes at 24 hpi or 48 hpi for each genotype.

3.2.6 Quantitative real-time PCR validation

Quantitative real-time PCR (qRT-PCR) was performed for the relative expression of selected eight DEGs to validate RNA-Seq data. All the randomly selected eight gene sequences were retrieved from Soybase (https://soybase.org/). RNA extraction was conducted as described above, and cDNA synthesis was carried out by qScript™ cDNA Synthesis Kit (New England Biolabs, Inc., Ipswich, MA, USA). Primers were designed using Primer-BLAST tool (YE *et al.* 2012) and listed in Appendix 1E. The qRT-PCR was performed in 96-well plates on Thermo Fisher Scientific Biosystems StepOnePlus™ Real-Time PCR system (Applied Biosystem, MA,

USA) using PerfeCTa SYBR Green ROX FastMix (Quantabio). The following conditions were used for amplification, 2 min at 95ºC; followed by 40 cycles of 5 sec at 95ºC and 10 sec at 58ºC plus melting curves to verify PCR products. The gene expression level of selected genes was calculated with the $2^{-\Delta\Delta CT}$ method (LIVAK AND SCHMITTGEN 2001), and Ubiquitin-conjugating protein endogenous control was used to normalize the variance among samples.

3.3 Results

3.3.1 RNA Sequencing data analysis and DEG in response to *C. cassiicola*

For RNA sequencing, raw reads ranged from 22,838,228 to 33,485,280 per sample, with an average GC content of 44-45%. After quality check, adapter and low-quality reads were discarded from the data. The clean data were ranged from 22,381,404 to 33,067,807 reads per sample (Table 3.1). The 88-91% sequence reads were successfully mapped to the soybean reference genome. A total of 11,263 DEGs were identified at 24 hpi, including 1,387 DEGs which were common in all four genotypes. Moreover, a total of 11,094 DEGs were determined at 48 hpi, including 1,184 DEGs which were common in all four genotypes (Figure 3.1).

When compared to the non-inoculated control, the resistant genotypes at 24hpi had 2,883 and 2,042 upregulated, and 4,565 and 1,410 downregulated DEGs in Bedford and Council, respectively, (Figure 3.1 A). At 48hr, 2,076 upregulated and 1,340 downregulated DEGs were noted in Bedford and 3,055 upregulated and 2,883 downregulated in Council. For the susceptible genotypes, at 24 hpi, Henderson had 4,710 DEGs (2,063 upregulated and 2,647 downregulated) while Pembina had 6,401 (2,757 upregulated and 3,544 downregulated). At 48 hpi, there were 7,042 (3,165 upregulated and 3,877 downregulated) DEGs in Henderson and 4,756 (2,198 upregulated and 2,558 downregulated) DEGs in Pembina (Figure 3.1 A).

Genes that were uniquely differentially expressed in resistant or susceptible genotypes after *C. cassiicola* infection were 5,172 at 24hpi and 5,384 at 48hpi. In the resistant genotypes, 3,403 and 2,760 genes were differentially expressed at 24 hpi and 48 hpi, respectively, but not in susceptible genotypes. Unique DEGs in Council were 895 at 24 hpi and 1,799 at 48 hpi; for Bedford, there were 2,353 at 24 hpi, and 798 at 48 hpi. A total of 155 and 163 DEGs were shared in both resistant genotypes at 24 hpi and 48 hpi, respectively (Figure 3.1 B,1C).

A total of 334 and 693 DEGs were shared between susceptible genotypes at 24 hpi and 48 hpi, respectively. The unique DEGs in Henderson were 578 (243 upregulated and 335 downregulated) at 24 hpi and 2,237 (1070 upregulated and 1167 downregulated) at 48 hpi; in Pembina there were 1,346 (610 upregulated and 736 downregulated) DEGs at 24 hpi and 550 (237 upregulated and 313 downregulated) DEGs at 48 hpi (Figure 3.1 B,1C).

3.3.2 Gene ontology enrichment analysis of DEGs

The results of a GO enrichment for the DEGs common in all four genotypes after *C. cassiicola* infection, regardless of time after inoculation, revealed 150 GO terms and 139 GO terms for upregulated and downregulated DEGs, respectively. The enriched GO terms in upregulated biological processes were secondary metabolic process, RNA modification, glucosinolate metabolic process, cell recognition, response to biotic stimulus, flavonoid biosynthetic process, regulation of protein serine/threonine phosphatase activity, defense response, and immune response indicating the importance of activation of a defense-related network in response to *C. cassiicola* infection. In addition, a cluster of GO terms related to oxidoreductase activity, monooxygenase activity, peroxidase activity, protein serine kinase activity, naringenin-chalcone synthase activity, protein serine/threonine kinase activity, hormone binding, glucosidase activity, and phenylalanine ammonia-lyase activity were also observed in

the molecular function category (Figure 3.2, Supplementary Table Appendix 1A). Thus, upregulated DEGs were assigned to GO terms mainly associated with defending the soybean plant in response to *C. cassiicola* infection.

The GO analysis (based on uniquely identified DEGs for each genotype) were performed. In Bedford, unique upregulated DEGs have enriched biological processes GO terms were: gene functions involved in biotic stimulus, defense response to other organisms, immune system process, immune response, defense response, innate immune response, plant type secondary cell wall biogenesis, response to external biotic stimulus, cell wall organization and biogenesis (Figure 3.3 A, Supplementary Table Appendix 1B). In Council, unique upregulated DEGs have enriched biological processes GO terms involved in defense response, chitin metabolic, chitin catabolic process, hormone-mediated signaling pathway, response to auxin, response to biotic stimulus, induced systemic resistance, activation of an innate immune response, and auxin activated signaling pathway. Protein serine/threonine kinase activity, chitin-binding, chitinase activity, and oxidoreductase activity were upregulated in molecular functions after infection compared to susceptible genotypes. Moreover, a cluster of GO terms related to meiotic nuclear division, meiotic cell cycle process, chromatin assembly, nucleosome assembly, chromosome segregation, chromosome organization, and nuclear division were downregulated after *C. cassiicola* infection in Council, which indicates that the cell division processes might be affected (Figure 3.3 B, Supplementary Table Appendix 1B).

In susceptible genotypes, the enriched GO terms assigned to unique DEGs were not involved directly with the defense network. The most enriched upregulated biological process GO terms in Henderson were rRNA processing, RNA metabolic process, and ribosome biogenesis. Likewise, in Pembina, RNA modification, phenylpropanoid metabolic process, and lignin metabolic

process were the most enriched upregulated GO terms in biological process (Supplementary Table Appendix 1B).

3.3.3 KEGG Pathway analysis

The KEGG enrichment analysis utilized the DEGs shared by all four genotypes at 24 hpi or 48 hpi against the *Glycine max* gene background were assigned to 98 pathways; 18 pathways were significantly enriched (FDR - adjusted $p \leq 0.05$) (Bu *et al.* 2021).The metabolic pathways (gmx01100), biosynthesis of secondary metabolites (gmx01110), phenylpropanoid biosynthesis (gmx00940), plant hormone signal transduction (gmx04075), starch and sucrose metabolism (gmx00500), MAPK signaling pathway (gmx04016), circadian rhythm (gmx04712), and flavonoid biosynthesis (gmx00941) were prominent pathways in response to *C. cassiicola* at 24 and 48 hpi across all genotypes. In addition, isoflavonoid biosynthesis, cutin, suberine and wax biosynthesis, fatty acid elongation, glycerolipid metabolism, ascorbate and aldarate metabolism, cyanoamino acid metabolism, thiamine metabolism, carotenoid biosynthesis, nitrogen metabolism, ubiquinone, and other terpenoid-quinone biosynthesis were also enriched (Table 3.2).

Closer investigation of the phenylpropanoid pathway revealed that many DEGs are involved in the biogenesis of various phenolic and lignin compounds. Figure 3.4 illustrates the activation of various phenolic polymers and lignin compounds in the phenylpropanoid pathway. The upregulated DEGs were phenylalanine ammonia-lyase (EC:4.3.1.24), caffeate O-methyltransferase (EC 2.1.1.68), coniferyl-aldehyde dehydrogenase (EC:1.2.1.68), shikimate O-hydroxycinnamoyltransferase (EC:2.3.1.133), coniferyl-alcohol glucosyltransferase (EC:2.4.1.111), 1-cys peroxiredoxin (EC:1.11.1.7). The above results indicates that these pathways may perform to increase soybean immunity against *C. cassiicola* (Figure 3.4).

Several upregulated DEGs were assigned to biosynthesis of flavonoids/isoflavonoids pathway; IFS1 (isoflavone synthase 1 precursor), IOMT1 (isoflavone 4'-O-methyltransferase), isoflavone 7-O-methyltransferase (EC:2.1.1.150), isoflavone 7-O-glucoside-6"-O-malonyltransferase (EC:2.3.1.115), isoflavone/4'-methoxyisoflavone 2'-hydroxylase (EC:1.14.14.90; 1.14.14.89), and CYP93A1 (3,9-dihydroxypterocarpan 6A-monooxygenase). Similarly in flavonoid biosynthesis pathway, upregulated DEGs were CHS8 (chalcone synthase), shikimate O-hydroxycinnamoyltransferase (EC:2.3.1.133), caffeoyl-CoA O-methyltransferase (EC:2.1.1.104), flavonoid 3'-monooxygenase (EC:1.14.14.82), flavonoid 4'-O-methyltransferase (EC:2.1.1.231), and flavanone 4-reductase (EC:1.1.1.219; 1.1.1.234) (Supplementary Figure Appendix 1C and D).

3.3.4 Identification of differentially expressed TFs

For transcription factors (TF), 574 DEGs were identified to 40 different TF families across all four genotypes. The highest represented TF families are 80 DEGs in ethylene responsive factor (ERFs), 74 MYB, 51 WRKY, 62 bHLH, and 38 NAC. Our study found that WRKY TFs were predominantly upregulated after infection, suggesting this particular group might have a critical role in response to *C. cassiicola* (Figure 3.5). Further, a total of 16 DEGs belonging to WRKY family were upregulated in all four genotypes after *C. cassiicola* infection (Figure 3.5).

In Council, 45 DEGs belonging to 22 TFs families were observed with upregulated TFs distributed between C2H2 family (3 TFs), WRKY (1 TFs), NAC (1 TFs), MYB (2 TFs), and MYB-related family (1 TFs). In Bedford, 66 genes belonging to 24 TFs families were differentially expressed after *C. cassiicola* infection. Among them, upregulated TFs were distributed in NAC (9 TFs), WRKY (4 TFs), bHLH (4 TFs), ERF (3 TFs), MYB (3 TFs), and MYB-related family (2 TFs). These upregulated TFs may enhance the immunity of the resistant

genotypes against *C. cassiicola* infection. Interestingly, there was no expression change of these TFs in susceptible genotypes after *C. cassiicola* infection.

3.3.5 Quantitative real-time expression analysis

Comparing log2 fold change from RNA-Seq analysis for eight differentially expressed genes with qPCR results revealed a positive correlation (r = 0.81) and had consistent expression trends (Figure 3.6). These results suggested the reliability of RNA-seq in analyzing the transcriptome of resistant and susceptible plants after *C. cassiicola* infection. In short, qRT-PCR endorsed the expression pattern of DEGs exhibited by RNA-Seq.

3.4. Discussion

Target spot caused by *C. cassiicola* is an emerging disease in regions with warm and humid climates. This RNA-Seq study involving two resistant genotypes (Bedford and Council) and two susceptible genotypes (Pembina and Henderson) sheds light on possible soybean responses to *C. cassiicola* infection and potential resistance mechanisms. The Illumina sequencing for 36 RNA-Seq libraries generating 94 GB of data was used for further analysis. An average of 25.72 million clean reads per library was obtained, of which 88.76% were mapped to the soybean genome. This indicates that our data is sufficient to conduct DEG analysis to identify defense related genes and pathways against *C. cassiicola*.

During their life cycle, soybean plants are attacked by various pathogens (fungi, nematodes, bacteria, and viruses), which impact plant growth and development, ultimately reducing yield (LEE *et al.* 2017). Plants respond to pathogen attack by the changing expression of genes which alters different pathways (DODDS AND RATHJEN 2010). The functional analysis of common DEGs in all four genotypes indicated that several genes belonging to biological processes, cellular components, and molecular function were influenced by *C. cassiicola* infection.

Enriched GO terms in all four genotypes in response to *C. cassiicola* infection were: defense response, response to biotic stimulus, cutin biosynthetic process, protein serine/threonine kinase, oxidoreductase activity, and peroxidase activity. These pathways have been highlighted in several plant-pathogen interaction transcriptomic studies such as *Sclerotium rolfsii* in peanut (BOSAMIA *et al.* 2020), and *Xanthomonas sp.* in pepper (GAO *et al.* 2021) and tomato (DU *et al.* 2015). Moreover, upregulated genes only in 'Bedford' and 'Council' after *C. cassiicola* infection were assigned to defense response, response to biotic stimulus, and immune response. Additionally, upregulated genes found only in Council exhibit activation of chitin-binding/catabolic, chitinase activity, salicylic acid biosynthetic process, hormone-mediated signaling pathway, and protein serine/threonine kinase activity. These unique upregulated DEGs found in the resistant genotypes after infection might contribute to resistance to *C. cassiicola*. Transcriptome studies involving *C. cassiicola* infection of rubber found similar activation of defense response and chitinase activity only in tolerant clone (ROY *et al.* 2019; RIBEIRO *et al.* 2021).

Fungal invasion of plants triggers two layers of the immune defense mechanisms. Transmembrane proteins such as receptor kinases (RLKs) and receptor-like proteins (RLPs) act as pattern-recognition receptors (PRRs), known as host sensors, allow plants to recognize microbial pathogens surrounding them through pathogen-associated molecular patterns (PAMPs)- triggered immunity (PTI) (TANG *et al.* 2017). This PTI is the first line of defense that restricts pathogen invasion. Effector-triggered immunity (ETI) is the second layer of defense against pathogen attack. This system recognizes the effector proteins that cause local cell death, often called the hypersensitive reaction (HR) in plants (DODDS AND RATHJEN 2010). A leucine-rich repeat (LRR) domain is generally present in immune receptors and defense genes of plants.

Several genes with the LRR domain were upregulated in all four genotypes after *C. cassiicola* infection. Specifically, 73 and 33 LRR genes were upregulated in Council and Bedford, while no expression difference was found in susceptible genotypes. These genes found in Council and Bedford were in the disease resistance protein (TIR-NBS-LRR class) family (10, 6), Leucine-rich receptor-like protein kinase family (10, 2), LRR family protein (22, 14), LRR protein kinase family (19, 8), LRR transmembrane protein kinase (10, 3), LRR and NB-ARC domains-containing disease resistance protein (1, 0) and MLLR family (1, 0), respectively.

Most commonly known disease resistant genes (R genes) contain a nucleotide-binding site (NSB) and LRR protein which helps to identify specialized pathogen- associated proteins (DEYOUNG AND INNES 2006). These genes can be further classified into proteins with north terminal toll and interleukin 1 receptors (TIR) domain, coiled-coil (CC) domain, and without any N domain (MCHALE *et al.* 2006). Interestingly, all upregulated NBS-LLR genes in Council and Bedford had only a TIR domain. Two different TIR-NB-LLRs were identified in *Arabidopsis* providing tolerance to *Leptosphaeria maculans* fungus (STAAL *et al.* 2006). Another gene, *RLM3,* encoding TIR-NB class, was found to provide immunity to different necrotrophic fungal pathogens in *Arabidopsis* (STAAL *et al.* 2008). Further investigation is needed to understand the role of TIR-NBS-LLR genes in soybean defense response to *C. cassiicola.*

Receptor kinases send downstream signals for appropriate cellular response to biotic and abiotic stress. Some receptor kinases contain cysteine-rich proteins, known as cysteine-rich receptor-like kinases (CRKs). Genes in the CRK family play important roles in disease resistance by interacting with PAMP and sending defense signaling for an HR-like cell death (CHEN *et al.* 2003; BOURDAIS *et al.* 2015; YADETA *et al.* 2017). A total of 46 and 42 CRKs genes were upregulated in Council and Bedford, respectfully, after *C. cassiicola* infection. Interestingly, two

copies of CRK 25 gene (Glyma.20g137400 and Glyma.20g139300) were highly upregulated in Bedford and a copy of CRK 4 (Glyma.20g118400) was upregulated in both resistant genotypes but had very low expression at 24 hpi in both susceptible genotypes (Figure 3.7 A-B). CRK 4 is an important receptor protein kinase identified to have a critical role in early PTI response by triggering HR (CHEN *et al.* 2003; BOURDAIS *et al.* 2015; YADETA *et al.* 2017). This suggests that early activation of CRK genes in resistance genotypes might reduce the spread of *C. cassiicola*.

Chitin is a cell wall component of fungi that is not present in the plant cell wall (PUSZTAHELYI 2018). Chitinase possesses antifungal properties, restricting the growth of many fungal pathogens like *Trichosanthes dioica*, *Aspergillus niger*, *Alternaria solani*, *Fusarium* sp., *Rhizoctonia solani*, and *Verticillium dahlia* (SCHLUMBAUM *et al.* 1986; KABIR *et al.* 2016; KUMAR *et al.* 2018; TOUFIQ *et al.* 2018). Council showed upregulation of additional genes related to chitin binding and chitinase activity after infection. Additionally, Chitinase A (Glyma.19G076200) was highly expressed in Council when compared to susceptible genotypes at 24hpi and 48 hpi (Figure 3.7 B). Tobacco plants overexpressing Chitinase A from *Autographa californica* nuclear polyhedrosis virus showed resistance to fungal pathogens (CORRADO *et al.* 2008). Thus, higher expression of Chitinase A might reduce colonization of *C. cassiicola*.

Several other defense-related genes were upregulated in resistant genotypes with significantly higher expression than susceptible genotypes at 24 and 48 hpi. In Council these genes belong to disease resistance-responsive (dirigent-like protein) (Glyma.03g147600), B-box zinc finger (Glyma.04g009200), mitogen-activated protein kinase (Glyma.12g097200), cysteine-rich secretory protein (Glyma.16g143300, *NPR1*), NB-ARC domain-containing disease resistance protein (Glyma.18g084400, *RPM1*) and Glyma.18g087000, *RPM4*), receptor serine/threonine kinase (Glyma.13g033100), and serine protease inhibitor (Glyma.20g205700) (Figure 3.7 B).

The *RPM1* gene provides resistance in *A. thaliana* against *P. syringae* (REUBER AND AUSUBEL 1996) and in wheat against *Puccinia striiformis* (WANG *et al.* 2020). *RPM* gene families have been identified as major players in the soybean defense mechanism against pathogens and to stress response (WHALEY *et al.* 2015; AFZAL *et al.* 2022). Pathogenesis-related proteins (PR) in plants participate in innate immune system defense response against pathogens. Several studies have found that the *NPR1* gene provides resistance to different species of fungus in cotton (PARKHI *et al.* 2010), *Arabidopsis* (STEIN *et al.* 2008), and *Brassica juncea* (ALI *et al.* 2017). In Bedford, these genes belong to wall-associated kinase family protein (Glyma.09g027500), mitogen-activated protein kinase (Glyma.18g060900, *MAPK*), *kunitz trypsin inhibitor 1* (Glyma.08g342000, *KTI1*), pentatricopeptide repeat (PPR) superfamily protein (Glyma.08g233900), scorpion toxin-like knottin superfamily protein (Glyma.06g160300), cytochrome P450 (Glyma.07g118200 and Glyma.05g042500), receptor serine/threonine kinase (Glyma.13g033100 and Glyma.13g033800), and peroxidase superfamily protein (Glyma.02g233800) (Figure 3.7 A). Transcriptomic study revealed expression of some DEGs associated with MAPK cascades in response to *Xanthomonas oryzae* infection in resistant rice genotype (YANG *et al.* 2015). Moreover, overexpression of KTI1 in tobacco enhances resistance against *Rhizoctonia solani* infection (HUANG *et al.* 2010). There is a need to understand the impact of these upregulated genes towards target spot resistance.

In this study, *C. cassiicola* infection was associated with diverse plant defense response TFs families, such as ERFs, WRKY, MYB, and bHLH. For the WRKY TF family, WRKY29 (Glyma.08G018300) in Council, and three WRKY6 (Glyma.08G320200, Glyma.18G092200), WRKY7(Glyma.17G239200), and WRKY41(Glyma.19G254800) in Bedford were upregulated after *C. cassiicola* infection while in susceptible genotypes there was no expression difference

between control and infected tissue. Previous studies show that expression of WRKY29 and WRKY41 TFs increased Fusarium head blight resistance in wheat (SAROWAR *et al.* 2019) and *Pseudomonas* resistance in Arabidopsis (HIGASHI *et al.* 2008).

Phenylpropanoid biosynthesis plays an essential role in plant stress response. This pathway leads to the biogenesis of various phenolic polymers, lignin compounds, and flavonoids, increasing plant immunity (DICKO *et al.* 2005; LOZOVAYA *et al.* 2007). The gene coding for phenylalanine ammonia-lyase (PAL), caffeate O-methyltransferase (CCoAOMT), coniferyl-aldehyde dehydrogenase, 1-cys peroxiredoxin, shikimate O-hydroxycinnamoyltransferase, and coniferyl-alcohol glucosyltransferase was found upregulated in all four genotypes. This indicates that this pathway might be stimulated early in all genotypes as part of the PTI response. Similar results were observed in the research conducted on rubber tree clones in response to *C. cassiicola* infection (RIBEIRO *et al.* 2021). Moreover, these gene coding enzymes form syringyl and guaiacyl, units of lignin polymers, which are major building blocks of lignin and end products of lignin biosynthesis. Phenylpropanoid polymer lignin acts as a physical barrier against pathogen invasion (MITCHELL *et al.* 1999). The phenylpropanoid pathway is regulated in stress conditions and associated with the lignification process in *Arabidopsis* and *Populous* (HAMBERGER *et al.* 2007). Similarly, lignin formation is an essential process for the defense of host plants under both abiotic and biotic stresses (VOGT 2010; MIEDES *et al.* 2014).

Flavonoid biosynthesis is an essential downstream branch of phenylpropanoid metabolism. The gene expression of chalcone synthase, the key enzyme in the flavonoid pathway, is induced in plants with fungal or bacterial infection (DAO *et al.* 2011). Similarly, isoflavonoids are a mainly legume-specific subclass of flavonoid metabolites with significant roles in plant defense (ALGAR *et al.* 2014). In the enriched isoflavonoid pathway, three cytochrome families were

involved in the biosynthesis process, CYP93C (cytochrome P45093C), CYP81E1/E7, (isoflavone/4'-methoxyisoflavone 2'-hydroxylase) and CYP93A1 (3,9-dihydroxypterocarpan 6A-monooxygenase). These pathways were activated in both resistant and susceptible genotypes, suggesting a common defense pathway activated in soybean plants when attacked by *C. cassiicola*. Additionally, few genes in the flavonoid biosynthesis pathway were expressed higher in Council and/or Bedford (Figure 3.8). In Bedford, two genes Glyma.09g038900 (*MYB111*) and Glyma.06g260200 (NAD(P)-linked oxidoreductase) involved in flavonoid biosynthesis were upregulated and more highly expressed than in the susceptible genotypes after infection. Similarly, in Council, four genes involved in flavonoid biosynthesis had higher expression than in susceptible genotypes after infection: Glyma.01g006800 (Pectin lyase-like superfamily protein), Glyma.02g013900 (MYB domain protein 12), Glyma.02g048400 (flavanone 3-hydroxylase), and Glyma.02g048600 (flavanone 3-hydroxylase). MYB transcription factors such as *MYB12* and *MYB111* modulate the production flavonoid pathway by regulating early biosynthesis genes such as chalcone synthase (*CHS*), chalcone isomerase (*CHI*), flavanone 3-hydroxylase (*F3H*), and flavonol synthase1 (*FLS1*) during normal development stages and in stress conditions (MEHRTENS *et al.* 2005; STRACKE *et al.* 2010; NAKABAYASHI *et al.* 2014; LI *et al.* 2019). Two different copies of flavanone 3-hydroxylase (*F3H*) were more highly expressed in Council compared to susceptible genotypes at 24 hpi and 48 hpi. Flavanone 3-hydroxylase (*F3H*) is responsible for producing different flavonoid compounds. Studies have associated upregulation of genes with tolerance/resistance to pathogens such as *Alternaria solani, Ascochyta rabiei (Pass) Labr., Xanthomonas Oryzae pv. oryzae* (CHO *et al.* 2005; MAHAJAN AND YADAV 2014; JAN *et al.* 2021). Thus, our study speculates a major role of genes involved in the flavonoid pathway in contributing to the resistance against *C. cassiicola*.

Jasmonic acid (JA), salicylic acid (SA), and brassinosteroids (BRs) are the plant defense hormones activated in the downstream process of PTI and ETI responses (VERHAGE *et al.* 2010; KOHLI *et al.* 2019). In this study, genes associated with biosynthesis of BRs, JA, and SA were upregulated in resistant genotypes compared to susceptible genotypes after *C. cassiicola* infection. Three and two genes related to BR biosynthesis were highly expressed after infection in Council and Bedford, respectively (Figure 3.8). Further, twelve genes for each SA and JA were upregulated at least in one resistant genotype while having lower expression in susceptible genotypes after infection. Such higher expression of genes involved in these defense-related hormones might play a vital role in conferring target spot resistance. Similar results were observed in other transcriptomic studies involving plant host-pathogen interactions (TARIQ *et al.* 2018; BOSAMIA *et al.* 2020; DASGUPTA *et al.* 2021).

Quantitative Disease Resistance (QDR) is a phenomenon when many genes with small effects are differentially expressed during the invasion of the pathogen, which results in a reduction of the fungal colonization. Pyramiding such small effect genes from the different resistant genotypes would be an effective strategy to develop an enduring disease-resistant variety (POLAND *et al.* 2009; FRENCH *et al.* 2016). Studies have demonstrated a higher level of disease resistance by pyramiding such QDRs (RICHARDSON *et al.* 2006; FUKUOKA *et al.* 2015). Further research needs to be conducted to understand the effects of QDRs and genetic gain for disease resistance by pyramiding these QDRs from Bedford and Council into a single germplasm. This would also help to develop germplasm with broad-spectrum resistance as generally QDR participates in defense mechanisms against wide range of microbial pathogen and multiple races (KRATTINGER *et al.* 2009; POLAND *et al.* 2009).

3.5 Conclusion

This study presents the first large-scale comparative transcriptomic profiling of resistant and susceptible soybean genotypes in response to the invasion of necrotrophic fungus *C. cassiicola*. Results have revealed a complex and massive gene network response, providing insight into mechanisms directing resistance to *C. cassiicola* in soybean. The analysis suggests that TIR-NBS-LRR, LRR, NB-ARC, CRKLs, and DIR genes play an essential role in understanding pathogen invasion through a downstream resistance mechanism. Furthermore, genes involved in flavonoid/isoflavonoid, phenylpropanoid, JA, SA, and BA are upregulated upon *C. cassiicola* attack, thereby inducing systemic resistance.

References

Afzal, M., S. S. Alghamdi, H. Nawaz, H. H. Migdadi, M. Altaf *et al.*, 2022 Genome-wide identification and expression analysis of CC-NB-ARC-LRR (NB-ARC) disease-resistant family members from soybean (*Glycine max L.*) reveal their response to biotic stress. Journal of King Saud University-Science 34**:** 101758.

Algar, E., F. J. Gutierrez-Mañero, A. Garcia-Villaraco, D. García-Seco, J. A. Lucas *et al.*, 2014 The role of isoflavone metabolism in plant protection depends on the rhizobacterial MAMP that triggers systemic resistance against *Xanthomonas axonopodis pv. glycines* in *Glycine max* (L.) Merr. cv. Osumi. Plant Physiology and Biochemistry 82**:** 9-16.

Ali, S., Z. A. Mir, A. Tyagi, H. Mehari, R. P. Meena *et al.*, 2017 Overexpression of NPR1 in *Brassica juncea* confers broad spectrum resistance to fungal pathogens. Frontiers In Plant Science 8**:** 1693.

Amaral, D. O. J. d., M. M. d. A. Lima, L. V. Resende and M. V. d. Silva, 2008 Differential gene expression, induced by salicylic acid and *Fusarium oxysporum f. sp. lycopersici* infection, in tomato. Pesquisa Agropecuária Brasileira 43**:** 1017-1023.

Blazquez, C., 1967 *Corynespora* leaf spot of cucumber, pp. 177-177 in Proceedings of the Florida State Horticultural Society.

Bosamia, T. C., S. M. Dodia, G. P. Mishra, S. Ahmad, B. Joshi *et al.*, 2020 Unraveling the mechanisms of resistance to *Sclerotium rolfsii* in peanut (*Arachis hypogaea L.*) using comparative RNA-Seq analysis of resistant and susceptible genotypes. PLOS One 15**:** e0236823.

Bourdais, G., P. Burdiak, A. Gauthier, L. Nitsch, J. Salojärvi *et al.*, 2015 Large-scale phenomics identifies primary and fine-tuning roles for CRKs in responses related to oxidative stress. PLOS Genetics 11**:** e1005373.

Bu, D., H. Luo, P. Huo, Z. Wang, S. Zhang *et al.*, 2021 KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. Nucleic Acids Research 49**:** W317-W325.

Chee, K., 1987 Studies of sporulation, pathogenicity and epidemiology of *Corynespora cassiicola* on *Hevea rubber*.

Chen, K., L. Du and Z. Chen, 2003 Sensitization of defense responses and activation of programmed cell death by a pathogen-induced receptor-like protein kinase in *Arabidopsis*. Plant Molecular Biology 53**:** 61-74.

Chen, S., Y. Zhou, Y. Chen and J. Gu, 2018 fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34**:** i884-i890.

Cho, S., W. Chen and F. J. Muehlbauer, 2005 Constitutive expression of the Flavanone 3-hydroxylase gene related to pathotype-specific ascochyta blight resistance in Cicer arietinum L. Physiological and Molecular Plant Pathology 67**:** 100-107.

Conner, K., A. Hagan and L. Zhang, 2013 First report of *Corynespora cassiicola*-incited target spot on cotton in Alabama. Plant Disease 97**:** 1379-1379.

Corrado, G., S. Arciello, P. Fanti, L. Fiandra, A. Garonna *et al.*, 2008 The Chitinase A from the *baculovirus AcMNPV* enhances resistance to both fungi and herbivorous pests in tobacco. Transgenic Research 17**:** 557-571.

Dao, T. T., H. J. Linthorst and R. Verpoorte, 2011 Chalcone synthase and its functions in plant resistance. Phytochem Rev 10**:** 397-412.

Dasgupta, U., G. P. Mishra, H. K. Dikshit, D. C. Mishra, T. Bosamia *et al.*, 2021 Comparative RNA-Seq analysis unfolds a complex regulatory network imparting yellow mosaic disease resistance in mungbean [*Vigna radiata* (L.) R. Wilczek]. PLOS One 16**:** e0244593.

de Mello, F. E., V. S. Lopes-Caitar, H. Prudente, S. A. Xavier-Valencio, S. Franzenburg *et al.*, 2021 Sensitivity of *Cercospora spp.* from soybean to quinone outside inhibitors and methyl benzimidazole carbamate fungicides in Brazil. Tropical Plant Pathology 46**:** 69-80.

DeMers, L. C., N. R. Redekar, A. Kachroo, S. A. Tolin, S. Li *et al.*, 2020 A transcriptional regulatory network of *Rsv3*-mediated extreme resistance against Soybean mosaic virus. PLOS One 15**:** e0231658.

DeYoung, B. J., and R. W. Innes, 2006 Plant NBS-LRR proteins in pathogen sensing and host defense. Nat Immunol 7**:** 1243-1249.

Dicko, M. H., H. Gruppen, C. Barro, A. S. Traoré, W. J. van Berkel *et al.*, 2005 Impact of phenolic compounds and related enzymes in sorghum varieties for resistance and susceptibility to biotic and abiotic stresses. Journal of Chemical Ecology 31**:** 2671-2688.

Dodds, P. N., and J. P. Rathjen, 2010 Plant immunity: towards an integrated view of plant–pathogen interactions. Nature Reviews Genetics 11**:** 539-548.

Dong, H., S. Shi, C. Zhang, S. Zhu, M. Li *et al.*, 2018 Transcriptomic analysis of genes in soybean in response to *Peronospora manshurica* infection. BMC Genomics 19**:** 1-13.

Du, H., Y. Wang, J. Yang and W. Yang, 2015 Comparative transcriptome analysis of resistant and susceptible tomato lines in response to infection by *Xanthomonas perforans* race T3. Frontiers in Plant Science 6**:** 1173.

Duan, Y., W. Xin, F. Lu, T. Li, M. Li *et al.*, 2019 Benzimidazole-and QoI-resistance in *Corynespora cassiicola* populations from greenhouse-cultivated cucumber: An emerging problem in China. Pesticide Biochemistry And Physiology 153**:** 95-105.

Faske, T., 2016. Target spot of soybean: What Do We Know? Arkansas Row Crop.https://agfax.com/2016/11/02/arkansas-soybeans-target-spot-what-do-we-know/

Fortunato, A. A., L. Araujo and F. Á. Rodrigues, 2017 Association of the production of phenylpropanoid compounds at the infection sites of *Corynespora cassiicola* with soybean resistance against target spot. Journal of Phytopathology 165**:** 131-142.

Fortunato, A. A., D. Debona, A. M. Bernardeli and F. A. Rodrigues, 2015 Defence-related enzymes in soybean resistance to target spot. Journal of Phytopathology 163**:** 731-742.

French, E., B. S. Kim and A. S. Iyer-Pascuzzi, 2016 Mechanisms of quantitative disease resistance in plants. Semin Cell Dev Biol 56**:** 201-208.

Fukuoka, S., N. Saka, Y. Mizukami, H. Koga, U. Yamanouchi *et al.*, 2015 Gene pyramiding enhances durable blast disease resistance in rice. Scientific Reports 5**:** 1-7.

Gao, S., F. Wang, J. Niran, N. Li, Y. Yin *et al.*, 2021 Transcriptome analysis reveals defense-related genes and pathways against *Xanthomonas campestris pv. vesicatoria* in pepper (*Capsicum annuum L.*). PLOS One 16**:** e0240279.

Ge, S. X., D. Jung and R. Yao, 2020 ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics 36**:** 2628-2629.

Godoy, C. V., 2015 Target spot, pp. 62-63 in *Compendium of soybean diseases and pests*, edited by G. L. Hartman, J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis *et al.* Am Phytopath Society.

Hamberger, B., M. Ellis, M. Friedmann, C. de Azevedo Souza, B. Barbazuk *et al.*, 2007 Genome-wide analyses of phenylpropanoid-related genes in *Populus trichocarpa*, *Arabidopsis thaliana*, and *Oryza sativa*: the Populus lignin toolbox and conservation and diversification of angiosperm gene families. Botany 85**:** 1182-1201.

Hartman, G. L., J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis *et al.*, 2015 *Compendium of soybean diseases and pests*. Am Phytopath Society.

Higashi, K., Y. Ishiga, Y. Inagaki, K. Toyoda, T. Shiraishi *et al.*, 2008 Modulation of defense signal transduction by flagellin-induced WRKY41 transcription factor in *Arabidopsis* thaliana. Molecular Genetics and Genomics 279**:** 303-312.

Huang, H., S. D. Qi, F. Qi, C. A. Wu, G. D. Yang *et al.*, 2010 NtKTI1, a Kunitz trypsin inhibitor with antifungal activity from Nicotiana tabacum, plays an important role in tobacco's defense response. The FEBS Journal 277**:** 4076-4088.

Jan, R., M. Aaqil Khan, S. Asaf, J.-R. Park, I.-J. Lee *et al.*, 2021 *Flavonone 3-hydroxylase* relieves bacterial leaf blight stress in rice via overaccumulation of antioxidant flavonoids and induction of defense genes and hormones. International Journal Of Molecular Sciences 22**:** 6152.

Jin, J., F. Tian, D.-C. Yang, Y.-Q. Meng, L. Kong *et al.*, 2016 PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Research**:** gkw982.

Kabir, S. R., M. M. Rahman, S. Tasnim, M. R. Karim, N. Khatun *et al.*, 2016 Purification and characterization of a novel chitinase from *Trichosanthes dioica* seed with antifungal activity. International Journal Of Biological Macromolecules 84**:** 62-68.

Kim, K. H., Y. J. Kang, D. H. Kim, M. Y. Yoon, J.-K. Moon *et al.*, 2011 RNA-Seq analysis of a soybean near-isogenic line carrying bacterial leaf pustule-resistant and-susceptible alleles. DNA Research 18**:** 483-497.

Koenning, S., T. Creswell, E. Dunphy, E. Sikora and J. Mueller, 2006 Increased occurrence of target spot of soybean caused by *Corynespora cassiicola* in the Southeastern United States. Plant Disease 90**:** 974-974.

Kofsky, J., H. Zhang and B.-H. Song, 2021 Novel resistance strategies to soybean cyst nematode (SCN) in wild soybean. Scientific Reports 11**:** 7967.

Kohli, S. K., S. Bali, K. Khanna, P. Bakshi, P. Sharma *et al.*, 2019 A current scenario on role of brassinosteroids in plant defense triggered in response to biotic challenges, pp. 367-388 in brassinosteroids: Plant Growth And Development. Springer.

Krattinger, S. G., E. S. Lagudah, W. Spielmeyer, R. P. Singh, J. Huerta-Espino *et al.*, 2009 A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science 323**:** 1360-1363.

Kumar, M., A. Brar, M. Yadav, A. Chawade, V. Vivekanand *et al.*, 2018 Chitinases-potential candidates for enhanced plant resistance towards fungal pathogens. Agriculture 8**:** 88.

Lanubile, A., U. K. Muppirala, A. J. Severin, A. Marocco and G. P. Munkvold, 2015 Transcriptome profiling of soybean (*Glycine max*) roots challenged with pathogenic and non-pathogenic isolates of *Fusarium oxysporum*. BMC Genomics 16**:** 1089.

Lee, D. S., Y. C. Kim, S. J. Kwon, C.-M. Ryu and O. K. Park, 2017 The *Arabidopsis* cysteine-rich receptor-like kinase CRK36 regulates immunity through interaction with the cytoplasmic kinase BIK1. Frontiers in Plant Science 8**:** 1856.

Li, B., R. Fan, S. Guo, P. Wang, X. Zhu *et al.*, 2019 The *Arabidopsis* MYB transcription factor, MYB111 modulates salt responses by regulating flavonoid biosynthesis. Environmental and Experimental Botany 166**:** 103807.

Lin, F., M. Zhao, D. D. Baumann, J. Ping, L. Sun *et al.*, 2014 Molecular response to the pathogen *Phytophthora sojae* among ten soybean near isogenic lines revealed by comparative transcriptomics. BMC Genomics 15**:** 1-13.

Liu, D., M. Xin, X. Zhou, C. Wang, Y. Zhang *et al.*, 2017 Expression and functional analysis of the transcription factor-encoding Gene *CsERF004* in cucumber during *Pseudoperonospora cubensis* and *Corynespora cassiicola* infection. BMC Plant Biol 17**:** 96.

Livak, K. J., and T. D. Schmittgen, 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods 25**:** 402-408.

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

Lozovaya, V. V., A. V. Lygin, O. V. Zernova, A. V. Ulanov, S. Li *et al.*, 2007 Modification of phenolic metabolism in soybean hairy roots through down regulation of chalcone synthase or isoflavone synthase. Planta 225**:** 665-679.

Mahajan, M., and S. K. Yadav, 2014 Overexpression of a tea flavanone 3-hydroxylase gene confers tolerance to salt stress and *Alternaria solani* in transgenic tobacco. Plant Molecular Biology 85**:** 551-573.

McCabe, C. E., S. R. Cianzio, J. A. O'Rourke and M. A. Graham, 2018 Leveraging RNA-Seq to characterize resistance to brown stem rot and the Rbs3 locus in soybean. Molecular Plant-Microbe Interactions 31**:** 1083-1094.

McHale, L., X. Tan, P. Koehl and R. W. Michelmore, 2006 Plant NBS-LRR proteins: adaptable guards. Genome Biology 7**:** 1-11.

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

Miedes, E., R. Vanholme, W. Boerjan and A. Molina, 2014 The role of the secondary cell wall in plant resistance to pathogens. Frontiers in Plant Science 5**:** 358.

Miraeiz, E., U. Chaiprom, A. Afsharifar, A. Karegar, J. M. Drnevich *et al.*, 2020 Early transcriptional responses to soybean cyst nematode HG Type 0 show genetic differences among resistant and susceptible soybeans. Theoretical and Applied Genetics 133**:** 87-102.

Mitchell, H. J., S. A. Hall, R. Stratford, J. L. Hall and M. S. Barber, 1999 Differential induction of cinnamyl alcohol dehydrogenase during defensive lignification in wheat (*Triticum aestivum L.*): characterisation of the major inducible form. Planta 208**:** 31-37.

Nakabayashi, R., K. Yonekura-Sakakibara, K. Urano, M. Suzuki, Y. Yamada *et al.*, 2014 Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. The Plant Journal 77**:** 367-379.

Parkhi, V., V. Kumar, L. M. Campbell, A. A. Bell, J. Shah *et al.*, 2010 Resistance against various fungal pathogens and reniform nematode in transgenic cotton plants expressing *Arabidopsis* NPR1. Transgenic Research 19**:** 959-975.

Patel, S., K. Bowen, J. Patel and J. Koebernick, 2022 Evaluating target spot (*Corynespora cassiicola*) resistance in soybean (*Glycine max* (L.) Merrill) in a controlled environment. Crop Protection**:** 106018.

Patro, R., G. Duggal, M. I. Love, R. A. Irizarry and C. Kingsford, 2017 Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods 14**:** 417-419.

Poland, J. A., P. J. Balint-Kurti, R. J. Wisser, R. C. Pratt and R. J. Nelson, 2009 Shades of gray: the world of quantitative disease resistance. Trends in Plant Science 14**:** 21-29.

Pusztahelyi, T., 2018 Chitin and chitin-related compounds in plant–fungal interactions. Mycology 9**:** 189-201.

Reuber, T. L., and F. M. Ausubel, 1996 Isolation of *Arabidopsis* genes that differentiate between resistance responses mediated by the *RPS2* and *RPM1* disease resistance genes. The Plant Cell 8**:** 241-249.

Ribeiro, S., P. Label, D. Garcia, P. Montoro and V. Pujade-Renaud, 2021 Transcriptome profiling in susceptible and tolerant rubber tree clones in response to cassiicolin Cas1, a necrotrophic effector from *Corynespora cassiicola*. PLOS One 16**:** e0254541.

Richardson, K., M. Vales, J. Kling, C. Mundt and P. Hayes, 2006 Pyramiding and dissecting disease resistance QTL to barley stripe rust. Theoretical and Applied Genetics 113**:** 485-495.

Roy, C. B., H. Liu, A. Rajamani and T. Saha, 2019 Transcriptome profiling reveals genetic basis of disease resistance against *Corynespora cassiicola* in rubber tree (*Hevea brasiliensis*). Current Plant Biology 17**:** 2-16.

Sarowar, S., S. T. Alam, R. Makandar, H. Lee, H. N. Trick *et al.*, 2019 Targeting the pattern-triggered immunity pathway to enhance resistance to *Fusarium graminearum*. Mol Plant Pathol 20**:** 626-640.

Schlumbaum, A., F. Mauch, U. Vögeli and T. Boller, 1986 Plant chitinases are potent inhibitors of fungal growth. Nature 324**:** 365-367.

Schmutz, J., S. B. Cannon, J. Schlueter, J. Ma, T. Mitros *et al.*, 2010 Genome sequence of the palaeopolyploid soybean. Nature 463**:** 178-183.

Staal, J., M. Kaliff, S. Bohman and C. Dixelius, 2006 Transgressive segregation reveals two *Arabidopsis* TIR-NB-LRR resistance genes effective against *Leptosphaeria maculans*, causal agent of blackleg disease. The Plant Journal 46**:** 218-230.

Staal, J., M. Kaliff, E. Dewaele, M. Persson and C. Dixelius, 2008 RLM3, a TIR domain encoding gene involved in broad-range immunity of *Arabidopsis* to necrotrophic fungal pathogens. The Plant Journal 55**:** 188-200.

Stein, E., A. Molitor, K.-H. Kogel and F. Waller, 2008 Systemic resistance in *Arabidopsis* conferred by the mycorrhizal fungus *Piriformospora indica* requires jasmonic acid signaling and the cytoplasmic function of NPR1. Plant and Cell Physiology 49**:** 1747-1751.

Stracke, R., O. Jahns, M. Keck, T. Tohge, K. Niehaus *et al.*, 2010 Analysis of production of flavonol glycosides-dependent flavonol glycoside accumulation in *Arabidopsis thaliana* plants reveals MYB11-, MYB12-and MYB111-independent flavonol glycoside accumulation. New Phytologist 188**:** 985-1000.

Tang, D., G. Wang and J. M. Zhou, 2017 Receptor kinases in plant-pathogen interactions: more than pattern recognition. Plant Cell 29**:** 618-637.

Tariq, R., C. Wang, T. Qin, F. Xu, Y. Tang *et al.*, 2018 Comparative transcriptome profiling of rice near-isogenic line carrying *Xa23* under infection of *Xanthomonas oryzae pv. oryzae*. International Journal of Molecular Sciences 19**:** 717.

Toufiq, N., B. Tabassum, M. U. Bhatti, A. Khan, M. Tariq *et al.*, 2018 Improved antifungal activity of barley derived *chitinase I* gene that overexpress a 32 kDa recombinant chitinase in Escherichia coli host. Brazilian Journal Of Microbiology 49**:** 414-421.

USDA, 2021. World agricultural production:Foreign Agricultural Service/USDA. Global Market Analysis.Circular Series WAP July 2021.
https://usda.library.cornell.edu/concern/publications/5q47rn72z?locale=en

Van Verk, M. C., R. Hickman, C. M. Pieterse and S. C. Van Wees, 2013 RNA-Seq: revelation of the messengers. Trends in Plant Science 18**:** 175-179.

Verhage, A., S. C. van Wees and C. M. Pieterse, 2010 Plant immunity: it's the hormones talking, but what do they say? Plant Physiology 154**:** 536-540.

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

Wang, J., W. Tian, F. Tao, J. Wang, H. Shang *et al.*, 2020 TaRPM1 positively regulates wheat high-temperature seedling-plant resistance to *Puccinia striiformis f. sp. tritici*. Frontiers in Plant Science 10**:** 1679.

Wang, X., D. Zhang, N. Cui, Y. Yu, G. Yu *et al.*, 2018 Transcriptome and miRNA analyses of the response to *Corynespora cassiicola* in cucumber. Scientific Reports 8**:** 7798.

Wang, Z., M. Gerstein and M. Snyder, 2009 RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10**:** 57-63.

Whaley, A., J. Sheridan, S. Safari, A. Burton, K. Burkey *et al.*, 2015 RNA-seq analysis reveals genetic response and tolerance mechanisms to ozone exposure in soybean. BMC Genomics 16**:** 426.

Xavier, S. A., M. G. Canteri, D. Barros and C. V. Godoy, 2013 Sensitivity of *Corynespora cassiicola* from soybean to carbendazim and prothioconazole. Tropical Plant Pathology 38**:** 431-435.

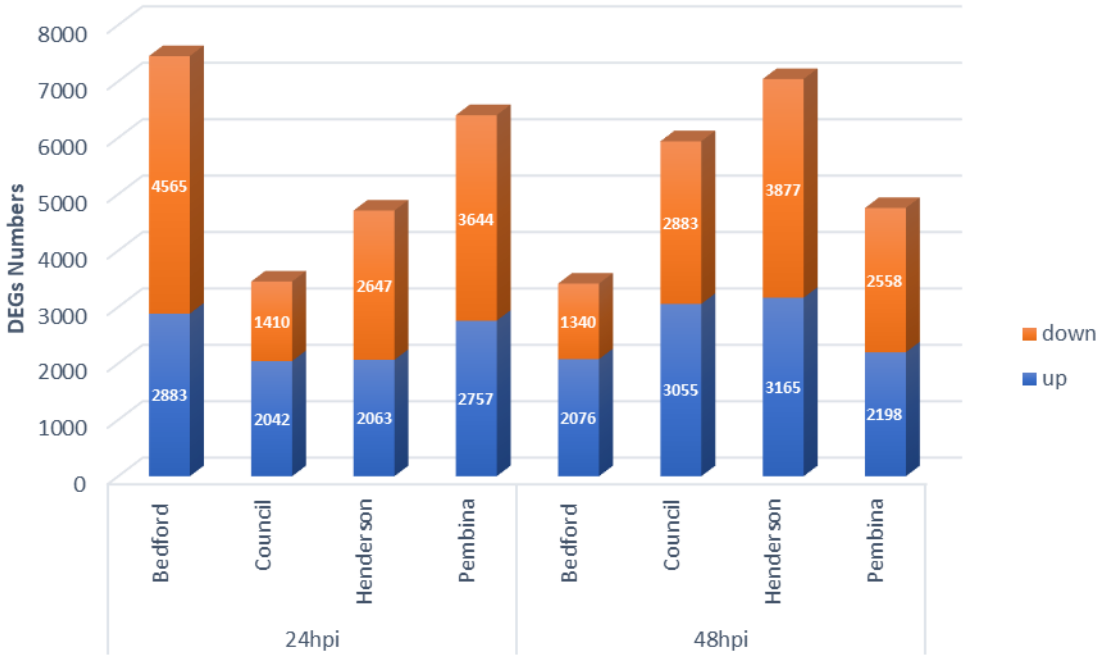
Yadeta, K. A., J. M. Elmore, A. Y. Creer, B. Feng, J. Y. Franco *et al.*, 2017 A Cysteine-rich protein kinase associates with a membrane immune complex and the cysteine residues are required for cell death. Plant Physiol 173**:** 771-787.

Yang, Z., H. Ma, H. Hong, W. Yao, W. Xie *et al.*, 2015 Transcriptome-based analysis of mitogen-activated protein kinase cascades in the rice response to *Xanthomonas oryzae* infection. Rice 8**:** 4.

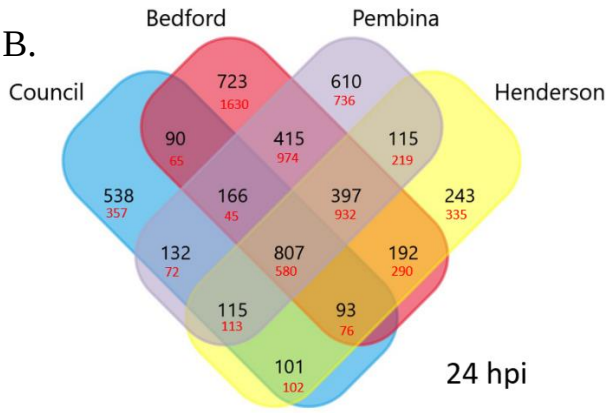
Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen *et al.*, 2012 Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13**:** 134.

Figure 3.1 Differntially expressed genes retrived from all four genotypes at 24 hpi and 48 hpi time of intervals. (A) Total numbers of DEGs (upregulated and downregulated) at each time points. (B) Venn diagram illustrating comparision of DEGs among all four genotypes, resistant(Bedford and Council) and susceptible (Henderson and Pembina). Black color numbers represents upregulated genes, and red color number represents downregulated genes.

A.



B.



C.

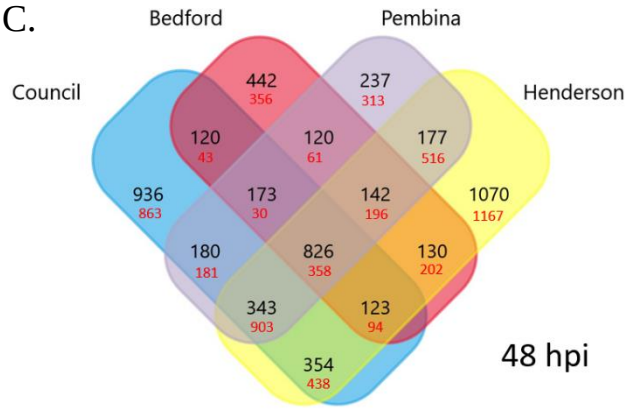


Figure 3.2 A hierarchical clustering tree of enriched biological processes. This clustering summarized the correlation between significant pathways and clustered together if these pathways share any common genes. The GO terms in the red box are some biosynthesis pathways and responses activated after a pathogen attack.

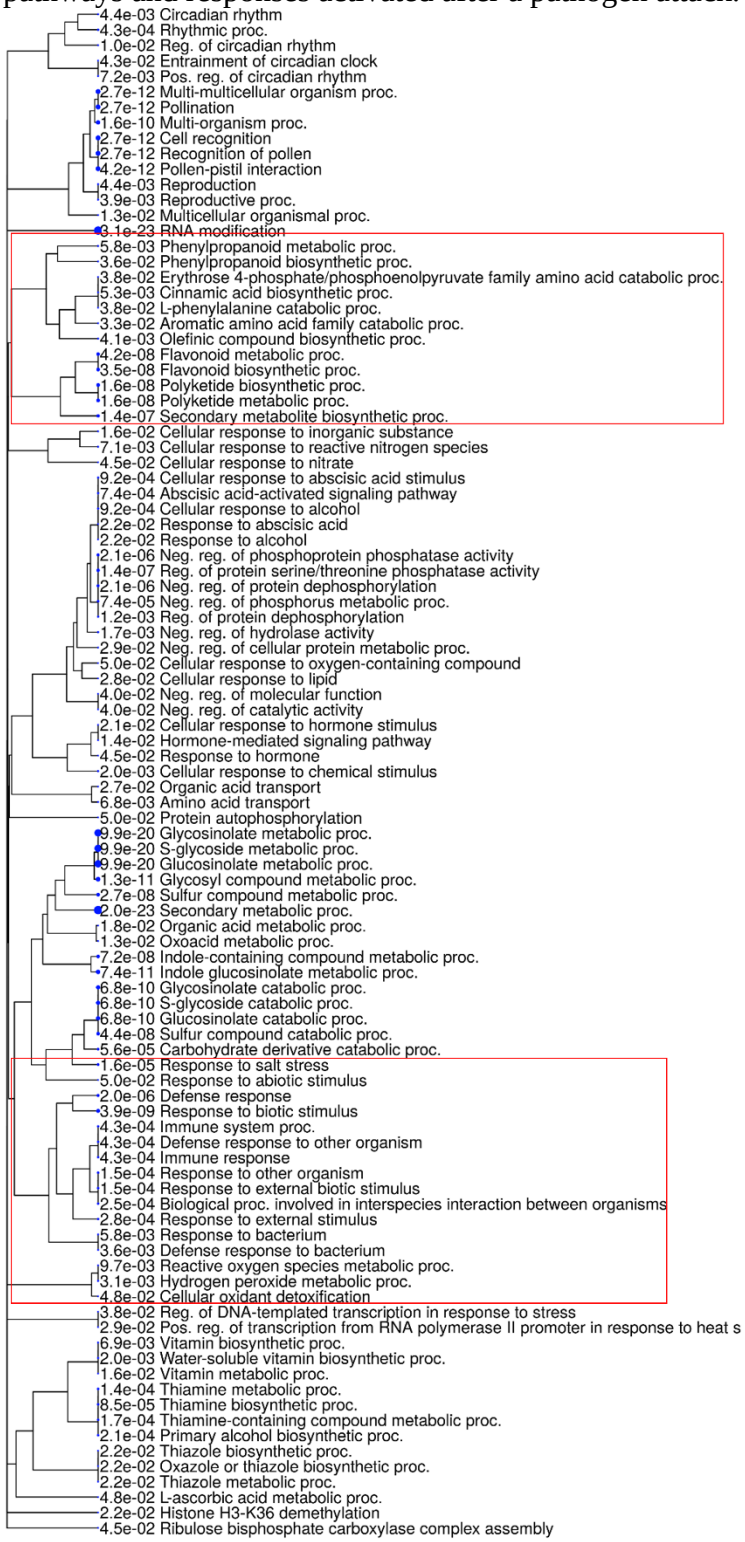
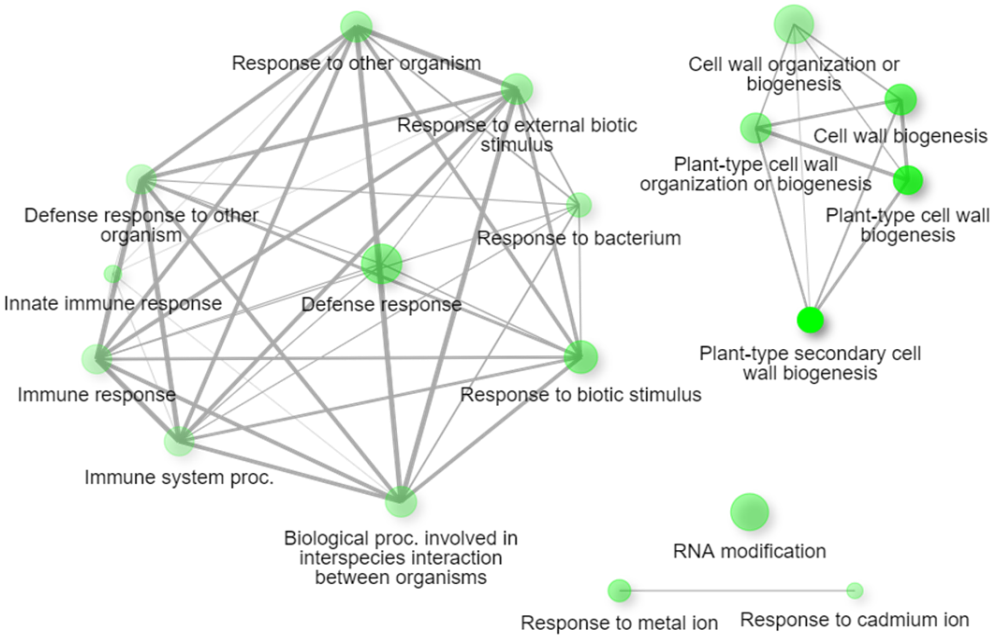


Figure 3.3 The interactive plot of significantly enriched biological processes. Activation of enrichment network of upregulated genes in resistant genotypes in response to *C. cassiicola* infection but not observed in susceptible, (A) Bedford (B) Council. Circles represent nodes (different biological processes), and lines represent a relation between two biological processes.

A



B

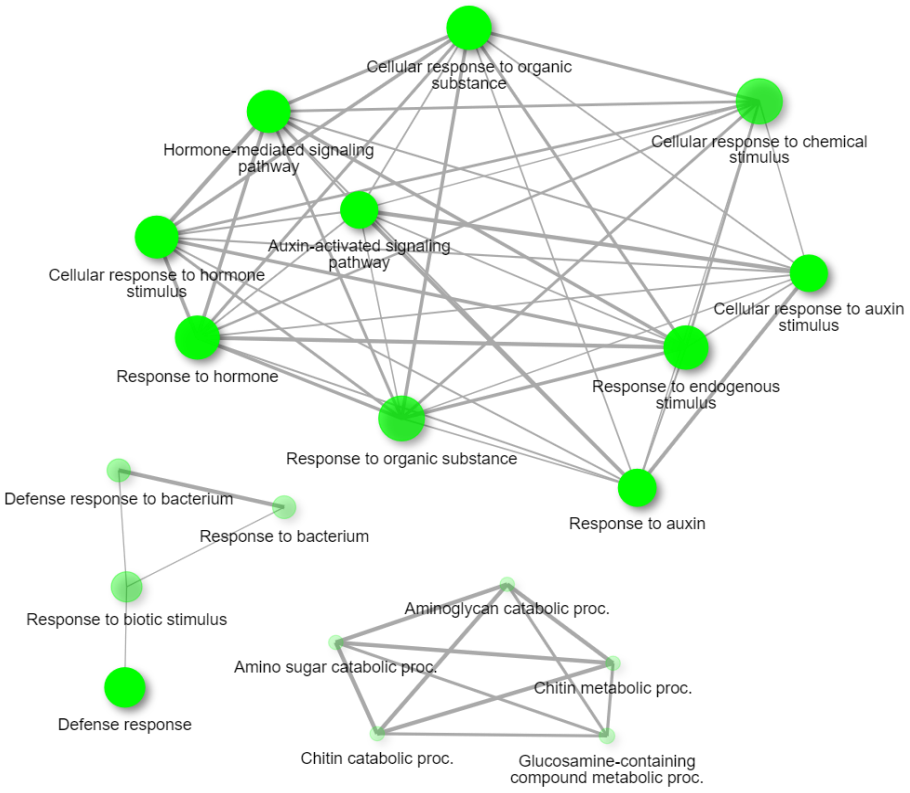


Figure 3.4 The illustration of the phenylpropanoid biosynthesis pathway. The numbers in the boxes represent coding for enzymes. The red color boxes represent upregulated genes.

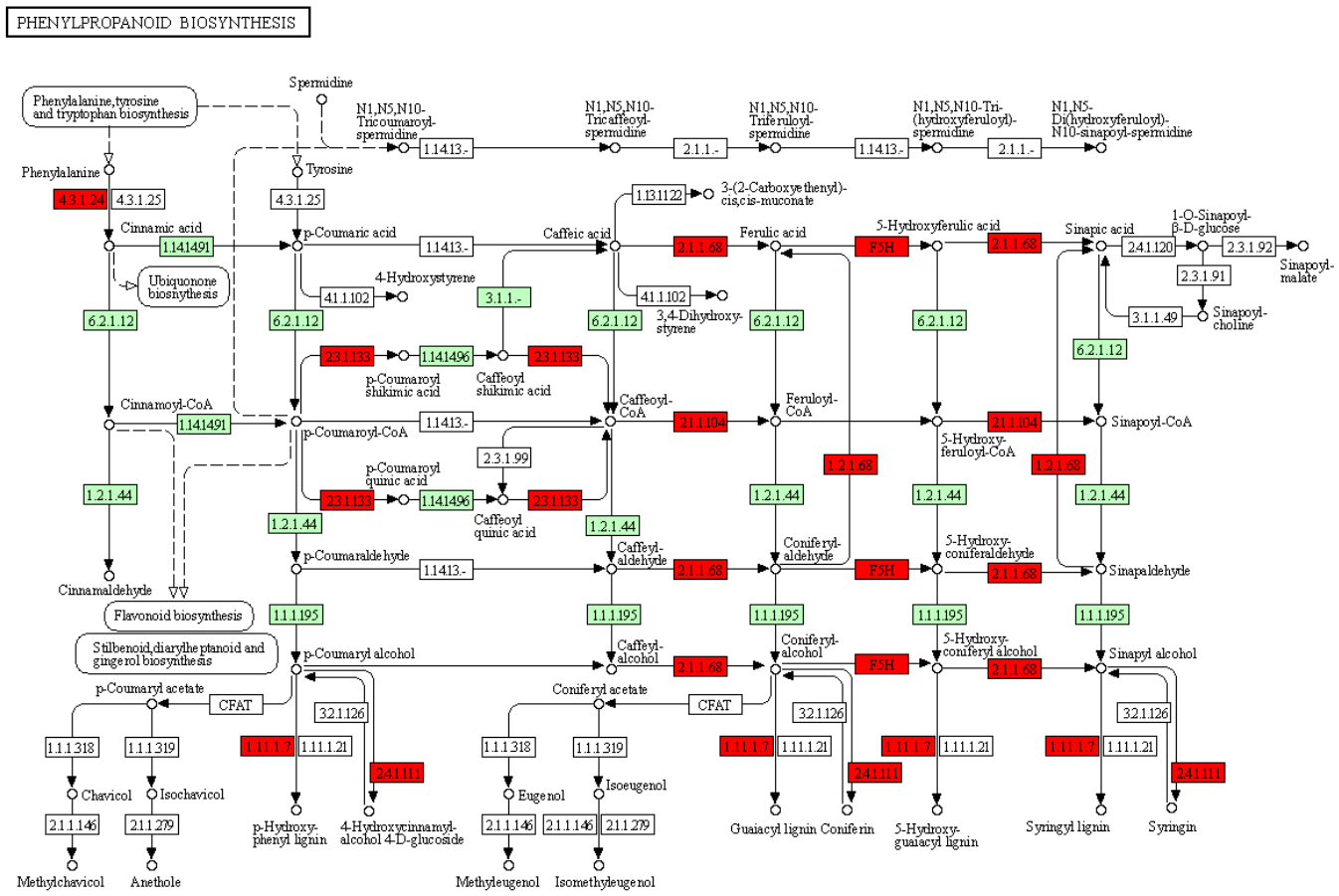


Figure 3.5 Classification of transcription factor families for four genotypes (A) Council (B) Bedford (C) Henderson (D) Pembina

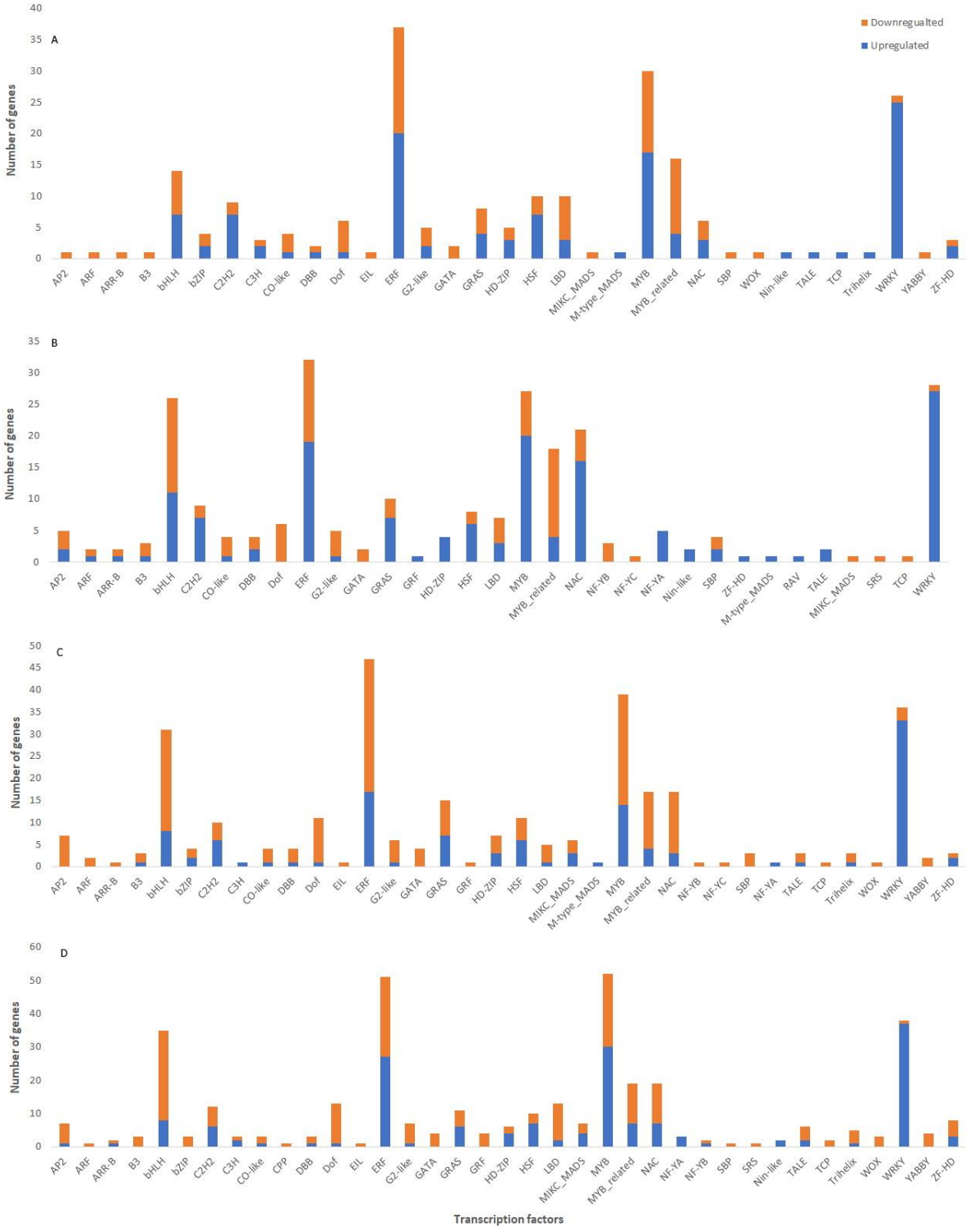


Figure 3.6 qRT-PCR based validation of DEGs in response to *C. cassiicola* inoculation at different time points. Relative gene expression is represented in Log2fold change obtained from RNA-Seq, and fold changes in qRT-PCR are calculated using the $2^{-\Delta\Delta CT}$ method.

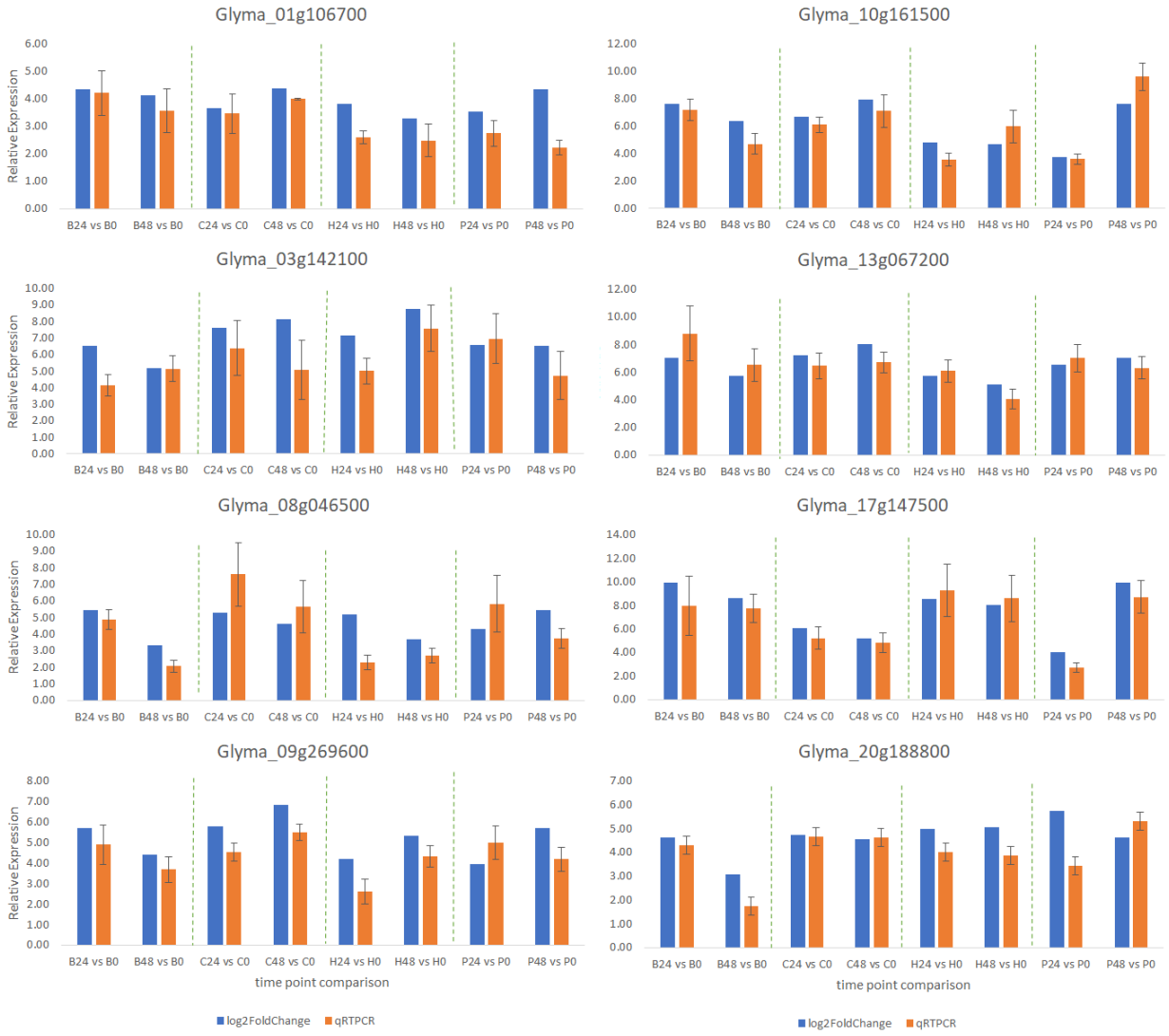


Figure 3.7 Heat maps represents the log2fold change based expression pattern of defense related genes in (A) Bedford and (B) Council after *C. cassiicola* inoculation compared to susceptible genotypes at different time points. The annotation of these genes was retrieved from Soybase.

A B

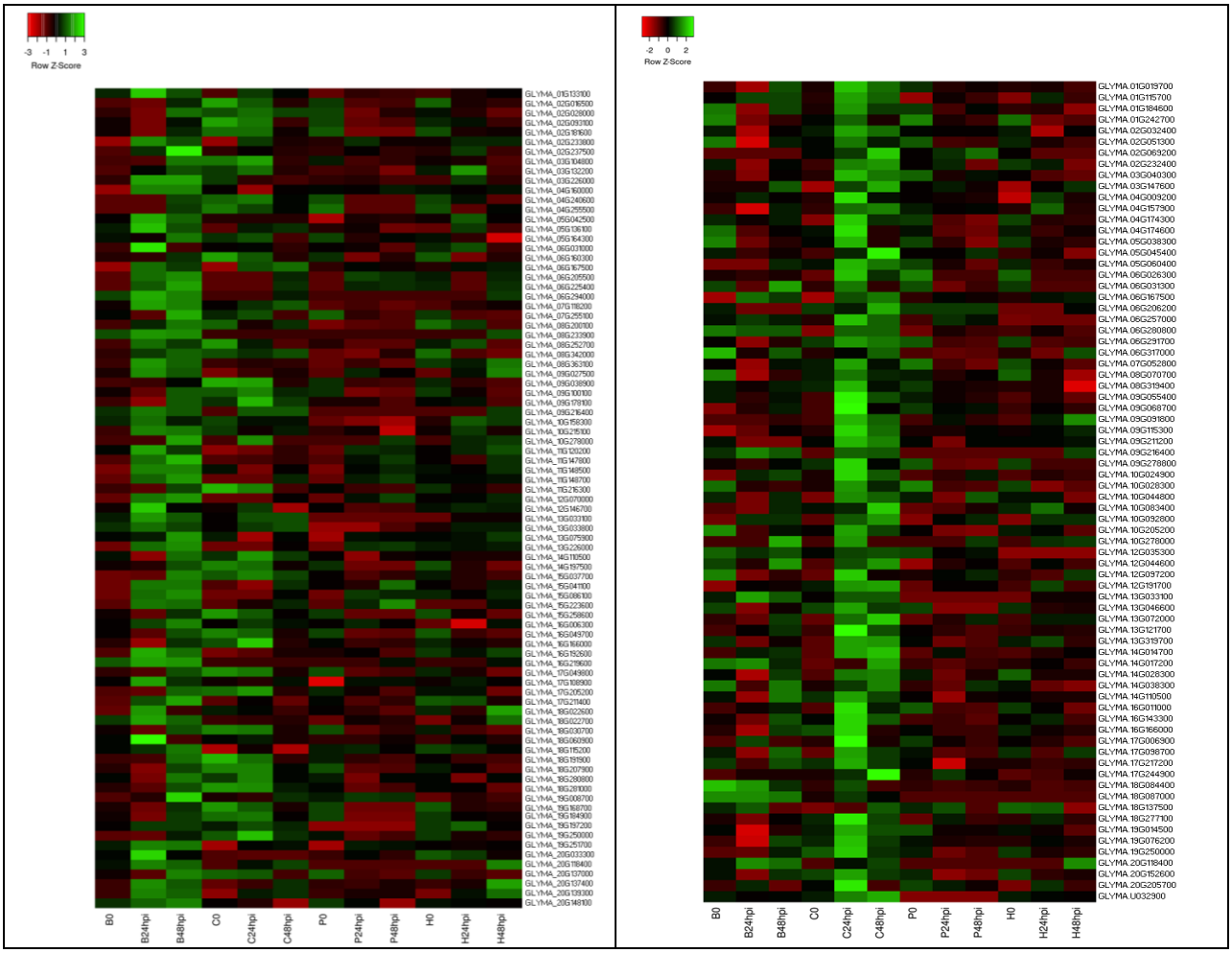


Figure 3.8 The heat map represents the log2fold change based expression pattern of secondary metabolites biosynthesis DEGs in Bedford and Council after *C. cassiicola* inoculation at different time points compared to susceptible genotypes. The annotation of these genes was retrieved from Soybase.

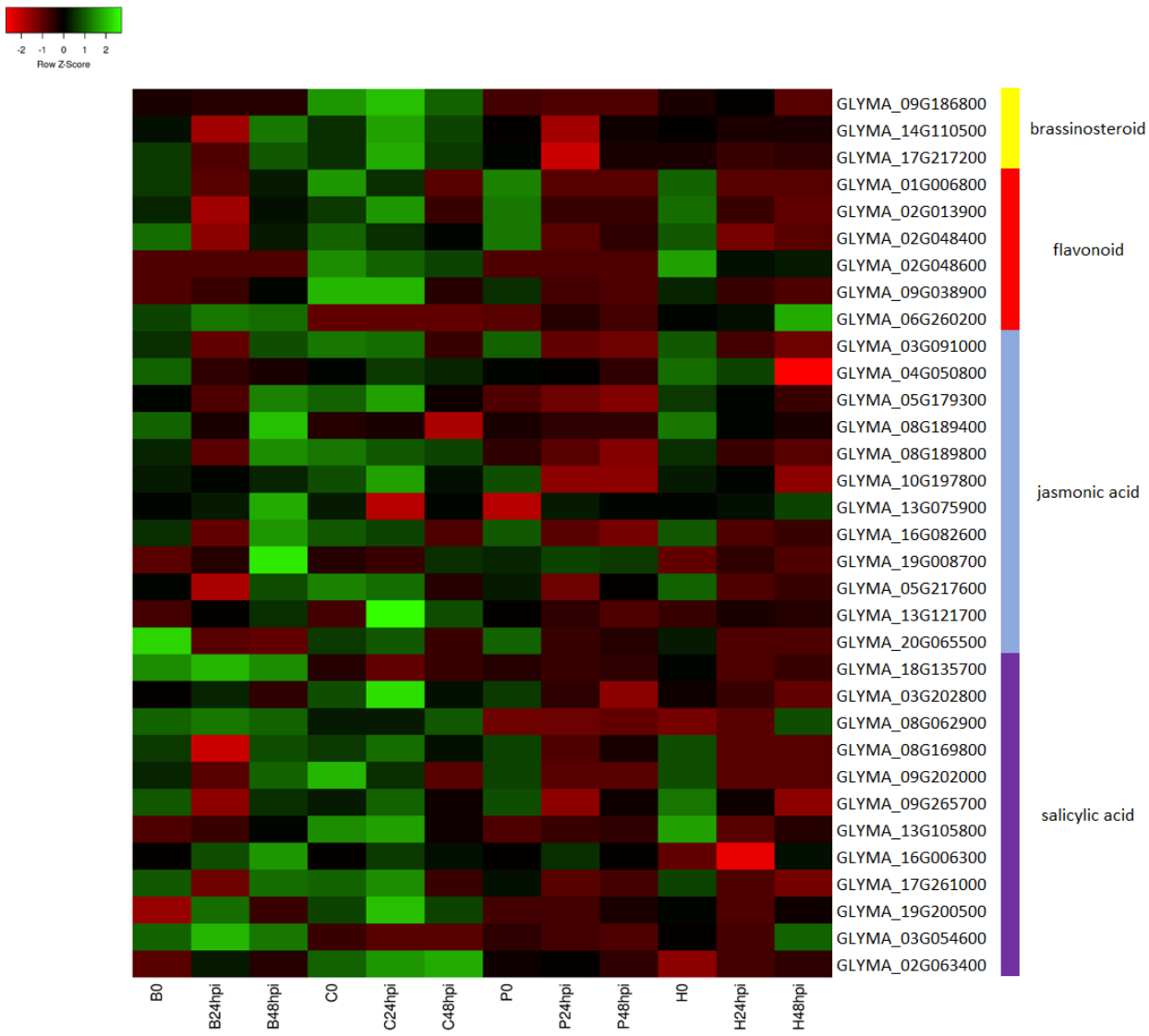


Table 3.1 An overview of statistics for the quality of sequencing data

Samples	Raw reads	Clean reads	Map reads	Raw data(G)	Clean data (G)	Q30 (%)	GC (%)
Bedford-0	26,390,793	25,793,638	23,582,978	7.9	7.8	93.39	45.04
Council-0	23,089,787	22,563,156	20,640,957	6.9	6.8	93.53	45.19
Henderson-0	22,838,228	22,381,404	20,209,026	6.9	6.7	93.19	44.79
Pembina-0	25,554,633	25,099,735	22,617,422	7.7	7.5	93.39	44.72
Bedford-24 hpi	28,423,021	28,174,656	24,664,671	8.6	8.5	93.19	44.69
Council-24 hpi	26,514,264	26,147,721	23,551,179	8.0	7.8	93.17	45.18
Henderson-24 hpi	27,493,410	27,128,017	23,310,773	8.2	8.1	93.05	44.69
Pembina-24 hpi	33,485,280	33,067,807	29,065,093	10.1	9.9	93.18	44.42
Bedford-48 hpi	25,860,451	25,641,661	22,746,284	7.8	7.7	93.37	44.94
Council-48 hpi	23,417,306	23,237,271	20,529,754	7.0	7.0	93.31	45.03
Henderson-48 hpi	25,075,023	24,887,851	21,578,946	7.5	7.4	93.09	44.53
Pembina-48 hpi	24,925,859	24,478,453	21,424,984	7.5	7.3	93.26	44.54

Table 3.2 Significantly enriched KEGG pathways of DEGs common in four genotypes.

Term	ID	Input number	Background number	*P*-Value	Corrected *P*-Value
Biosynthesis of secondary metabolites	gmx01110	129	2121	1.83E-19	1.79E-17
Metabolic pathways	gmx01100	180	4144	6.06E-13	2.97E-11
Phenylpropanoid biosynthesis	gmx00940	35	349	2.87E-11	9.37E-10
Circadian rhythm - plant	gmx04712	18	106	1.29E-09	3.17E-08
Cyanoamino acid metabolism	gmx00460	15	87	2.53E-08	4.97E-07
Flavonoid biosynthesis	gmx00941	15	94	6.30E-08	1.03E-06
Isoflavonoid biosynthesis	gmx00943	10	42	4.13E-07	5.78E-06
Starch and sucrose metabolism	gmx00500	23	299	4.73E-06	5.80E-05
Thiamine metabolism	gmx00730	7	43	0.000191	0.00199
Ubiquinone and other terpenoid-quinone biosynthesis	gmx00130	10	93	0.000203	0.00199
Fatty acid elongation	gmx00062	7	49	0.000391	0.003485
MAPK signaling pathway - plant	gmx04016	19	307	0.00043	0.003513
Plant hormone signal transduction	gmx04075	30	675	0.00233	0.017564
Ascorbate and aldarate metabolism	gmx00053	8	90	0.002635	0.018442
Cutin, suberine, and wax biosynthesis	gmx00073	6	54	0.003245	0.021197
Nitrogen metabolism	gmx00910	6	65	0.007429	0.043641
Carotenoid biosynthesis	gmx00906	6	66	0.007942	0.043641
Glycerolipid metabolism	gmx00561	10	158	0.008016	0.043641

Chapter 4. The genetic architecture of resistance to *Corynespora cassiicola* in soybean revealed by combining Association mapping and RNA-Seq studies

Abstract

Target spot caused by *Corynespora cassiicola* is a problematic disease in soybean growing regions, especially in humid tropical and subtropical. Resistant soybean (*Glycine Max*) genotypes do exist however, to date, the genetic mechanisms underlying this resistance to target spot has not been studied. Therefore, this is the first genome-wide association study (GWAS) conducted using 246 soybean accessions with the SoySNP50K array to understand the genetic architecture of *Glycine max* resistance to *C. cassiicola*. The results showed that 14 and 33 loci were significantly associated with resistance to LIM01 and SSTA *C. cassiicola* isolates, respectively. Of these, six loci had significant associations with resistance across both isolates. The gene expression level changes in GWAS-identified loci were assessed between resistant and susceptible genotypes through RNA-seq analysis of the leaf tissue collected at different time points after inoculation. Integrated results of GWAS and RNA-Seq analyses identified 238 differentially expressed genes within the 200 kb of the significant QTLs for disease severity ratings trait. These genes were involved in defense response to pathogen, innate immune response, chitinase activity, histone H3-K9 methylation, salicylic acid mediated signaling pathway, kinase activity, and biosynthesis of flavonoid, jasmonic acid, phenylpropanoid, and wax. Additionally, comparing our study with previous GWAS research involving different diseases aid in identifying 11 hotspot regions for resistance to multiple diseases and abiotic stress. This research provides insight into potentially valuable genetic resources that could potentially be valuable for the simultaneous improvement of soybean resistance to target spot and other diseases.

4.1 Introduction

Soybean (*Glycine max* L. Merrill) is a principal legume crop grown in over 50 countries and providing high-quality oil and protein for human and animal feed. Soybean production in the U.S. was about 120 million metric tons in 2021, which accounts for 33% of the world's production (USDA 2021). However, about 11% of soybean production is annually constrained by soybean diseases in the U.S. (HARTMAN *et al.* 2015). Target spot was first reported on soybean in 1945 (OLIVE *et al.* (1945) and is now considered widespread in the mid-south and the southeast U.S. due to frequent warm and humid conditions that tend to occur during crucial stages of soybean development. The occurrence of target spot, caused by *Corynespora cassiicola*, (Berk & M.A. Curtis) C.T. Wei (1950) is of concern for farmers since it can lead to 18 to 32% yield loss (DIXON *et al.* 2009; GODOY 2015; FASKE 2016).

Disease symptoms initially appear on foliage as lesions with alternating concentric rings of light and dark brown bands encircled by a yellow halo (GODOY 2015). Leaves prematurely senesce when infection is severe, resulting in yield loss and decreased seed quality (SEAMAN *et al.* 1965; KOENNING *et al.* 2006; GODOY 2015). The foliar pathogen, *C. cassiicola*, belongs to the phylum Ascomycota and has a wide host range, including cotton, tomato, rubber, and cucumber (BLAZQUEZ 1967; CHEE 1987; BLAZQUEZ 1991; CONNER *et al.* 2013). Host-specific *C. cassiicola* isolates have also been reported (DIXON *et al.* 2009; SUMABAT *et al.* 2018). Target spot is favored by warm temperatures (20 to 29 ◦C) and high humidity (> 90 %) (MOORE *et al.* 2021). .

Current disease control for target spot depends on field management practices and chemical applications. Although fungicides can control target spot, there is a concern over the development of fungicide resistance (XAVIER *et al.* 2013) which would reduce the effectiveness of the fungicide products (DUAN *et al.* 2019; RONDON AND LAWRENCE 2019). In addition,

fungicidal control of target spot is not always economical or environmentally friendly. Therefore, target spot resistant cultivars are the most economical and long-lasting strategy to manage this disease.

Conventional breeding requires a long selection process to develop a resistant soybean cultivar with competitive yield levels. Therefore, with the availability of plentiful molecular markers and affordable genotyping resources, it would be more efficient to identify genomic regions and incorporate them into a marker assisted breeding program which would accelerate the selection process and reduce the overall cost of developing new varieties.

The soybean *Glycine max* genome 'Williams 82' was assembled in 2010, covering 950 megabases (Mb) distributed in 20 linkage groups (SCHMUTZ *et al.* 2010). In addition, reference genomes for two additional soybean accessions, Zhonghuang 13 (ZH13) and wild soybean (W05), are available (SHEN *et al.* 2019; XIE *et al.* 2019). Six cultivated and two wild relatives of soybean were sequenced and analyzed to develop SoySNP50K, high-throughput SNP assay (SONG *et al.* 2013). This SNP assay has helped genotype 18,480 *G. max* accessions and 1168 *G. soja* accessions available in the United States Department of Agriculture (USDA) Soybean Germplasm Collection located at the University of Illinois. (SONG *et al.* 2015). Moreover, SoySNP50K has been used to conduct Genome-wide Association Studies (GWAS) and Quantitative Trait Locus (QTL) studies.

A GWAS associates variation across the whole genome with phenotypes and identifies genetic variations of essential traits. A panel of diverse individuals in GWAS provides a higher mapping resolution than the classic bi-parental QTL mapping method (KORTE AND FARLOW 2013). The major limitation of QTL mapping is the limited number of recombination events, while an association panel has accumulated recombination events over several generations,

providing higher mapping resolution. In soybean, GWAS studies have identified genes associated with resistance to: *Sclerotinia sclerotiorum* (IQUIRA *et al.* 2015; WEN *et al.* 2018), *Pythium sylvaticum* (LIN *et al.* 2020), *Heterodera glycines* Inchinohe (VUONG *et al.* 2015; RAVELOMBOLA *et al.* 2020), *Fusarium virguliforme* (BAO *et al.* 2015; ZHANG *et al.* 2015), *Phytophthora sojae* (LI *et al.* 2016), and *Macrophomina phaseolina* (COSER *et al.* 2017).

Target spot is also a concern in other crops such as rubber, cucumber, tomato, and cotton. Transcriptome profiling studies have been conducted in cucumber and rubber to identify genetic mechanisms conferring resistance to *C. cassiicola* (WANG et al. 2018; ROY et al. 2019; RIBEIRO et al. 2021). These studies have identified defense-related genes, biosynthesis of plant hormones, transcription factors, Ca2+ signaling pathways, secondary metabolites, and miRNA involved in the defense response of host plants against *C. cassiicola* infection. Molecular markers linked to tolerance to *C. cassiicola* infection in rubber were identified by screening 104 F_1 clones with 28 SSR markers and performing single marker analysis (OKTAVIA *et al.* 2021). In cucumber, a combination of SSR and SNP markers were used to screen a large number of F_2 plants to fine map a recessive resistance gene, *cca-3*, against *C. cassiicola*. Further screening of candidate genes in the region identified *RGA* (Resistance Gene Analog) gene containing CC-domain, nucleotide-binding domain (NB-ARC), and Leucine-rich repeat (LRR) domain that might play a role in hypersensitive responses to attack by the target spot pathogen (Wen et al. 2015). Resources like reference genomes, a diverse germplasm collection, markers, and other genotyping options are available for soybean; however, *C. cassiicola* resistance has not been investigated at a genomic level. Therefore, this study aims to determine genomic regions associated with resistance against *C. cassiicola* and identify candidate genes located within identified QTL regions.

4.2 Materials and Methods

4.2.1 Plant material and experiment design

A collection of 246 soybean genotypes were obtained from National Germplasm Database (GRIN). Two genotypes, 'Council' (PI 587091) and 'Henderson' (PI 665225) were used as resistant and susceptible checks, respectively (PATEL *et al.* 2022). Two experiments were performed with two *C. cassiicola* isolates (LIM01 and SSTA) in a greenhouse to identify genomic regions involved in the target spot resistance and locate candidate genes.

4.2.2 Phenotyping

Soybean genotypes were arranged in a randomized complete block design (RCBD), where treatments (genotypes) were replicated three times. Plants were grown in polypropylene Deepots D40L (Stuewe & Sons, Inc., Tangent, OR) filled with potting mix (PRO-MIX BX, Premier Tech Horticulture, Quakertown, PA). Three seeds per deepot were sown and thinned to one plant five days post-emergence. Both experiments were conducted in Plant Science Research Center, Auburn University, maintained between 24°C to 34°C, with 1000-watt halide bulbs providing supplemental light to ensure 14 photoperiods. Plants were grown to the V3-V4 stage before inoculation and watered as needed.

Inoculation and disease assessment methods were adapted from PATEL *et al.* (2022). Each 10-day old *C. cassiicola* culture, growing on V8 agar, was flooded with sterile distilled water and the colony surface was gently rubbed with a glass rod. The conidial suspension was filtered through double layers of sterile cheesecloth and the concentration were adjusted to 50,000 per ml. Plants were inoculated with conidia applied to the upper and lower leaf surfaces using a fine mist professional spray bottle. The inoculated plants were then moved to a mist chamber. The mist ran 2 seconds every 10 minutes, for 72 h, to create high humidity (>90%). At 14 days after

inoculation, resistance assessment was evaluated primarily on the percentage of damaged leaf area. The severity rating scale was 0 to 90% (VINCELLI AND HERSHMAN 2011). Genotypes were categorized as resistant if disease severity was below 25%, moderately susceptible with disease severity between 25 and 50%, and susceptible if more than 50% severity. Descriptive statistics of phenotypic data were performed with PROC GLIMMIX (Generalized linear mixed model) in SAS (version 9.4; SAS Institute Inc., Cary, NC).

4.2.3 Genotyping

The 246 soybean accessions were genotyped using Illumina Infinium SoySNP50K BeadChip (SONG *et al.* 2013; SONG *et al.* 2015). This association panel identified 42,180 SNPs, individual markers with a minor allele frequency <5 % or missing rate >10 % were omitted from the further analysis and 33,378 SNPs were used for the GWAS analysis.

4.2.4 Population structure and linkage disequilibrium estimation

The population structure was analyzed through STRUCTURE v.2.3.4 software to determine subpopulations (PRITCHARD *et al.* 2000). The 33,378 SNPs and following parameters were set for structure analysis, admixture model with five iterations of 50,000 burn-in and 50,000 Monte Carlo Markov Chain (MCMC) replications for k=1 to k=10. Admixture model with five iterations of 50,000 burn-in and 50,000 Monte Carlo Markov Chain (MCMC) replications for k=1 to k=10 were used for structure analysis of 33,378 SNPs. The number of subpopulations was identified by estimating DeltaK using STRUCTURE HARVESTER (EARL AND VONHOLDT 2012). Moreover, LD between pairs of SNP markers was determined by TASSEL 5 (BRADBURY *et al.* 2007) with 50 SNPs as a sliding window. LD decay rate was estimated as the distance from where r^2 dropped to half its maximum value (REMINGTON *et al.* 2001; HUANG *et al.* 2010).

4.2.5 GWAS analysis

The mean scores of disease severity ratings were used for the GWAS study and filtered SNPs were conducted in GAPIT v3. A statistical model FarmCPU (Fixed and random model Circulating Probability Unification), using PCA (Principal component analysis) and Kinship as covariates (FarmCPU_ (PCA+K), were used in this analysis (LIU *et al.* 2016; WANG AND ZHANG 2021). FarmCPU controls false positives and false negatives which make the model superior than generalized linear model (GLM) and the mixed linear model (MLM) and is widely used in soybean GWAS studies (KALER *et al.* 2018; CHAMARTHI *et al.* 2021). The significant threshold for the association of SNP and trait was considered if $-\log10(P) \geq 3.5$.

4.2.6 Prediction of candidate genes

Selection of candidate genes located within 200 kb genomic region on both sides of significant associated SNPs were examined and then annotated with a Glyma.Wm.82.a2 soybean reference genome in SoyBase (www.soybase.org) (CHENG *et al.* 2017; SUN *et al.* 2020).

4.2.7 Characterization of candidate genes based on RNA-Seq

The RNA-Seq data (From previous chapter) were collected to identify causative candidate genes around GWAS- identified loci. Briefly, two resistant genotypes (Bedford and Council) and two susceptible genotypes (Henderson and Pembina) were inoculated with a *C. cassiicola* isolate and with their respective controls (non-inoculated). Afterward, total RNA was extracted from three biological replications of young leaves using Direct-zol RNA Mini-Prep Kit (Zymo Research, Irvine, CA) according to the manufacturer's instructions. The cDNA library construction and 150-bp paired-end sequencing were performed by Illumina NovaSeq 6000 instrument (Novo gene Bioinformatics Technology Co., Ltd., CA, US). The raw data were collected, and the low-quality reads, adapters, and Poly A sequences were removed by using fastp software v0.23.1 (CHEN *et al.* 2018). The cleaned reads were mapped to the soybean

(*Glycine max* (L.) Merr.) reference genome (Schmutz *et al.* 2010), and transcript quantification was conducted by Salmon ver.0.9.1 (PATRO *et al.* 2017). To identify genes related to target spot resistance, we selected genes with differential expression levels significantly altered by *C. cassiicola* inoculation compared with the control using R-package DESeq2 (Version 1.37.6) (LOVE *et al.* 2014). Moreover, the genes with an absolute value of log2 fold change ≥ 2 (upregulated genes) or ≤ -2 (downregulated genes) and a false discovery rate (FDR) < 0.01 were considered significant DEGs.

4.3 Results

4.3.1 Phenotypic analysis

A total of 246 soybean accessions were evaluated for disease reaction to two *C. cassiicola* isolates (LIM01 and SSTA). The phenotypic data for both isolates exhibited near-normal distribution (Figure 4.1). The disease severity scores for isolate LIM01 ranged from 0 to 80 % with an overall mean of 34.20. Similarly, disease severity scores for isolate SSTA ranged from 3.3 to 76.6% with an overall mean of 32.69. The correlation for disease severity ratings for both isolates was 0.85. Analysis of variance indicated that genotypes and the genotype by isolate interaction were significantly different (P <0.0001), while there was no significant difference between isolates (P = 0.08) (Table 4.1).

4.3.2 Distribution of SNP Markers and Linkage Disequilibrium

Genotypic data using SoySNP50K Bead Chip was obtained for 246 accessions. A total of 42,180 SNPs were polymorphic in this population. Further filtering removed markers with a minor allele frequency <5 % or a missing rate >10 %, reducing the total number of SNPs to 33,378 (Figure 4.2). These SNPs were distributed in all twenty chromosomes with an average of 1,669 SNPs per chromosome. The highest number of SNPs were found on chromosome 18

(2,771 SNPs), while the lowest number of SNPs were found on Chr 20 (1106 SNPs). The average distance between two markers was around 28 kb. The markers on Chr. 18 have the lowest distance of 21 kb between two markers, while Chr. 20 has highest distance of 43.3 kb.

In this study, Linkage Disequilibrium (LD) analysis was performed for the filtered 33,378 SNPs which was calculated in r^2 and plotted against the physical distance. The fitting curve (red line) in the graph (Figure 4.3) estimates the rate of LD decay. The blue lines represent the half value of maximum r^2, and the intersection of green and blue lines represent the distance in kilobase pair (kb) where the power of maximum r^2 was reduced to half. For this study, the estimated LD decay rate dropped to half from its maximum value at about 249.47 kb which aligns with previous soybean GWAS studies (HWANG *et al.* 2014; DHANAPAL *et al.* 2015).

4.3.3 Population Structure

Population structure analysis was conducted for 246 genotypes using STRUCTURE software to determine the number of subpopulations. The likelihood value [LnP(D)] was obtained through analysis using the 33,378 SNPs. The likelihood value [LnP(D)] was further used to calculate the delta K value, determining the best-fitted subpopulation of 246 genotypes. The delta K graph had two peak values at K=2 and K=6. However, the best-fitted subpopulation group was considered as K=6 because it was able to differentiate soybean accessions collected from different locations (Figure 4.4). Under K=6, the 246 accessions were distributed in group 1 (includes 12 accessions), group 2 (34), group 3 (19), group 4 (87), group 5 (41), and group 6 (53). Further, groups 1 and 4 consist mainly of accessions from South Korea. At the same time, China and Vietnam dominated group 2, and the third group was mostly made of accessions from China. Accessions from the United States was largely found in group 5, and accessions from Japan prevailed in group 6 (Supplementary Table Appendix 2A).

4.3.4 Genomic regions identified for target spot

A total of 14 marker-trait associations at −Log10 (P) ≥ 3.5; P ≤ 0.0003 were identified with disease severity ratings for LIM01 isolate. These 14 SNPs were distributed among 12 chromosomes of the soybean genome, with Chr. 3 and 15 having two SNPs associated to the trait (Table 4.2, Figure 4.5). The allelic effect for disease severity was between -7.92 (ss715584729, Chr. 3) and 6.5 (ss715637816, Chr. 20). For traits involving disease ratings, a major allele is desirable when the allele effect is negative. In contrast a minor allele is desirable for the disease resistance when allelic effect is positive.

Thirty-three marker trait associations were discovered for disease severity ratings for the SSTA isolate. These markers were distributed on 10 chromosomes. Interestingly, Chr. 10 had 13 SNPs and Chr. 20 had seven SNPs associated with SSTA. Additionally, the allelic effect for disease severity ranged from -8.96 (ss715612259, Chr. 12) to 9.98 (ss715632608, Chr. 18).

Common significantly associated SNPs with disease severity ratings for both *C. cassiicola* isolates are important to develop a marker-based selection strategy for target spot. Six SNPs were identified as having a significant association with disease severity for both *C. cassiicola* isolates. A negative allelic effect was observed in only one of the SNP (ss715621861 on Chr. 15), while the rest five common SNPs had a positive allele effect for disease severity ratings for both isolates (Table 4.2).

4.3.5 Candidate genes associated with target spot resistance

A total of 41 SNPs were significantly associated with disease severity ratings in response to two *C. cassiicola* isolates. Candidate genes were searched for in the vicinity of 200kb region of either side of the associated SNP. Several markers were identified in less than 200kb region of chromosome 10 and 20 associated with disease severity to SSTA isolate; instead of gene mining

for all the SNPs on these chromosomes, only a few markers were selected that can span the entire region. For association on Chr. 10, only three markers were selected, covering all the 13 SNPs in the region to identify the genes. Similarly, for Chr, 20 only two markers were selected that cover the region for all seven SNPs to identify the genes in the vicinity. In total, 615 and 640 genes were found in the vicinity of associated SNPs for disease ratings for LIM01 and SSTA isolates, respectively. Additionally, the six SNPs found associated with disease severity ratings of both isolates had 262 genes in their vicinity of 200 kb. Further, functional annotation of these genes determined their role in many important pathways, such as defense against fungi, defense against other organisms such as bacteria, viruses, and oomycetes, chitin binding and chitinase activity, MAPKKK cascade, protein serine/threonine kinase activity, response to ethylene, biosynthesis of lignin, flavonoid, salicylic acid, brassinosteroid, and jasmonic acid (GRANT *et al.* 2010; MORALES *et al.* 2013).

4.3.6 Integration of information from GWAS and RNA-Seq analysis

A transcriptomic study was conducted for two resistant (Council and Bedford) and two susceptible (Henderson and Pembina) genotypes at different timepoints of *C. cassiicola* infection. RNA-Seq analysis identified 6,480 and 7,890 genes that are upregulated and downregulated, respectively, in either one of these genotypes. Integration of the RNA-Seq DEGs and candidate genes identified in GWAS analysis was performed. A total of 131genes in the vicinity of associated markers were upregulated after *C. cassiicola* infection in at least one of the genotypes, out of which 24 genes were upregulated only in resistant genotypes and 31 in susceptible genotypes. Similarly, 115 genes in the vicinity of associated markers were downregulated after *C. cassiicola* infection, out of which 29 were downregulated only in resistant genotypes and 16 in susceptible genotypes. Additionally, in these DEGs, eight genes

appeared to have interesting patterns after infection of *C. cassiicola* in which they were upregulation in one line while downregulated in another (Supplementary table Appendix 2B).

4.4 Discussion

Target spot can cause severe economic losses in warm and humid soybean growing regions. Since, soybean has a well assembled genome sequence, plentiful marker information, and a diverse germplasm collection, (SONG *et al.* 2013), there exists the ability to evaluate phenotypic variation against genetic information. To our best knowledge, this is the first study reporting molecular markers associated with target spot resistance in soybean.

Disease severity ratings differed widely among genotypes for each isolate. These data and genotypic information aid in identifying 14 and 33 marker-trait associations for LIM01 and SSTA isolates of *C. cassiicola*. Candidate genes and their putative functions were determined around the associated SNPs to target spot disease. A total of 993 candidate genes were discovered around associated SNPs. RNA-Seq is a powerful approach to evaluate differential expression of genes in response to pathogen attacks. Scrutinizing the genes exhibited in GWAS with the associated SNPs and combining soybean gene expression data after *C. cassiicola* infection, can increase the confidence in determining candidate target spot defense-associated genes. Transcriptome profiling of four genotypes at different time points of *C. cassiicola* infection and GWAS aid in reducing the list of 993 candidate genes to 238, which were differentially expressed. Among the 238 genes, several were determined to have a role in defense response to pathogen attack. These genes were involved in defense response to pathogen, innate immune response, chitinase activity, histone H3-K9 methylation, salicylic acid mediated signaling pathway, kinase activity, and biosynthesis of flavonoid, jasmonic acid, phenylpropanoid, and wax.

Leucine-rich repeat (LRR) domains are generally part of plant's immune receptors and defense genes. LRR-receptor-like kinase (RLK) genes are mainly involved in development and stress responses. A LRR-RLK gene known as flagellin-sensitive 2 (FLS2) binds bacterial flagellin and initiates pathogen-associated molecular patterns (PAMPs)-triggered immunity (CHINCHILLA *et al.* 2007). LRR-RLK also contributes to resistance to downy mildew in wheat and *Arabidopsis* (HOK *et al.* 2011; HU *et al.* 2018). A homologous LRR-RLK found in barley and wheat provided resistance to Fusarium head blight (THAPA *et al.* 2018). RPS2 belongs to a LRR Class of disease resistance that plays a vital role in conferring resistance to broad spectrum of pathogens (LI *et al.* 2019). We found four LRR-RLK and four genes with LRR belonging to a disease resistance family to be differentially expressed and present in the vicinity of associated SNPs.

RPM1-interacting protein 4 (RIN4) was found near ss715621861 (Chr.15) and was differentially expressed after *C. cassicola* infection. It is a plant immunity regulator that plays a major role in both PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI) (ZHAO *et al.* 2021). It is an inherently disordered protein (IDP) except at the region for pathogen-induced post-translational modifications. It is likely the center of the immune signaling pathway that can bind to multiple proteins to serve as a hub for protein complex formation (CUI *et al.* 2010; LIU *et al.* 2011; OLDFIELD AND DUNKER 2014; TORUÑO *et al.* 2018).

Plant-specific NAC domain proteins are transcription factors that regulate stress tolerance and defense response against various abiotic and biotic challenges (OH *et al.* 2005; PURANIK *et al.* 2012; YUAN *et al.* 2019). Four genes were in the narrowed-down candidate genes list, namely, NAC1, NAC47, and two copies of NAC90. The NAC1 gene originates a defense signal upon infection of *Pseudomonas syringae pv*. in tomato (KUD 2017). NAC090, along with other NAC

transcription factors, is essential in regulating salicylic acid accumulation and leaf senescence (KIM *et al.* 2018; ZHANG *et al.* 2020). In cotton, both NAC1 and NAC90 are involved in stress response to the fungal pathogen *Verticillium dahlia* and positively regulates verticillium wilt resistance (WANG *et al.* 2016; BAI *et al.* 2022).

Plant U-box protein PUB23 (two copies) and PUB25 were identified in the GWAS and differentially expressed after *C. cassiicola* infection. PUB23, along with other U-box proteins, negatively impacts tolerance to drought and *Fusarium oxysporum* (CHO et al. 2008; CHEN et al. 2014). Similarly, PUB25, along with PUB26, plays a role in the degradation of MYB6 and negatively impacts resistance to *Verticillium dahlia (MA et al. 2021).* It would be interesting to understand the structural and functional aspects of these U-box proteins and their impact on *C. cassiicola* resistance.

Durable resistance to multiple pathogens might be acquired by manipulating the genes encoding a small number of pathogen recognition proteins. Several genes and pathways are common response mechanisms for abiotic stress or pathogen attack. For example, a phenylpropanoid pathway is activated in response to pathogen attack and iron deficiency chlorosis (ZABALA *et al.* 2006; WATERS *et al.* 2018). RLK, *GmRLK18-1* provides pleiotropic resistance to sudden death syndrome and soybean cyst nematode (SCN) (SROUR *et al.* 2012). Likewise, a set of two genes (*GmSNAP18* and *GmSNAP11*), conferred resistance to both SCN and reniform nematode (RN) (USOVSKY *et al.* 2021). A close linkage of favorable alleles and/or the pleiotropic effect of genes can play an essential role in marker-based selection process, which could help to improve several traits simultaneously. This is more useful in traits like disease resistance that are controlled by a few genes. Collating the associated genomic regions from the current study and the previous publication assisted in identifying 11 hotspot regions in the

genome linked to different biotic and abiotic stress, such as iron deficiency chlorosis, SCN, RN, *Phytophthora sojae*, sclerotinia stem rot (*Sclerotinia sclerotiorum*), sudden death syndrome, and yellow mosaic virus (supplementary Table Appendix 2C). A block of less than two Mb on Chr.10 contains 13 marker-trait association (MTA) for target spot, 14 for *Phytophthora sojae*, one each for RN and SCN. This region consists of many genes that are involved in defense mechanisms such as pathogenesis-related 4, LRR-RLK, *Arabidopsis* toxicos en levadura 6 (ATL6), enhanced disease resistance 1 (EDR1), WRKY3, WRKY4, MYB15, pentatricopeptide repeat, and genes belonging to protein kinase family. Pathogenesis-related protein (PR) are genes that are encoded by the host only under pathogen or parasitic attack to provide a long-term resistance. These genes play a pivotal role in systemic acquired resistance in tissue that are uninfected and distal from the original infection site (SUDISHA *et al.* 2012; JAIN AND KHURANA 2018). WRKY3 and WRKY4 belong to one of the most prominent transcription factor families involved in large-scale transcriptional reprogramming as part of plant immune responses during pathogen attacks (EULGEM AND SOMSSICH 2007; BOSAMIA *et al.* 2020). *WRKY3* and *WRKY4* improved plant resistance to necrotrophic pathogens such as *Botrytis cinerea* in *Arabidopsis* (Lai et al. 2008). Another transcription factor, MYB15, activated lignin biosynthesis genes as part of effector-triggered immune responses providing resistance to *Pseudomonas syringae* (Kim et al. 2020). Another block of 1.86 Mb region on Chr. 16 had two MTAs for target spot, each one for *Sclerotinia sclerotiorum*, iron deficiency chlorosis, and SCN. This region contains 17 gene copies of disease resistance family protein / LRR family protein, five copies of disease resistance protein belonging to TIR-NBS-LRR family, four copies of receptor like protein (RLP), three copies of cysteine-rich RLK protein kinase 25, three copies of cytochrome P450 family protein, two copies of LRR-RLK, and two copies of LRR transmembrane protein kinase. These findings

suggest the presence of closely linked blocks affecting different biotic and abiotic stress. Markers from these regions can be integrated into soybean breeding programs to develop broad-spectrum resistance to pathogens.

4.5 Conclusion

This study demonstrates the first GWAS in soybean to elucidate the genetic basis of resistance to a fungal-incited disease, target spot. The GWAS analysis identified six common SNPs associated with disease traits in response to two different *C. cassiicola* isolates. A total of 238 DEGs were obtained by GWAS and RNA-Seq screening, and several genes related to disease resistance were investigated, including LRR-RLK, LRR protein belonging to disease resistance disease resistance family protein, NAC transcription factor, cytochrome P450, RIN4, Pentatricopeptide repeat (PPR) protein, and plant U-box protein. Further, 11 colocalized regions contributing to multiple disease resistance were distributed on different chromosomes. The QTLs and genes obtained from this research will be useful to breeders for improving target spot disease resistance in soybean.

References

Bai, S., Q. Niu, Y. Wu, K. Xu, M. Miao *et al.*, 2022 Genome-wide identification of the NAC transcription factors in gossypium hirsutum and analysis of their responses to *verticillium* wilt. Plants (Basel) 11.

Bao, Y., J. E. Kurle, G. Anderson and N. D. Young, 2015 Association mapping and genomic prediction for resistance to sudden death syndrome in early maturing soybean germplasm. Molecular Breeding 35**:** 1-14.

Blazquez, C., 1967 Corynespora leaf spot of cucumber, pp. 177-177 in Proceedings of the Florida State Horticultural Society.

Blazquez, C., 1991 Target spot. Compendium of tomato diseases.

Bosamia, T. C., S. M. Dodia, G. P. Mishra, S. Ahmad, B. Joshi *et al.*, 2020 Unraveling the mechanisms of resistance to *Sclerotium rolfsii* in peanut (Arachis hypogaea L.) using comparative RNA-Seq analysis of resistant and susceptible genotypes. PLOS One 15**:** e0236823.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss *et al.*, 2007 TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23**:** 2633-2635.

Chamarthi, S. K., A. S. Kaler, H. Abdel-Haleem, F. B. Fritschi, J. D. Gillman *et al.*, 2021 Identification and confirmation of loci associated with canopy wilting in soybean using Genome-wide association mapping. Frontiers in Plant Science 12.

Chee, K., 1987 Studies of sporulation, pathogenicity and epidemiology of *Corynespora cassiicola* on *Hevea rubber*.

Chen, S., Y. Zhou, Y. Chen and J. Gu, 2018 fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34**:** i884-i890.

Chen, Y. C., C. L. Wong, F. Muzzi, I. Vlaardingerbroek, B. N. Kidd *et al.*, 2014 Root defense analysis against *Fusarium oxysporum* reveals new regulators to confer resistance. Scientific Reports 4**:** 5584.

Cheng, Y., Q. Ma, H. Ren, Q. Xia, E. Song *et al.*, 2017 Fine mapping of a *Phytophthora*-resistance gene *RpsWY* in soybean (*Glycine max* L.) by high-throughput genome-wide sequencing. Theoretical and Applied Genetics 130**:** 1041-1051.

Chinchilla, D., C. Zipfel, S. Robatzek, B. Kemmerling, T. Nürnberger *et al.*, 2007 A flagellin-induced complex of the receptor *FLS2* and *BAK1* initiates plant defence. Nature 448**:** 497-500.

Cho, S. K., M. Y. Ryu, C. Song, J. M. Kwak and W. T. Kim, 2008 *Arabidopsis* PUB22 and PUB23 are homologous U-Box E3 ubiquitin ligases that play combinatory roles in response to drought stress. Plant Cell 20**:** 1899-1914.

Conner, K., A. Hagan and L. Zhang, 2013 First report of *Corynespora cassiicola*-incited target spot on cotton in Alabama. Plant Disease 97**:** 1379-1379.

Coser, S. M., R. V. Chowda Reddy, J. Zhang, D. S. Mueller, A. Mengistu *et al.*, 2017 Genetic architecture of charcoal rot (*macrophomina phaseolina*) resistance in soybean revealed using a diverse panel. Frontiers in Plant Science 8.

Cui, H., Y. Wang, L. Xue, J. Chu, C. Yan *et al.*, 2010 *Pseudomonas syringae* effector protein AvrB perturbs *Arabidopsis* hormone signaling by activating MAP kinase 4. Cell host & Microbe 7**:** 164-175.

Dhanapal, A. P., J. D. Ray, S. K. Singh, V. Hoyos-Villegas, J. R. Smith *et al.*, 2015 Genome-wide association study (GWAS) of carbon isotope ratio (δ 13 C) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. Theoretical and Applied Genetics 128**:** 73-91.

Dixon, L. J., R. L. Schlub, K. Pernezny and L. Datnoff, 2009 Host specialization and phylogenetic diversity of *Corynespora cassiicola*. Phytopathology 99**:** 1015-1027.

Duan, Y., W. Xin, F. Lu, T. Li, M. Li *et al.*, 2019 Benzimidazole-and QoI-resistance in *Corynespora cassiicola* populations from greenhouse-cultivated cucumber: An emerging problem in China. Pesticide Biochemistry And Physiology 153**:** 95-105.

Earl, D. A., and B. M. Vonholdt, 2012 STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4**:** 359-361.

Eulgem, T., and I. E. Somssich, 2007 Networks of WRKY transcription factors in defense signaling. Current Opinion in Plant Biology 10**:** 366-371.

Faske, T., 2016. Target spot of soybean: What Do We Know? Arkansas Row Crop.https://agfax.com/2016/11/02/arkansas-soybeans-target-spot-what-do-we-know/

Godoy, C. V., 2015 Target spot, pp. 62-63 in *Compendium of soybean diseases and pests*, edited by G. L. Hartman, J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis *et al.* Am Phytopath Society.

Grant, D., R. T. Nelson, S. B. Cannon and R. C. Shoemaker, 2010 SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res 38**:** D843-846.

Hartman, G. L., J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis *et al.*, 2015 *Compendium of soybean diseases and pests*. Am Phytopath Society.

Hok, S., E. G. J. Danchin, V. Allasia, F. PanabiÈRes, A. Attard *et al.*, 2011 An *Arabidopsis* (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease. Plant, Cell & Environment 34**:** 1944-1957.

Hu, P., J. Liu, J. Xu, C. Zhou, S. Cao *et al.*, 2018 A malectin-like/leucine-rich repeat receptor protein kinase gene, RLK-V, regulates powdery mildew resistance in wheat. Mol Plant Pathol 19**:** 2561-2574.

Huang, X., T. Sang, Q. Zhao, Q. Feng, Y. Zhao *et al.*, 2010 Genome-wide association studies of 14 agronomic traits in rice landraces. Nature Genetics 42**:** 961-967.

Hwang, E.-Y., Q. Song, G. Jia, J. E. Specht, D. L. Hyten *et al.*, 2014 A genome-wide association study of seed protein and oil content in soybean. BMC Genomics 15**:** 1-12.

Iquira, E., S. Humira and B. François, 2015 Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. BMC Plant Biology 15**:** 1-12.

Jain, D., and J. P. Khurana, 2018 Role of pathogenesis-related (PR) proteins in plant defense mechanism, pp. 265-281 in Molecular Aspects of Plant-Pathogen Interaction. Springer.

Kaler, A. S., J. D. Ray, W. T. Schapaugh, M. K. Davies, C. A. King *et al.*, 2018 Association mapping identifies loci for canopy coverage in diverse soybean genotypes. Molecular Breeding 38**:** 1-12.

Kim, H. J., J.-H. Park, J. Kim, J. J. Kim, S. Hong *et al.*, 2018 Time-evolving genetic networks reveal a NAC troika that negatively regulates leaf senescence in *Arabidopsis*. Proceedings of the National Academy of Sciences 115**:** E4930-E4939.

Kim, S. H., P. Y. Lam, M.-H. Lee, H. S. Jeon, Y. Tobimatsu *et al.*, 2020 The Arabidopsis R2R3 MYB transcription factor MYB15 is a key regulator of lignin biosynthesis in effector-triggered immunity. Frontiers in Plant Science 11**:** 583153.

Koenning, S., T. Creswell, E. Dunphy, E. Sikora and J. Mueller, 2006 Increased occurrence of target spot of soybean caused by *Corynespora cassiicola* in the Southeastern United States. Plant Disease 90**:** 974-974.

Korte, A., and A. Farlow, 2013 The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9**:** 1-9.

Kud, J., 2017 *Regulation of the Defense-Related SlNAC1 Transcription Factor During Defense Signaling in Tomato*. University of Idaho.

Lai, Z., K. M. Vinod, Z. Zheng, B. Fan and Z. Chen, 2008 Roles of Arabidopsis WRKY3 and WRKY4 Transcription Factors in Plant Responses to Pathogens. BMC Plant Biology 8**:** 68.

Li, L., N. Guo, J. Niu, Z. Wang, X. Cui *et al.*, 2016 Loci and candidate gene identification for resistance to *Phytophthora sojae* via association analysis in soybean [*Glycine max* (L.) Merr.]. Molecular Genetics and Genomics 291**:** 1095-1103.

Li, Z., J. Huang, Z. Wang, F. Meng, S. Zhang *et al.*, 2019 Overexpression of Arabidopsis nucleotide-binding and leucine-rich repeat genes RPS2 and RPM1 (D505V) confers broad-spectrum disease resistance in rice. Frontiers in Plant Science 10**:** 417.

Lin, F., S. H. Wani, P. J. Collins, Z. Wen, W. Li *et al.*, 2020 QTL mapping and GWAS for identification of loci conferring partial resistance to *Pythium sylvaticum* in soybean (*Glycine max* (L.) Merr). Molecular Breeding 40**:** 1-11.

Liu, J., J. M. Elmore, Z.-J. D. Lin and G. Coaker, 2011 A receptor-like cytoplasmic kinase phosphorylates the host target RIN4, leading to the activation of a plant innate immune receptor. Cell Host & Microbe 9**:** 137-146.

Liu, X., M. Huang, B. Fan, E. S. Buckler and Z. Zhang, 2016 Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. PLoS Genetics 12**:** e1005767.

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

Ma, A., D. Zhang, G. Wang, K. Wang, Z. Li *et al.*, 2021 *Verticillium dahliae* effector VDAL protects MYB6 from degradation by interacting with PUB25 and PUB26 E3 ligases to enhance Verticillium wilt resistance. Plant Cell 33**:** 3675-3699.

Moore, C., N. Sharma, K. L. Bowen and J. Koebernick, 2021 Developing a greenhouse protocol for evaluating resistance to *Corynespora cassiicola* in cotton. Plant Disease.

Morales, A. M., J. A. O'Rourke, M. Van De Mortel, K. T. Scheider, T. J. Bancroft *et al.*, 2013 Transcriptome analyses and virus induced gene silencing identify genes in the Rpp4-mediated Asian soybean rust resistance pathway. Functional Plant Biology 40**:** 1029-1047.

Oh, S.-K., S. Lee, S. H. Yu and D. Choi, 2005 Expression of a novel NAC domain-containing transcription factor (CaNAC1) is preferentially associated with incompatible interactions between chili pepper and pathogens. Planta 222**:** 876-887.

Oktavia, F., Sudarsono and D. Dinarti, 2021 Preliminary QTL detection for Corynespora Leaf Fall disease resistance in rubber plant. Nusantara Bioscience 13**:** 52-60.

Oldfield, C. J., and A. K. Dunker, 2014 Intrinsically disordered proteins and intrinsically disordered protein regions. Annu Rev Biochem 83**:** 553-584.

Olive, L. S., D. Bain and C. Lefebvre, 1945 A leaf spot of Cowpea and Soybean caused by an undescribed species of *Helminthosporium*. Phytopathology 35**:** 822-831.

Patel, S., K. Bowen, J. Patel and J. Koebernick, 2022 Evaluating target spot (*Corynespora cassiicola*) resistance in soybean (*Glycine max* (L.) Merrill) in a controlled environment. Crop Protection**:** 106018.

Patro, R., G. Duggal, M. I. Love, R. A. Irizarry and C. Kingsford, 2017 Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods 14**:** 417-419.

Pritchard, J. K., M. Stephens and P. Donnelly, 2000 Inference of population structure using multilocus genotype data. Genetics 155**:** 945-959.

Puranik, S., P. P. Sahu, P. S. Srivastava and M. Prasad, 2012 NAC proteins: regulation and role in stress tolerance. Trends in Plant Science 17**:** 369-381.

Ravelombola, W. S., J. Qin, A. Shi, L. Nice, Y. Bao *et al.*, 2020 Genome-wide association study and genomic selection for tolerance of soybean biomass to soybean cyst nematode infestation. PLoS One 15**:** e0235089.

Remington, D. L., J. M. Thornsberry, Y. Matsuoka, L. M. Wilson, S. R. Whitt *et al.*, 2001 Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proceedings of The National Academy Of Sciences 98**:** 11479-11484.

Ribeiro, S., P. Label, D. Garcia, P. Montoro and V. Pujade-Renaud, 2021 Transcriptome profiling in susceptible and tolerant rubber tree clones in response to cassiicolin Cas1, a necrotrophic effector from *Corynespora cassiicola*. PLOS One 16**:** e0254541.

Rondon, M. N., and K. S. Lawrence, 2019 *Corynespora cassiicola* isolates from soybean in Alabama detected with G143A mutation in the cytochrome b gene. Plant Health Progress 20**:** 247-249.

Roy, C. B., H. Liu, A. Rajamani and T. Saha, 2019 Transcriptome profiling reveals genetic basis of disease resistance against *Corynespora cassiicola* in rubber tree (*Hevea brasiliensis*). Current Plant Biology 17**:** 2-16.

Schmutz, J., S. B. Cannon, J. Schlueter, J. Ma, T. Mitros *et al.*, 2010 Genome sequence of the palaeopolyploid soybean. Nature 463**:** 178.

Seaman, W. L., R. A. Shoemaker and E. A. Peterson, 1965 Pathogenicity of *Corynespora cassiicola* on soybean. Canadian Journal of Botany 43**:** 1461-1469.

Shen, Y., H. Du, Y. Liu, L. Ni, Z. Wang *et al.*, 2019 Update soybean Zhonghuang 13 genome to a golden reference. Sci. China Life Sci 62**:** 1257-1260.

Song, Q., D. L. Hyten, G. Jia, C. V. Quigley, E. W. Fickus *et al.*, 2013 Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLOS One 8**:** e54985.

Song, Q., D. L. Hyten, G. Jia, C. V. Quigley, E. W. Fickus *et al.*, 2015 Fingerprinting Soybean Germplasm and Its Utility in Genomic Research. G3 Genes|Genomes|Genetics 5**:** 1999-2006.

Srour, A., A. J. Afzal, L. Blahut-Beatty, N. Hemmati, D. H. Simmonds *et al.*, 2012 The receptor like kinase at Rhg1-a/Rfs2 caused pleiotropic resistance to sudden death syndrome and soybean cyst nematode as a transgene by altering signaling responses. BMC Genomics 13**:** 368.

Sudisha, J., R. Sharathchandra, K. Amruthesh, A. Kumar and H. S. Shetty, 2012 Pathogenesis related proteins in plant defense response, pp. 379-403 in Plant defence: Biological Control. Springer.

Sumabat, L. G., R. C. Kemerait Jr and M. T. Brewer, 2018 Phylogenetic diversity and host specialization of *Corynespora cassiicola* responsible for emerging target spot disease of cotton and other crops in the southeastern United States. Phytopathology 108**:** 892-901.

Sun, M., Y. Jing, X. Zhao, W. Teng, L. Qiu *et al.*, 2020 Genome-wide association study of partial resistance to *sclerotinia* stem rot of cultivated soybean based on the detached leaf method. PLoS One 15**:** e0233366.

Thapa, G., L. R. Gunupuru, J. G. Hehir, A. Kahla, E. Mullins *et al.*, 2018 A pathogen-responsive leucine rich receptor like kinase contributes to Fusarium resistance in cereals. Frontiers in Plant Science 9**:** 867.

Toruño, T. Y., M. Shen, G. Coaker and D. Mackey, 2018 Regulated Disorder: Posttranslational Modifications Control the RIN4 Plant Immune Signaling Hub. Molecular Plant-Microbe Interactions® 32**:** 56-64.

USDA, 2021. World agricultural production:Foreign Agricultural Service/USDA. Global Market Analysis.Circular Series WAP July 2021.
https://usda.library.cornell.edu/concern/publications/5q47rn72z?locale=en

Usovsky, M., N. Lakhssassi, G. B. Patil, T. D. Vuong, S. Piya *et al.*, 2021 Dissecting nematode resistance regions in soybean revealed pleiotropic effect of soybean cyst and reniform nematode resistance genes. The Plant Genome 14**:** e20083.

Vincelli, P., and D. E. Hershman, 2011 Assessing Foliar Diseases of Corn, Soybeans, and Wheat, pp. UK Cooperative Extension Service.

Vuong, T., H. Sonah, C. Meinhardt, R. Deshmukh, S. Kadam *et al.*, 2015 Genetic architecture of cyst nematode resistance revealed by genome-wide association study in soybean. BMC Genomics 16**:** 1-13.

Wang, J., and Z. Zhang, 2021 GAPIT Version 3: boosting power and accuracy for genomic association and prediction. Genomics, Proteomics & Bioinformatics 19**:** 629-640.

Wang, W., Y. Yuan, C. Yang, S. Geng, Q. Sun *et al.*, 2016 Characterization, Expression, and Functional Analysis of a Novel NAC Gene Associated with Resistance to Verticillium Wilt and Abiotic Stress in Cotton. G3 (Bethesda) 6**:** 3951-3961.

Wang, X., D. Zhang, N. Cui, Y. Yu, G. Yu *et al.*, 2018 Transcriptome and miRNA analyses of the response to *Corynespora cassiicola* in cucumber. Scientific Reports 8**:** 7798.

Waters, B. M., K. Amundsen and G. Graef, 2018 Gene expression profiling of iron deficiency chlorosis sensitive and tolerant soybean indicates key roles for phenylpropanoids under alkalinity stress. Frontiers in Plant Science 9**:** 10.

Wei, C., 1950 Notes on Corynespora. Mycological papers 34.

Wen, C., A. Mao, C. Dong, H. Liu, S. Yu *et al.*, 2015 Fine genetic mapping of target leaf spot resistance gene *cca-3* in cucumber, *Cucumis sativus* L. Theor Appl Genet 128**:** 2495-2506.

Wen, Z., R. Tan, S. Zhang, P. J. Collins, J. Yuan *et al.*, 2018 Integrating GWAS and gene expression data for functional characterization of resistance to white mould in soya bean. Plant Biotechnology Journal 16**:** 1825-1835.

Xavier, S. A., M. G. Canteri, D. Barros and C. V. Godoy, 2013 Sensitivity of *Corynespora cassiicola* from soybean to carbendazim and prothioconazole. Tropical Plant Pathology 38**:** 431-435.

Xie, M., C. Y.-L. Chung, M.-W. Li, F.-L. Wong, X. Wang *et al.*, 2019 A reference-grade wild soybean genome. Nature Communications 10**:** 1-12.

Yuan, X., H. Wang, J. Cai, D. Li and F. Song, 2019 NAC transcription factors in plant immunity. Phytopathology Research 1**:** 3.

Zabala, G., J. Zou, J. Tuteja, D. O. Gonzalez, S. J. Clough *et al.*, 2006 Transcriptome changes in the phenylpropanoid pathway of glycine max in response to *Pseudomonas syringae* infection. BMC Plant Biology 6**:** 1-18.

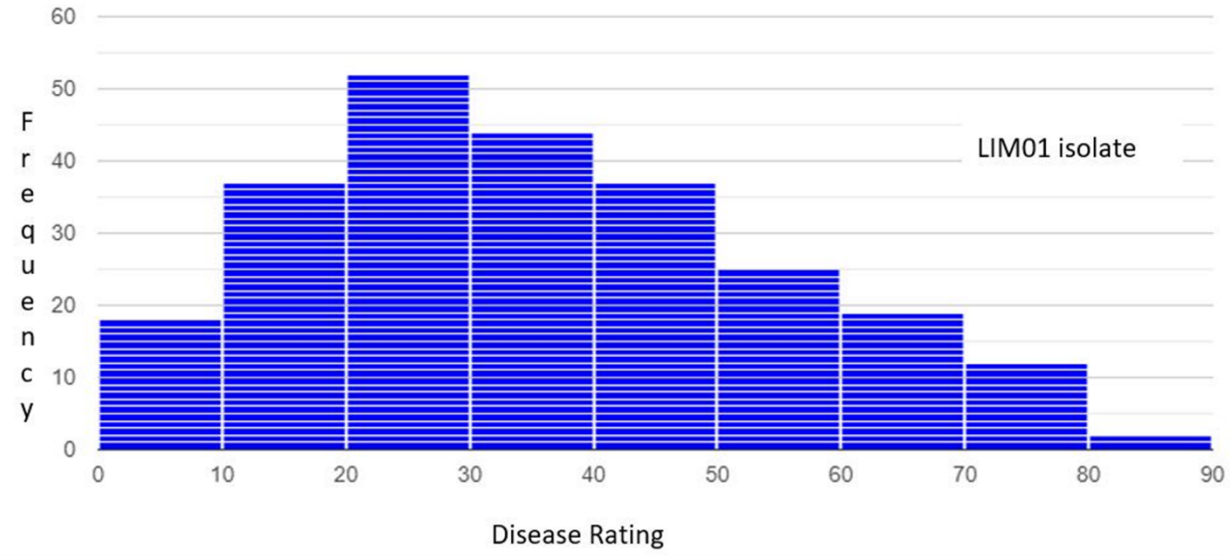
Zhang, J., A. Singh, D. S. Mueller and A. K. Singh, 2015 Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. The Plant Journal 84**:** 1124-1136.

Zhang, Y., H.-L. Wang, Z. Li and H. Guo, 2020 Genetic network between leaf senescence and plant immunity: crucial regulatory nodes and new insights. Plants 9**:** 495.

Zhao, G., D. Guo, L. Wang, H. Li, C. Wang *et al.*, 2021 Functions of RPM1-interacting protein 4 in plant immunity. Planta 253**:** 11.

Figure 4.1 Frequency distribution of disease severity ratings among 246 soybean accessions against *Corynespora cassiicola* infection. (A) LIM01 (B) SSTA

A



B

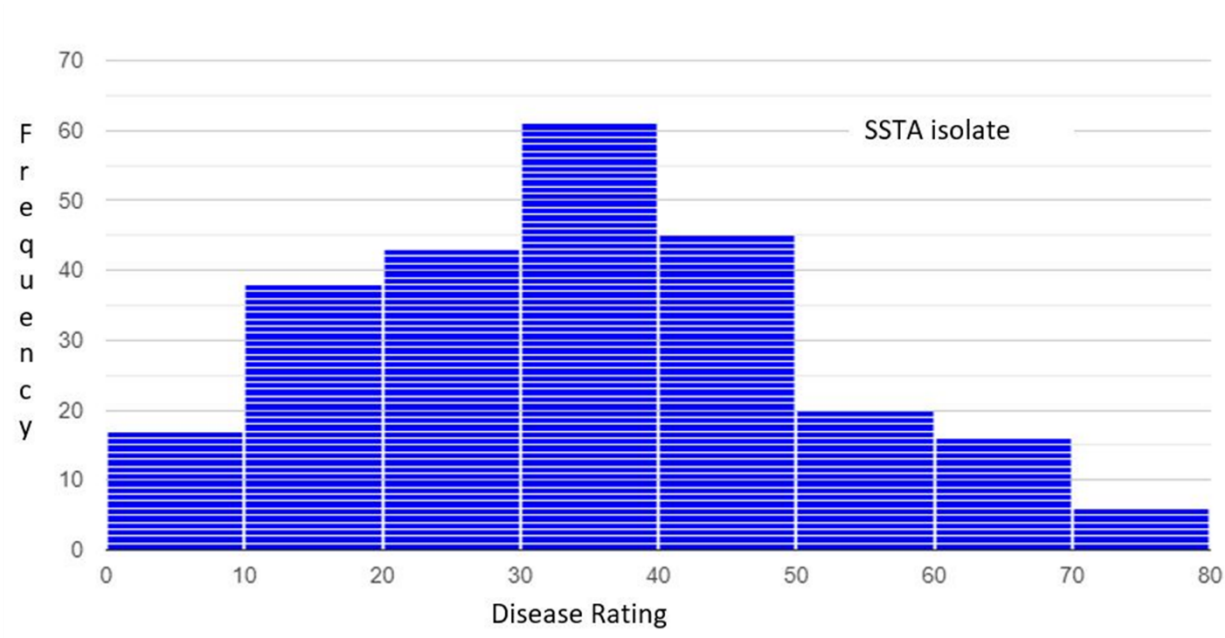


Figure 4.2 Distribution of SNP markers in the soybean genome (among 20 chromosomes)

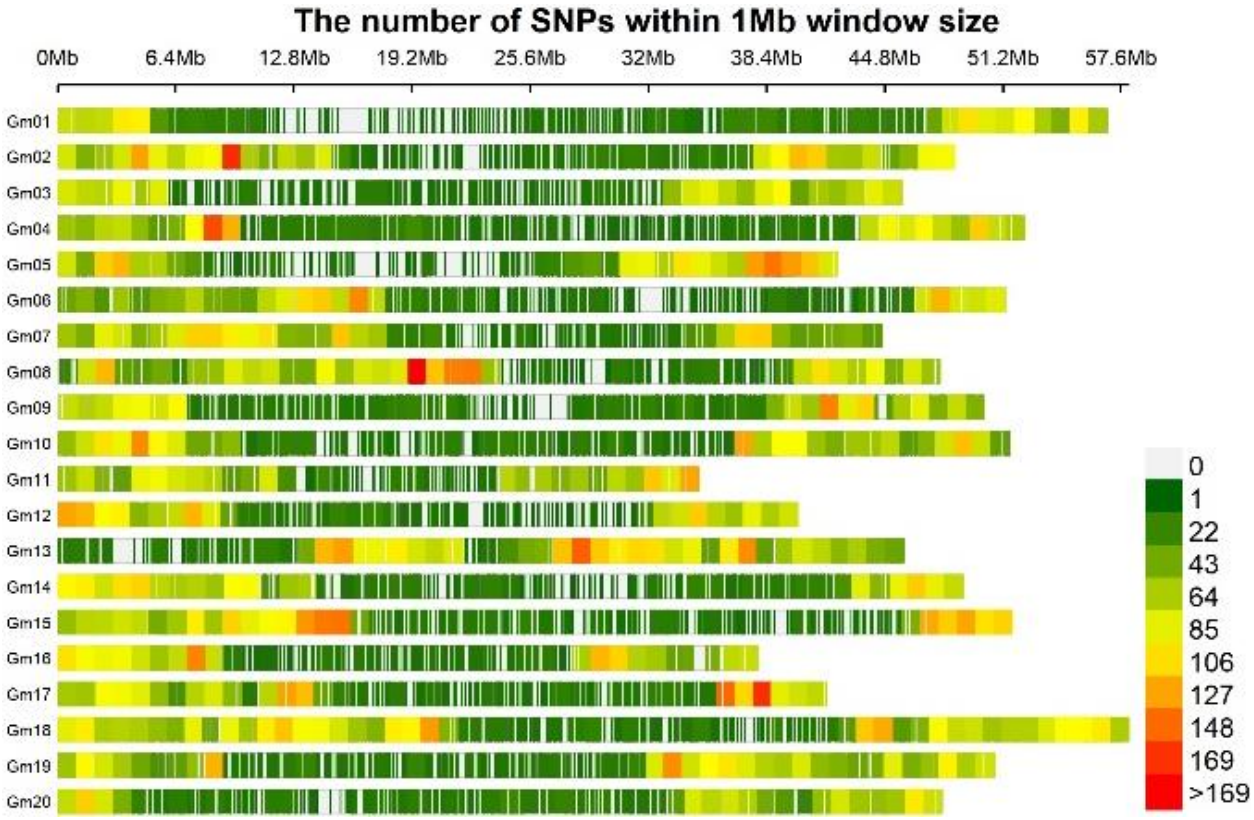


Figure 4.3 Genome-wide linkage disequilibrium (LD) decay rate estimated in 246 soybean accessions. The value on X- axis represents genetic distance in Kb (killo bases) and Y-axis represents the squared correlation coefficient r^2

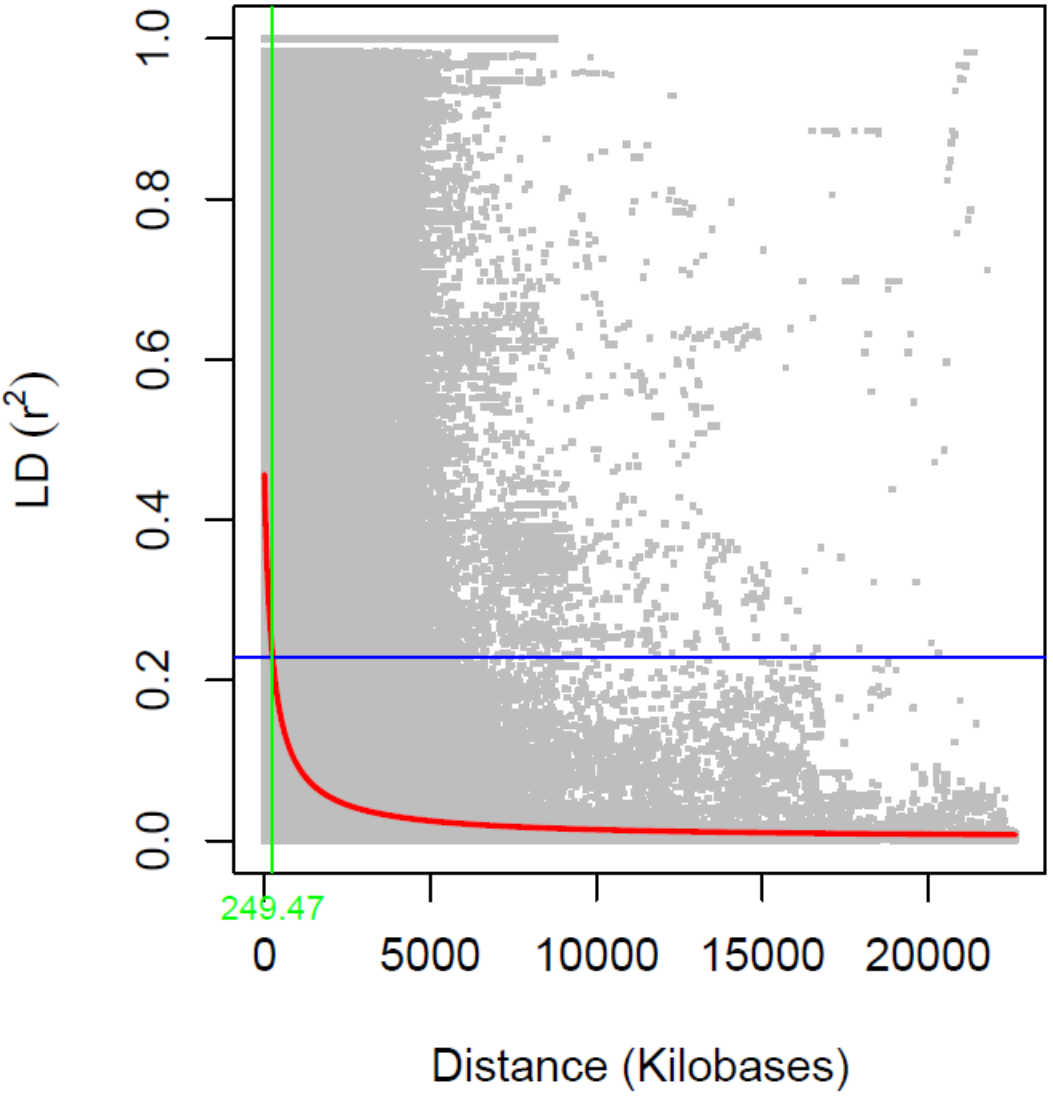


Figure 4.4 Population structure of the 246 soybean accessions.

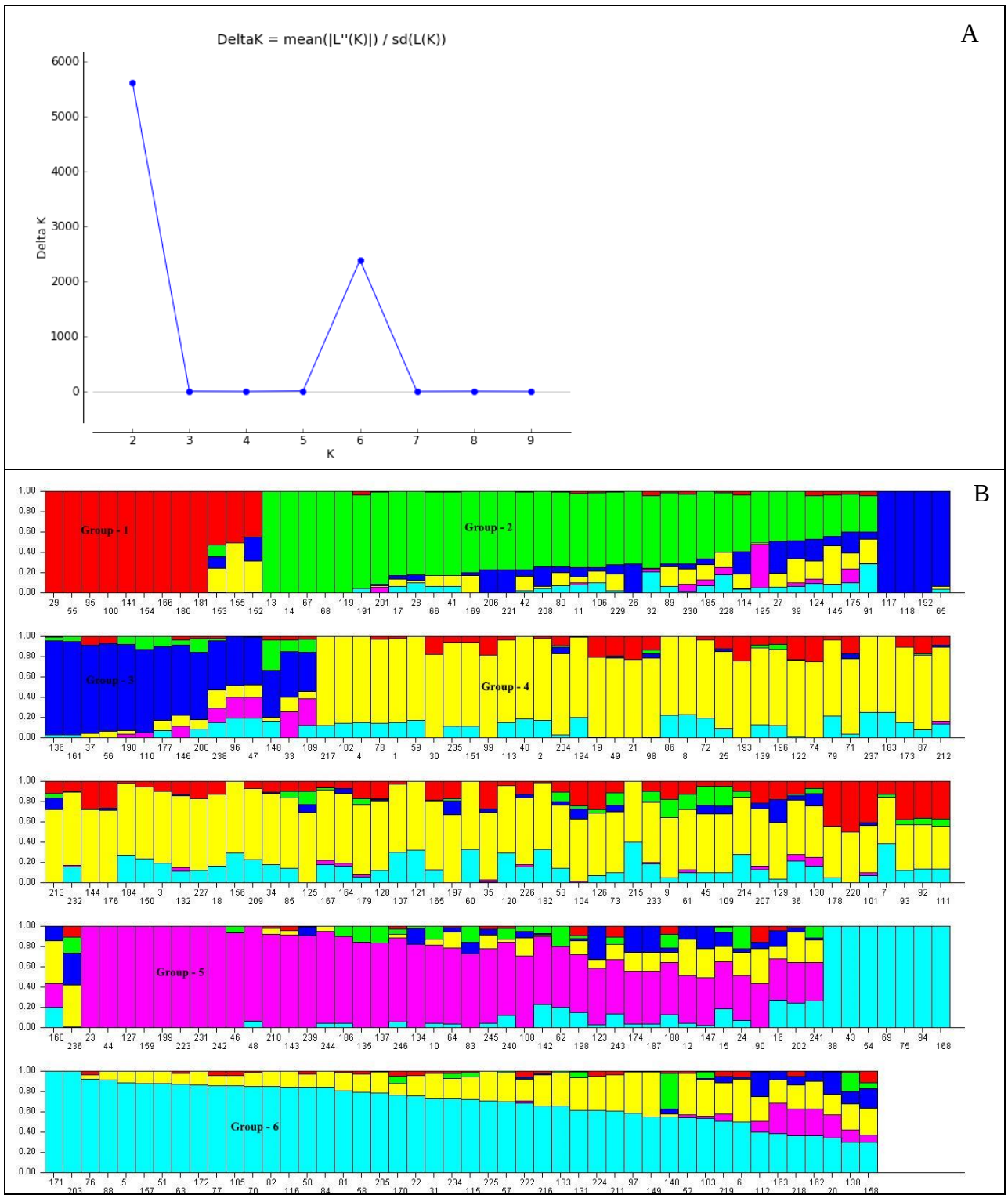
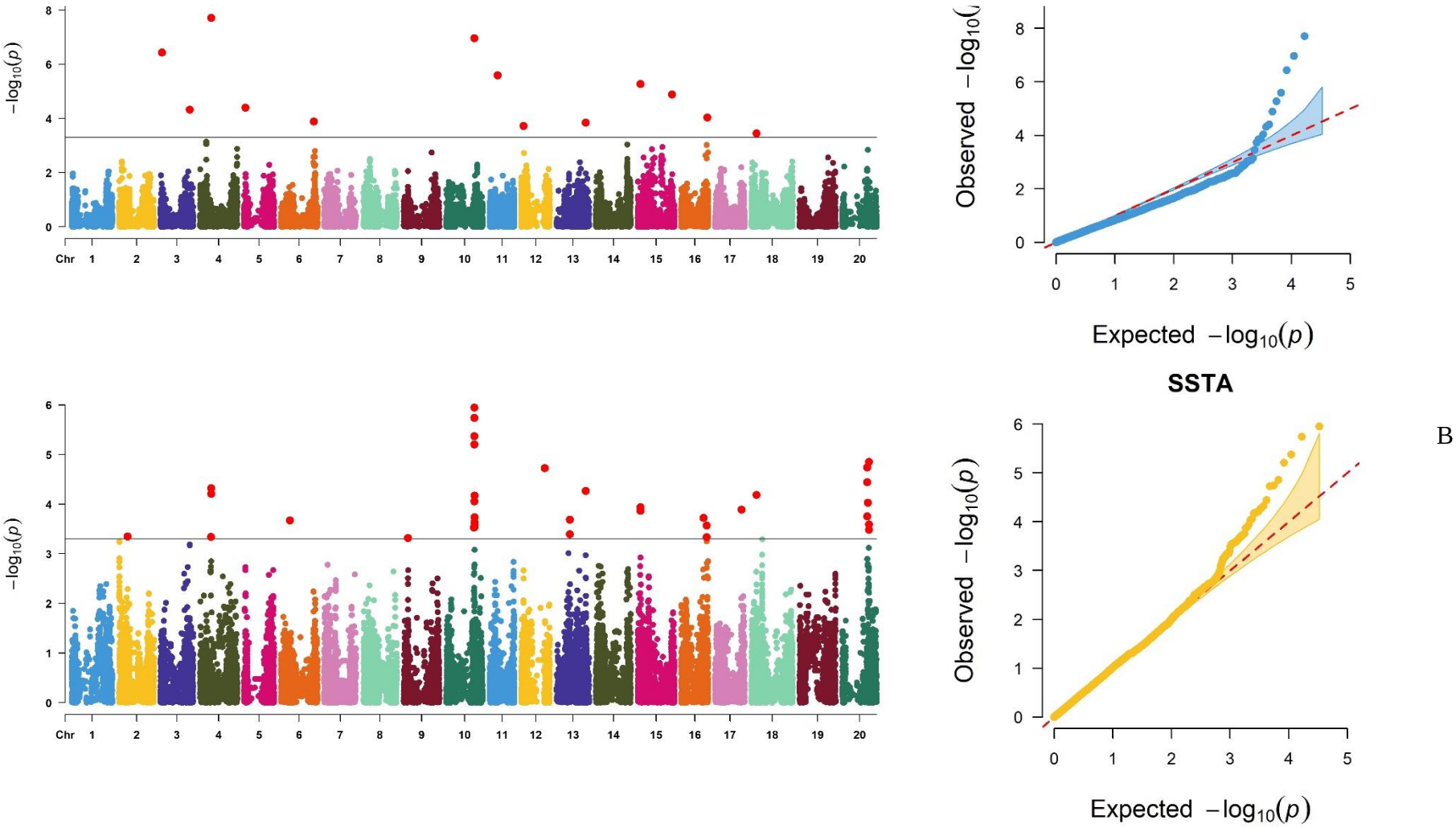


Figure 4.5 Manhattan and quantile-quantile (QQ) plots of FarmCPU from the genome-wide association study (GWAS) displaying significantly associated SNPs for target spot against two *C. cassiicola* isolates: (**A**) LIM01 and (**B**)SSTA. The cut-off threshold is −log10 (p) ≥ 3, represented by a black horizontal line.

A



B

Table 4.1 Analysis of variance for soybean accessions with two *Corynespora cassiicola* isolates.

Source of variance	Degree of freedom	Den degree of freedom	F value	*P* Value
Genotype	245	980	63.08	<0.0001
Isolate	1	2	9.8	0.0887
Genotype × Isolate	245	980	3.19	<0.0001

Table 4.2 The SNPs (single-nucleotide polymorphisms) associated with target spot resistance to each of two *C. cassiicola* isolates

Isolate	SNP Marker	Chr[a]	Position (bp)	−log10 (*P*-Value)	FDR[b]	Allelic effect	MAF[c]
LIM01	ss715637816	20	37,162,511	10.039	3.05E-06	6.454	0.264
	ss715587169	04	15,841,584	7.713	0.0003231	5.720	0.246
	ss715606800	10	39,534,982	6.964	0.0012097	4.825	0.329
	ss715584729	03	2,091,140	6.436	0.0030556	-7.915	0.132
	ss715609438	11	11,087,536	5.593	0.0170229	-4.842	0.248
	ss715621861	15	4,137,385	5.273	0.0296592	-5.218	0.236
	ss715622335	15	48,361,099	4.884	0.0622925	-4.857	0.173
	ss715590348	05	1,952,710	4.400	0.1661507	-4.012	0.421
	ss715586174	03	40,847,194	4.321	0.1769222	-6.190	0.108
	ss715624869	16	36,571,566	4.035	0.3079899	-3.614	0.374
	ss715594534	06	46,295,298	3.885	0.3956736	-3.575	0.423
	ss715616244	13	40,878,186	3.845	0.3973279	4.061	0.272
	ss715612322	12	3,373,730	3.725	0.483678	3.445	0.262
	ss715632608	18	7,161,394	3.449	0.8470135	5.940	0.057
SSTA	ss715606800	10	39,534,982	5.943	0.0306067	6.255	0.329
	ss715606796	10	39,521,125	5.737	0.0306067	6.258	0.301
	ss715606801	10	39,543,852	5.371	0.0473767	5.847	0.333
	ss715606793	10	39,500,426	5.205	0.0520579	6.288	0.266
	ss715637816	20	37,162,511	4.850	0.0902092	5.846	0.264
	ss715637481	20	34,776,314	4.738	0.0902092	-5.631	0.419
	ss715612259	12	33,175,894	4.723	0.0902092	-8.969	0.091
	ss715637482	20	34,795,864	4.444	0.1501294	-5.463	0.415
	ss715587167	04	15,781,556	4.322	0.1730753	-5.389	0.423
	ss715616244	13	40,878,186	4.268	0.1730753	6.350	0.272
	ss715587169	04	15,841,584	4.207	0.1730753	5.502	0.246
	ss715632608	18	7,161,394	4.185	0.1730753	9.985	0.057
	ss715606833	10	39,716,038	4.171	0.1730753	5.775	0.248
	ss715606779	10	39,456,010	4.055	0.2084827	5.411	0.289
	ss715637615	20	35,651,832	4.028	0.2084827	-4.514	0.498
	ss715621861	15	4,137,385	3.937	0.2413941	-6.513	0.236
	ss715627202	17	37,097,907	3.889	0.2518998	-5.520	0.346
	ss715621851	15	4,119,316	3.867	0.2518998	6.639	0.230
	ss715637471	20	34,699,114	3.755	0.3014464	5.219	0.396
	ss715606822	10	39,670,594	3.734	0.3014464	-5.779	0.189
	ss715624393	16	31,349,749	3.722	0.3014464	-5.101	0.350
	ss715615362	13	18,693,392	3.686	0.3078804	-6.957	0.140
	ss715592934	06	12,584,572	3.672	0.3078804	5.686	0.167
	ss715606823	10	39,682,612	3.636	0.3078804	4.904	0.254

ss715606832	10	39,714,825	3.610	0.3078804	4.695	0.321
ss715637820	20	37,173,335	3.594	0.3078804	4.947	0.335
ss715637817	20	37,167,430	3.591	0.3078804	4.950	0.327
ss715624785	16	35,789,070	3.568	0.3078804	5.876	0.283
ss715606829	10	39,709,662	3.562	0.3078804	-5.948	0.167
ss715606817	10	39,666,999	3.547	0.3078804	4.861	0.252
ss715606824	10	39,687,502	3.544	0.3078804	5.279	0.323
ss715606653	10	38,753,114	3.528	0.3091247	-5.303	0.222

[a] Chromosome
[b] False discovery rate (FDR) adjusted *p* values
[c] Minor allele frequency

Appendix 1

Appendix 1A: Common DEGs in all four genotypes

Upregulated DEGs after *C. cassiicola* infection common in all genotypes				
Biological process				
Pathway (GO term #)	Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment
Secondary metabolic proc. (GO:0019748)	2.03E-23	41	232	8.796392821
RNA modification (GO:0009451)	3.07E-23	65	657	4.924438644
S-glycoside metabolic proc. (GO:0016143)	9.86E-20	23	70	16.35454777
Glycosinolate metabolic proc. (GO:0019757)	9.86E-20	23	70	16.35454777
Glucosinolate metabolic proc. (GO:0019760)	9.86E-20	23	70	16.35454777
Recognition of pollen (GO:0048544)	2.66E-12	24	161	7.419832636
Cell recognition (GO:0008037)	2.66E-12	24	161	7.419832636
Pollination (GO:0009856)	2.66E-12	26	193	6.705401428
Multi-multicellular organism proc. (GO:0044706)	2.66E-12	26	193	6.705401428
Pollen-pistil interaction (GO:0009875)	4.17E-12	24	165	7.239957905
Glycosyl compound metabolic proc. (GO:1901657)	1.26E-11	24	174	6.865477324
Indole glucosinolate metabolic proc. (GO:0042343)	7.40E-11	12	33	18.09989476
Multi-organism proc. (GO:0051704)	1.58E-10	27	252	5.333004707
S-glycoside catabolic proc. (GO:0016145)	6.79E-10	11	31	17.66199408
Glycosinolate catabolic proc. (GO:0019759)	6.79E-10	11	31	17.66199408
Glucosinolate catabolic proc. (GO:0019762)	6.79E-10	11	31	17.66199408
Response to biotic stimulus (GO:0009607)	3.92E-09	33	428	3.837769742
Polyketide metabolic proc. (GO:0030638)	1.56E-08	9	23	19.47706067
Polyketide biosynthetic proc. (GO:0030639)	1.56E-08	9	23	19.47706067
Sulfur compound metabolic proc. (GO:0006790)	2.73E-08	36	541	3.312180372
Flavonoid biosynthetic proc. (GO:0009813)	3.48E-08	12	55	10.85993686
Flavonoid metabolic proc. (GO:0009812)	4.15E-08	12	56	10.66600941
Sulfur compound catabolic proc. (GO:0044273)	4.38E-08	11	45	12.16715148
Indole-containing compound metabolic proc. (GO:0042430)	7.24E-08	13	72	8.987100524

Reg. of protein serine/threonine phosphatase activity (GO:0080163)	1.35E-07	11	50	10.95043633
Secondary metabolite biosynthetic proc. (GO:0044550)	1.42E-07	14	91	7.657647784
Defense response (GO:0006952)	2.03E-06	41	796	2.563772782
Neg. reg. of phosphoprotein phosphatase activity (GO:0032515)	2.05E-06	11	65	8.423412563
Neg. reg. of protein dephosphorylation (GO:0035308)	2.05E-06	11	65	8.423412563
Response to salt stress (GO:0009651)	1.60E-05	15	154	4.848186097
Carbohydrate derivative catabolic proc. (GO:1901136)	5.60E-05	13	129	5.016056107
Neg. reg. of phosphorus metabolic proc. (GO:0010563)	7.37E-05	11	94	5.824700176
Thiamine biosynthetic proc. (GO:0009228)	8.52E-05	6	22	13.57492107
Thiamine metabolic proc. (GO:0006772)	0.000145	6	24	12.44367765
Response to external biotic stimulus (GO:0043207)	0.000153	22	368	2.975662047
Response to other organism (GO:0051707)	0.000153	22	368	2.975662047
Thiamine-containing compound metabolic proc. (GO:0042723)	0.00017	6	25	11.94593054
Primary alcohol biosynthetic proc. (GO:0034309)	0.000213	6	26	11.48647168
Biological proc. involved in interspecies interaction between organisms (GO:0044419)	0.00025	22	382	2.866606369
Response to external stimulus (GO:0009605)	0.000279	29	595	2.425994298
Defense response to other organism (GO:0098542)	0.000432	20	340	2.927924153
Immune system proc. (GO:0002376)	0.000433	20	341	2.919337865
Immune response (GO:0006955)	0.000433	20	341	2.919337865
Rhythmic proc. (GO:0048511)	0.000435	9	77	5.817823316
Abscisic acid-activated signaling pathway (GO:0009738)	0.000743	12	146	4.091072104
Cellular response to abscisic acid stimulus (GO:0071215)	0.000915	12	150	3.981976848
Cellular response to alcohol (GO:0097306)	0.000915	12	150	3.981976848
Reg. of protein dephosphorylation (GO:0035304)	0.001153	11	131	4.179555852
Neg. reg. of hydrolase activity (GO:0051346)	0.001685	12	161	3.709916318
Water-soluble vitamin biosynthetic proc. (GO:0042364)	0.001952	9	95	4.715498899
Cellular response to chemical stimulus	0.001976	56	1657	1.68218696

(GO:0070887)				
Hydrogen peroxide metabolic proc. (GO:0042743)	0.00306	13	199	3.25161426
Defense response to bacterium (GO:0042742)	0.003563	11	151	3.625972295
Reproductive proc. (GO:0022414)	0.003919	30	741	2.015170469
Olefinic compound biosynthetic proc. (GO:0120255)	0.004108	5	30	8.295785099
Circadian rhythm (GO:0007623)	0.004355	7	65	5.360353449
Reproduction (GO:0000003)	0.004379	30	748	1.996311922
Cinnamic acid biosynthetic proc. (GO:0009800)	0.005347	3	8	18.66551647
Response to bacterium (GO:0009617)	0.005787	11	162	3.3797643
Phenylpropanoid metabolic proc. (GO:0009698)	0.005841	9	113	3.964357481
Amino acid transport (GO:0006865)	0.006821	13	220	2.941232899
Vitamin biosynthetic proc. (GO:0009110)	0.006887	9	116	3.861830995
Cellular response to reactive nitrogen species (GO:1902170)	0.007129	4	20	9.954942119
Pos. reg. of circadian rhythm (GO:0042753)	0.007192	3	9	16.5915702
Reactive oxygen species metabolic proc. (GO:0072593)	0.009655	13	230	2.813353208
Reg. of circadian rhythm (GO:0042752)	0.010483	5	38	6.549304026
Multicellular organismal proc. (GO:0032501)	0.012964	39	1146	1.693903764
Oxoacid metabolic proc. (GO:0043436)	0.013142	49	1537	1.586832023
Hormone-mediated signaling pathway (GO:0009755)	0.014391	36	1040	1.722970751
Cellular response to inorganic substance (GO:0071241)	0.015847	5	42	5.925560785
Vitamin metabolic proc. (GO:0006766)	0.016084	9	133	3.368213499
Organic acid metabolic proc. (GO:0006082)	0.01762	50	1603	1.552548677
Cellular response to hormone stimulus (GO:0032870)	0.021398	36	1068	1.677799234
Oxazole or thiazole biosynthetic proc. (GO:0018131)	0.022414	2	4	24.8873553
Thiazole biosynthetic proc. (GO:0052837)	0.022414	2	4	24.8873553
Thiazole metabolic proc. (GO:0052838)	0.022414	2	4	24.8873553
Histone H3-K36 demethylation (GO:0070544)	0.022414	2	4	24.8873553
Response to abscisic acid (GO:0009737)	0.022414	12	226	2.642904987
Response to alcohol (GO:0097305)	0.022414	12	226	2.642904987

Organic acid transport (GO:0015849)	0.026624	14	295	2.362189655
Cellular response to lipid (GO:0071396)	0.027945	14	297	2.346282654
Neg. reg. of cellular protein metabolic proc. (GO:0032269)	0.029073	14	299	2.330588456
Pos. reg. of transcription from RNA polymerase II promoter in response to heat s (GO:0061408)	0.029119	5	50	4.97747106
Aromatic amino acid family catabolic proc. (GO:0009074)	0.032684	4	32	6.221838825
Phenylpropanoid biosynthetic proc. (GO:0009699)	0.036461	5	53	4.695727415
Erythrose 4-phosphate/phosphoenolpyruvate family amino acid catabolic proc. (GO:1902222)	0.037997	3	17	8.783772458
L-phenylalanine catabolic proc. (GO:0006559)	0.037997	3	17	8.783772458
Reg. of DNA-templated transcription in response to stress (GO:0043620)	0.038064	5	54	4.6087695
Neg. reg. of catalytic activity (GO:0043086)	0.040455	16	381	2.09027656
Neg. reg. of molecular function (GO:0044092)	0.040455	16	381	2.09027656
Entrainment of circadian clock (GO:0009649)	0.042898	3	18	8.295785099
Cellular response to nitrate (GO:0071249)	0.045461	2	6	16.5915702
Ribulose bisphosphate carboxylase complex assembly (GO:0110102)	0.045461	2	6	16.5915702
Response to hormone (GO:0009725)	0.045461	39	1258	1.543095162
L-ascorbic acid metabolic proc. (GO:0019852)	0.047614	3	19	7.859164831
Cellular oxidant detoxification (GO:0098869)	0.047614	14	322	2.164117852
Protein autophosphorylation (GO:0046777)	0.049611	17	428	1.977032898
Cellular response to oxygen-containing compound (GO:1901701)	0.049611	17	428	1.977032898
Response to abiotic stimulus (GO:0009628)	0.049611	27	797	1.686219807
Cellular components				
Apoplast (GO:0048046)	0.0008	20	311	3.200946019
Extracellular region (GO:0005576)	0.001279	52	1395	1.855401399
Molecular functions				
Oxidoreductase activity, acting on paired donors, with incorporation or reductio (GO:0016709)	2.23E-11	32	307	5.18824345
Beta-glucosidase activity (GO:0008422)	2.50E-10	23	175	6.541819107

Protein serine kinase activity (GO:0106310)	2.50E-10	31	328	4.70431716
Glucosidase activity (GO:0015926)	3.32E-09	23	202	5.667417543
Heme binding (GO:0020037)	7.72E-09	46	766	2.989081837
Naringenin-chalcone synthase activity (GO:0016210)	1.64E-08	9	22	20.36238161
Tetrapyrrole binding (GO:0046906)	5.14E-08	46	821	2.78883884
Protein serine/threonine kinase activity (GO:0004674)	2.06E-07	74	1764	2.088054753
Abscisic acid binding (GO:0010427)	3.13E-07	11	52	10.5292657
Monooxygenase activity (GO:0004497)	6.26E-07	34	553	3.060289621
Isoprenoid binding (GO:0019840)	7.14E-07	11	57	9.605645905
Hormone binding (GO:0042562)	7.93E-07	11	58	9.44003132
Protein phosphatase inhibitor activity (GO:0004864)	2.35E-06	11	65	8.423412563
Iron ion binding (GO:0005506)	2.35E-06	35	621	2.805337956
Phosphatase inhibitor activity (GO:0019212)	3.04E-06	11	67	8.171967411
Alcohol binding (GO:0043178)	8.15E-06	11	74	7.398943467
Polysaccharide binding (GO:0030247)	1.16E-05	14	130	5.360353449
Oxidoreductase activity, acting on paired donors, with incorporation or reductio (GO:0016705)	2.72E-05	35	697	2.499447447
Sequence-specific DNA binding (GO:0043565)	0.000126	69	1945	1.765786648
Ubiquitin-like protein ligase binding (GO:0044389)	0.000128	14	161	4.328235704
Monocarboxylic acid binding (GO:0033293)	0.000175	11	103	5.315745792
Ubiquitin protein ligase binding (GO:0031625)	0.000368	13	155	4.174653147
O-methyltransferase activity (GO:0008171)	0.000425	10	94	5.295181978
Acyltransferase activity, transferring groups other than amino-acyl groups (GO:0016747)	0.000537	33	745	2.204785839
Dioxygenase activity (GO:0051213)	0.00083	22	414	2.64503293
Carbohydrate binding (GO:0030246)	0.0013	22	428	2.558513162
Organic acid binding (GO:0043177)	0.001342	11	132	4.14789255
Phosphatase regulator activity (GO:0019208)	0.001342	11	133	4.116705388
Symporter activity (GO:0015293)	0.001342	20	373	2.668885287
FAD binding (GO:0071949)	0.001695	11	137	3.99650961
Acyltransferase activity (GO:0016746)	0.002929	33	828	1.983774698

Carboxylic acid binding (GO:0031406)	0.007204	11	163	3.359029549
Molecular transducer activity (GO:0060089)	0.00871	23	534	2.143854576
Signaling receptor activity (GO:0038023)	0.009724	22	506	2.164117852
Hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553)	0.010372	30	788	1.894976292
Phenylalanine ammonia-lyase activity (GO:0045548)	0.011399	3	10	14.93241318
Secondary active transmembrane transporter activity (GO:0015291)	0.013694	30	805	1.854958159
Coproporphyrinogen oxidase activity (GO:0004109)	0.014783	2	3	33.1831404
Peroxidase activity (GO:0004601)	0.017511	14	274	2.543233388
Oxidoreductase activity, acting on peroxide as acceptor (GO:0016684)	0.017675	14	275	2.533985267
NADPH dehydrogenase (quinone) activity (GO:0008753)	0.019583	4	26	7.657647784
Ammonia-lyase activity (GO:0016841)	0.027685	3	14	10.66600941
Alditol:NADP+ 1-oxidoreductase activity (GO:0004032)	0.027685	4	29	6.865477324
Alcohol dehydrogenase (NADP+) activity (GO:0008106)	0.027685	4	29	6.865477324
Aldo-keto reductase (NADP) activity (GO:0004033)	0.030758	4	30	6.63662808
Lipid binding (GO:0008289)	0.034519	20	503	1.979113741
TRNA (guanine-N2-)-methyltransferase activity (GO:0004809)	0.038492	2	5	19.90988424
Zinc ion binding (GO:0008270)	0.038492	39	1229	1.579506683
Oxidoreductase activity, acting on CH-OH group of donors (GO:0016614)	0.045019	17	413	2.048837966
Downregulated DEGs in after C. cassiicola infection common in all genotypes				
Biological process				
	Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment
Response to heat (GO:0009408)	8.26E-07	17	192	6.807721515
Small molecule biosynthetic proc. (GO:0044283)	1.08E-06	39	1000	2.9986011
Organic hydroxy compound biosynthetic proc. (GO:1901617)	7.1E-06	18	269	5.144868917
Organic hydroxy compound metabolic proc. (GO:1901615)	7.1E-06	23	441	4.009990424
Sterol biosynthetic proc. (GO:0016126)	7.15E-06	10	72	10.67877885
Response to temperature stimulus	7.25E-06	17	250	5.228330124

(GO:0009266)				
Sterol metabolic proc. (GO:0016125)	1.74E-05	13	152	6.575879606
Steroid biosynthetic proc. (GO:0006694)	5.29E-05	12	142	6.49751051
Steroid metabolic proc. (GO:0008202)	6.46E-05	13	174	5.744446553
Fluid transport (GO:0042044)	8.19E-05	8	58	10.6051321
Water transport (GO:0006833)	8.19E-05	8	58	10.6051321
Response to abiotic stimulus (GO:0009628)	8.54E-05	29	797	2.797652476
Inositol biosynthetic proc. (GO:0006021)	0.000134	4	8	38.44360385
Cellular response to heat (GO:0034605)	0.000134	10	110	6.989746155
Carboxylic acid biosynthetic proc. (GO:0046394)	0.000134	26	687	2.909850655
Lipid biosynthetic proc. (GO:0008610)	0.000146	30	877	2.63012113
Inositol metabolic proc. (GO:0006020)	0.000242	6	33	13.97949231
Lipid metabolic proc. (GO:0006629)	0.00042	43	1594	2.074121663
Response to inorganic substance (GO:0010035)	0.000661	14	266	4.046695142
Protein complex oligomerization (GO:0051259)	0.001065	7	65	8.28016083
Circadian rhythm (GO:0007623)	0.001065	7	65	8.28016083
Oxoacid metabolic proc. (GO:0043436)	0.001595	40	1537	2.00096832
Response to hydrogen peroxide (GO:0042542)	0.002018	7	73	7.372745944
Microtubule-based proc. (GO:0007017)	0.002018	18	465	2.976279008
Carboxylic acid metabolic proc. (GO:0019752)	0.002018	38	1454	2.009431838
Rhythmic proc. (GO:0048511)	0.00262	7	77	6.989746155
Branched-chain amino acid biosynthetic proc. (GO:0009082)	0.002824	5	34	11.30694231
Organic acid metabolic proc. (GO:0006082)	0.003113	40	1603	1.91858285
Alcohol biosynthetic proc. (GO:0046165)	0.003122	8	108	5.695348719
Alcohol metabolic proc. (GO:0006066)	0.003137	11	206	4.105627596
Microtubule-based movement (GO:0007018)	0.003437	10	174	4.41880504
Sphingolipid biosynthetic proc. (GO:0030148)	0.004191	6	60	7.68872077
Polyol metabolic proc. (GO:0019751)	0.00482	8	117	5.257244971
Response to salt stress (GO:0009651)	0.006089	9	154	4.493408242
Response to reactive oxygen species (GO:0000302)	0.00741	8	126	4.881727473
Movement of cell or subcellular component (GO:0006928)	0.00741	11	232	3.645514158
Cellular lipid metabolic proc.	0.007549	30	1141	2.02157426

(GO:0044255)				
Cellular amino acid biosynthetic proc. (GO:0008652)	0.007633	14	358	3.006762312
Protein folding (GO:0006457)	0.008247	16	450	2.733767385
Sphingolipid metabolic proc. (GO:0006665)	0.008469	7	100	5.382104539
Fatty acid metabolic proc. (GO:0006631)	0.009069	13	324	3.084980556
Polyol biosynthetic proc. (GO:0046173)	0.010256	5	49	7.845633439
Response to oxygen-containing compound (GO:1901700)	0.013315	22	769	2.199634031
Response to water (GO:0009415)	0.015118	5	54	7.119185898
Water-soluble vitamin metabolic proc. (GO:0006767)	0.015118	7	112	4.805450481
Reg. of anaphase-promoting complex-dependent catabolic proc. (GO:1905784)	0.015163	3	14	16.47583022
Pos. reg. of ubiquitin protein ligase activity (GO:1904668)	0.01805	3	15	15.37744154
Asparagine metabolic proc. (GO:0006528)	0.01805	3	15	15.37744154
Cutin biosynthetic proc. (GO:0010143)	0.021533	4	35	8.787109452
Alpha-amino acid biosynthetic proc. (GO:1901607)	0.022067	11	277	3.053282616
Calcium import into the mitochondrion (GO:0036444)	0.023319	3	17	13.56833077
Mitochondrial calcium ion homeostasis (GO:0051560)	0.023319	3	17	13.56833077
Cytosolic calcium ion transport (GO:0060401)	0.023319	3	17	13.56833077
Calcium ion import (GO:0070509)	0.023319	3	17	13.56833077
Mitochondrial calcium ion transmembrane transport (GO:0006851)	0.023319	3	17	13.56833077
Response to acid chemical (GO:0001101)	0.024425	5	63	6.102159341
Water-soluble vitamin biosynthetic proc. (GO:0042364)	0.026414	6	95	4.856034171
Vitamin metabolic proc. (GO:0006766)	0.031139	7	133	4.046695142
Response to herbicide (GO:0009635)	0.039838	2	6	25.62906923
Spindle pole body organization (GO:0051300)	0.039838	3	21	10.98388681
Microtubule organizing center organization (GO:0031023)	0.044417	4	45	6.834418462
Pos. reg. of protein ubiquitination (GO:0031398)	0.049727	3	23	10.02876622
Pos. reg. of ubiquitin-protein transferase activity (GO:0051443)	0.049727	3	23	10.02876622
Cellular components				

Amyloplast (GO:0009501)	0.017141	3	17	13.56833077
Ion channel complex (GO:0034702)	0.017141	3	17	13.56833077
Cation channel complex (GO:0034703)	0.017141	3	17	13.56833077
Calcium channel complex (GO:0034704)	0.017141	3	17	13.56833077
Uniplex complex (GO:1990246)	0.017141	3	17	13.56833077
Nuclear microtubule (GO:0005880)	0.017141	4	34	9.045553847
Supramolecular polymer (GO:0099081)	0.017141	14	407	2.644768815
Supramolecular fiber (GO:0099512)	0.017141	14	407	2.644768815
Polymeric cytoskeletal fiber (GO:0099513)	0.017141	14	407	2.644768815
Cell wall (GO:0005618)	0.017141	23	749	2.361022399
External encapsulating structure (GO:0030312)	0.017141	23	776	2.278873424
Kinesin complex (GO:0005871)	0.022405	7	133	4.046695142
Microtubule (GO:0005874)	0.024697	13	394	2.536887564
Acetolactate synthase complex (GO:0005948)	0.02535	2	6	25.62906923
Perinuclear region of cytoplasm (GO:0048471)	0.027126	3	22	10.48461923
Vacuolar lumen (GO:0005775)	0.030786	2	7	21.96777363
NatA complex (GO:0031415)	0.03679	2	8	19.22180193
Microtubule associated complex (GO:0005875)	0.03679	8	198	3.106553847
Supramolecular complex (GO:0099080)	0.03679	15	534	2.159753025
Cytoskeleton (GO:0005856)	0.047467	17	660	1.980428077
Molecular functions				
Ribulose-1,5-bisphosphate carboxylase/oxygenase activator activity (GO:0046863)	3.94E-06	5	7	54.91943407
Inositol-3-phosphate synthase activity (GO:0004512)	4.94E-06	4	4	76.8872077
C-4 methylsterol oxidase activity (GO:0000254)	4.94E-06	6	15	30.75488308
Sphingosine hydroxylase activity (GO:0000170)	1.14E-05	5	10	38.44360385
Channel activity (GO:0015267)	3.42E-05	18	327	4.23232336
Passive transmembrane transporter activity (GO:0022803)	3.42E-05	18	327	4.23232336
Water transmembrane transporter activity (GO:0005372)	5.56E-05	8	58	10.6051321
Water channel activity (GO:0015250)	5.56E-05	8	58	10.6051321
Enzyme activator activity (GO:0008047)	0.000159	5	18	21.3575577
Protein self-association (GO:0043621)	0.00048	7	57	9.442288665
Hydrolase activity, hydrolyzing O-glycosyl	0.000906	26	788	2.536887564

compounds (GO:0004553)				
Iron ion binding (GO:0005506)	0.001254	22	621	2.723862431
Microtubule motor activity (GO:0003777)	0.002683	10	165	4.65983077
Intramolecular lyase activity (GO:0016872)	0.0028	4	18	17.08604616
Unfolded protein binding (GO:0051082)	0.003038	14	318	3.384971408
Molecular function activator activity (GO:0140677)	0.003233	5	36	10.67877885
Starch synthase activity (GO:0009011)	0.003588	3	8	28.83270289
Cytoskeletal motor activity (GO:0003774)	0.00402	11	216	3.915552244
Monooxygenase activity (GO:0004497)	0.00402	19	553	2.641694297
Calcium activated cation channel activity (GO:0005227)	0.004132	4	22	13.97949231
Ion gated channel activity (GO:0022839)	0.004132	4	22	13.97949231
Polygalacturonase activity (GO:0004650)	0.004132	8	119	5.168887913
Glucosyltransferase activity (GO:0046527)	0.004949	15	390	2.957200296
Chlorophyllide a oxygenase [overall] activity (GO:0010277)	0.007274	3	11	20.96923846
Hydrolase activity, acting on glycosyl bonds (GO:0016798)	0.007986	27	1004	2.067683872
Oxidoreductase activity, acting on single donors with incorporation of molecular (GO:0016703)	0.008866	3	12	19.22180193
Lactaldehyde dehydrogenase activity (GO:0008911)	0.009713	2	3	51.25813847
Glycogen (starch) synthase activity (GO:0004373)	0.010599	3	13	17.74320178
Cation channel activity (GO:0005261)	0.011019	7	110	4.892822308
Ubiquitin ligase activator activity (GO:1990757)	0.012469	3	14	16.47583022
Ubiquitin-protein transferase activator activity (GO:0097027)	0.014938	3	15	15.37744154
3-oxo-arachidoyl-CoA synthase activity (GO:0102336)	0.016088	4	35	8.787109452
3-oxo-cerotoyl-CoA synthase activity (GO:0102337)	0.016088	4	35	8.787109452
3-oxo-lignoceronyl-CoA synthase activity (GO:0102338)	0.016088	4	35	8.787109452
Very-long-chain 3-ketoacyl-CoA synthase activity (GO:0102756)	0.016088	4	35	8.787109452
Anaphase-promoting complex binding (GO:0010997)	0.017405	3	17	13.56833077
Xyloglucan:xyloglucosyl transferase activity (GO:0016762)	0.017405	5	62	6.200581266

Inorganic molecular entity transmembrane transporter activity (GO:0015318)	0.017405	28	1146	1.87857052
2 iron, 2 sulfur cluster binding (GO:0051537)	0.018928	5	64	6.006813102
Microtubule binding (GO:0008017)	0.018928	12	330	2.795898462
GDP-mannose 3,5-epimerase activity (GO:0047918)	0.020456	2	5	30.75488308
Tubulin binding (GO:0015631)	0.030699	12	353	2.61372944
Acetolactate synthase activity (GO:0003984)	0.039406	2	7	21.96777363
Inositol oxygenase activity (GO:0050113)	0.039406	2	7	21.96777363
GTP cyclohydrolase II activity (GO:0003935)	0.042619	2	8	19.22180193
Asparagine synthase (glutamine-hydrolyzing) activity (GO:0004066)	0.042619	2	8	19.22180193
Hydroxymethylglutaryl-CoA reductase (NADPH) activity (GO:0004420)	0.042619	2	8	19.22180193
Prephenate dehydrogenase (NADP+) activity (GO:0004665)	0.042619	2	8	19.22180193
3,4-dihydroxy-2-butanone-4-phosphate synthase activity (GO:0008686)	0.042619	2	8	19.22180193
L-leucine:2-oxoglutarate aminotransferase activity (GO:0050048)	0.042619	2	8	19.22180193
L-leucine transaminase activity (GO:0052654)	0.042619	2	8	19.22180193
L-valine transaminase activity (GO:0052655)	0.042619	2	8	19.22180193
L-isoleucine transaminase activity (GO:0052656)	0.042619	2	8	19.22180193
Long-chain fatty acid-CoA ligase activity (GO:0004467)	0.046783	3	27	8.543023078
Oxidoreductase activity, acting on paired donors, with incorporation or reductio (GO:0016705)	0.046783	18	697	1.985609381
Electron transfer activity (GO:0009055)	0.049036	12	389	2.371841883

Appendix 1B: Unique DEGs

Unique DEGs upregulated in Bedford				
Biological process				
Pathway (GO term #)	Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment
Plant-type secondary cell wall biogenesis (GO:0009834)	0.000127483	13	92	6.522286648
Plant-type cell wall biogenesis	0.00121978	17	199	3.943121885

(GO:0009832)				
Cell wall biogenesis (GO:0042546)	0.00386574	20	296	3.118764925
Defense response (GO:0006952)	0.011665385	36	796	2.087535116
Response to metal ion (GO:0010038)	0.012226608	10	99	4.66239605
Plant-type cell wall organization or biogenesis (GO:0071669)	0.012226608	19	317	2.766551094
Response to biotic stimulus (GO:0009607)	0.012226608	23	428	2.480438272
RNA modification (GO:0009451)	0.012226608	31	657	2.177913771
Response to external biotic stimulus (GO:0043207)	0.021835394	20	368	2.508571788
Response to other organism (GO:0051707)	0.021835394	20	368	2.508571788
Cell wall organization or biogenesis (GO:0071554)	0.021835394	34	799	1.964158336
Response to bacterium (GO:0009617)	0.024803034	12	162	3.419090436
Response to cadmium ion (GO:0046686)	0.025262789	5	27	8.547726091
Innate immune response (GO:0045087)	0.026741948	6	43	6.440612217
Biological proc. involved in interspecies interaction between organisms (GO:0044419)	0.026741948	20	382	2.416634602
Defense response to other organism (GO:0098542)	0.03848498	18	340	2.443644047
Immune system proc. (GO:0002376)	0.03848498	18	341	2.436477936
Immune response (GO:0006955)	0.03848498	18	341	2.436477936
Cellular components				
VCB complex (GO:0030891)	0.005308928	3	5	34.65769231
Extracellular region (GO:0005576)	0.005308928	45	1395	1.863316791
Unique DEGs downregulated in Bedford				
Biological process				
Phenylpropanoid catabolic proc.	7.73E-15	21	51	11.82147497
Lignin catabolic proc.	7.73E-15	21	51	11.82147497
Lignin metabolic proc.	2.16E-13	23	74	8.923159677
Cell wall organization or biogenesis	2.16E-13	78	799	2.802659719
Cellular polysaccharide metabolic proc.	5.74E-12	53	453	3.358924297
Polysaccharide metabolic proc.	5.74E-12	73	770	2.721790434
Phenylpropanoid metabolic proc.	2.99E-11	25	113	6.351614237
Cell wall macromolecule metabolic proc.	2.27E-09	34	247	3.951886947
Polysaccharide biosynthetic proc.	3.02E-09	38	306	3.565206737
Cellular carbohydrate metabolic proc.	3.35E-09	64	727	2.527365841
Cell wall polysaccharide metabolic proc.	1.15E-08	31	224	3.973161549
Cell wall biogenesis	5.42E-08	35	296	3.394680312
DNA replication initiation	7.06E-08	15	56	7.689990095

Cellular polysaccharide biosynthetic proc.	0.000000153	32	265	3.466782956
Plant-type cell wall biogenesis	0.000000214	27	199	3.895231164
External encapsulating structure organization	0.000000616	53	633	2.403779947
Plant-type cell wall organization or biogenesis	0.000000796	34	317	3.079230524
Cell wall organization	0.000000796	51	603	2.428149443
Reg. of jasmonic acid mediated signaling pathway	0.000000852	12	41	8.402720884
Hemicellulose metabolic proc.	0.00000113	24	176	3.914904048
Secondary metabolic proc.	0.0000042	27	232	3.34116811
Cell cycle	0.00000541	68	975	2.002289387
Xylan metabolic proc.	0.00000613	15	79	5.451132219
Glucan metabolic proc.	0.00000613	35	365	2.752946226
Cellular glucan metabolic proc.	0.00000656	34	350	2.788903074
Carbohydrate biosynthetic proc.	0.00000732	42	490	2.46079683
Cellulose metabolic proc.	0.0000121	19	133	4.10132805
Cell wall polysaccharide biosynthetic proc.	0.0000434	16	105	4.374749921
Pre-replicative complex assembly	0.0000506	8	24	9.569765451
Pre-replicative complex assembly involved in nuclear cell cycle DNA replication	0.0000506	8	24	9.569765451
Cell wall macromolecule catabolic proc.	0.0000586	11	51	6.192201174
Nuclear DNA replication	0.0000586	11	51	6.192201174
Double-strand break repair via break-induced replication	0.0000624	8	25	9.186974833
Jasmonic acid mediated signaling pathway	0.0000624	12	62	5.556638004
Cellular response to jasmonic acid stimulus	0.0000624	12	62	5.556638004
Cellular component macromolecule biosynthetic proc.	0.000088	16	113	4.065033112
Mitotic cell cycle	0.0000986	41	531	2.21672533
Cytoskeleton-dependent cytokinesis	0.000104806	12	66	5.219872064
Mitotic cytokinesis	0.000104806	12	66	5.219872064
Beta-glucan metabolic proc.	0.000104806	19	157	3.474373444
Plant-type secondary cell wall biogenesis	0.000141134	14	92	4.368805967
Cellulose biosynthetic proc.	0.000157273	14	93	4.321829559
Glucan biosynthetic proc.	0.000192615	21	195	3.091770377
Aromatic compound catabolic proc.	0.000192615	35	436	2.304645349
Response to jasmonic acid	0.000207961	12	71	4.85227544
Cell cycle proc.	0.00032044	48	702	1.96302881
Xylan biosynthetic proc.	0.000337808	11	63	5.012734284
DNA-dependent DNA replication	0.000337808	23	236	2.797939899
Organic cyclic compound catabolic proc.	0.000337808	35	450	2.232945272
Mitotic cell cycle proc.	0.000606247	34	444	2.198459631

Movement of cell or subcellular component	0.00074264	22	232	2.722433275
Cell wall polysaccharide catabolic proc.	0.000819565	8	36	6.379843634
Plant-type primary cell wall biogenesis	0.001341357	10	61	4.706442025
Glycerol-3-phosphate catabolic proc.	0.001494927	5	13	11.04203706
Beta-glucan biosynthetic proc.	0.001683235	14	117	3.435300418
Cellular carbohydrate catabolic proc.	0.001818904	14	118	3.406187703
Carbohydrate catabolic proc.	0.003003512	36	525	1.968637464
Cell cycle DNA replication	0.003873021	11	83	3.804846505
Cytokinesis	0.005063752	12	100	3.445115562
Glycerol-3-phosphate metabolic proc.	0.00571848	5	17	8.443910692
Alditol phosphate metabolic proc.	0.00571848	5	17	8.443910692
Polysaccharide catabolic proc.	0.005930931	24	308	2.237088028
Microtubule-based proc.	0.006000732	32	465	1.975693512
Cellular polysaccharide catabolic proc.	0.006671577	10	75	3.82790618
Double-strand break repair via homologous recombination	0.008386575	13	122	3.059187316
NADH oxidation	0.009514325	4	11	10.43974413
Microtubule-based movement	0.009805707	16	174	2.639935297
Response to wounding	0.010753397	8	53	4.333478695
Cortical cytoskeleton organization	0.010753397	8	53	4.333478695
Borate transport	0.012964904	4	12	9.569765451
Carbohydrate derivative catabolic proc.	0.013069923	13	129	2.893184904
Polyamine metabolic proc.	0.013180058	8	55	4.175897651
Plasmodesmata-mediated intercellular transport	0.017415537	5	22	6.52484008
Reg. of defense response	0.017415537	10	86	3.338290274
Recombinational repair	0.02042557	13	136	2.744270975
Intercellular transport	0.021030633	5	23	6.241151381
Aminoglycan metabolic proc.	0.024706668	6	35	4.921593661
Amine metabolic proc.	0.02568689	16	193	2.380045294
DNA replication, synthesis of RNA primer	0.026551018	3	7	12.30398415
Microtubule cytoskeleton organization	0.033817397	20	275	2.087948826
Arabinan metabolic proc.	0.036006534	4	16	7.177324088
Arabinan catabolic proc.	0.036006534	4	16	7.177324088
Xylan catabolic proc.	0.036006534	4	16	7.177324088
Cytoskeleton organization	0.036006534	30	483	1.783186109
Cellular biogenic amine metabolic proc.	0.041652918	10	98	2.929520036
Cell division	0.043054423	23	343	1.925113167
Lipid metabolic proc.	0.049411024	77	1594	1.38683552
Cellular components				
Extracellular region	6.44E-13	110	1395	2.263815483

Apoplast	1.97E-10	40	311	3.692514
Anchored component of membrane	2.41E-09	40	341	3.367659396
Anchored component of plasma membrane	3.2E-09	35	275	3.653910445
External encapsulating structure	0.000000015	64	776	2.367777019
Cell wall	2.05E-08	62	749	2.376470459
Golgi apparatus	0.000000321	93	1429	1.868414668
MCM complex	0.000000735	9	22	11.74471214
Cell-cell junction	0.000000735	50	613	2.341704433
Plasmodesma	0.000000735	50	613	2.341704433
Symplast	0.000000735	50	613	2.341704433
Anchoring junction	0.000000735	50	613	2.341704433
Cell junction	0.000000735	50	614	2.337890583
Intrinsic component of plasma membrane	0.00000466	45	560	2.306997028
Cytoskeleton	0.0000135	49	660	2.13144776
Plant-type cell wall	0.0000172	24	222	3.103707714
Microtubule	0.0000225	34	394	2.47745197
Supramolecular polymer	0.0000392	34	407	2.398319597
Supramolecular fiber	0.0000392	34	407	2.398319597
Polymeric cytoskeletal fiber	0.0000392	34	407	2.398319597
Trans-Golgi network	0.0000632	26	276	2.704498932
Microtubule cytoskeleton	0.000169726	36	476	2.171291321
Golgi apparatus subcompartment	0.000171148	27	311	2.49244695
Kinesin complex	0.000181269	16	133	3.453749937
Supramolecular complex	0.001565842	36	534	1.935458181
Alpha DNA polymerase:primase complex	0.001581819	4	9	12.75968727
Glycerol-3-phosphate dehydrogenase complex	0.00550201	4	12	9.569765451
Golgi membrane	0.011212516	31	491	1.812603232
Microtubule associated complex	0.014619896	16	198	2.31994314
Nuclear replication fork	0.03684619	7	59	3.406187703
Cytoplasmic microtubule	0.041044896	6	46	3.744690829
Vacuole	0.041044896	44	838	1.507409355
Origin recognition complex	0.041275653	3	11	7.829808096
Nuclear origin of replication recognition complex	0.041275653	3	11	7.829808096
Nuclear pore inner ring	0.041275653	3	11	7.829808096
Nucleotide-excision repair factor 1 complex	0.048036613	2	4	14.35464818
CAF-1 complex	0.048036613	2	4	14.35464818
Molecular functions				
Oxidoreductase activity, acting on diphenols and related substances as donors, o	1.54E-16	27	82	9.453060994
Hydroquinone:oxygen oxidoreductase	2.39E-15	21	51	11.82147497

activity				
Copper ion binding	6.32E-13	32	166	5.534322189
Oxidoreductase activity, acting on metal ions	5.37E-12	24	98	7.030848087
Hydrolase activity, hydrolyzing O-glycosyl compounds	0.000000447	62	788	2.258853266
DNA replication origin binding	0.000000567	12	39	8.833629647
Glycosyltransferase activity	0.000000826	82	1201	1.960168444
Hydrolase activity, acting on glycosyl bonds	0.00000402	70	1004	2.001644168
Hexosyltransferase activity	0.0000142	59	820	2.065668884
UDP-glycosyltransferase activity	0.0000555	46	602	2.193733608
Calcium-dependent phospholipid binding	0.0000881	9	33	7.829808096
O-acetyltransferase activity	0.0000978	18	138	3.744690829
O-acyltransferase activity	0.000114034	24	230	2.995752663
Single-stranded DNA helicase activity	0.000129552	6	13	13.25044447
Jasmonic acid hydrolase	0.0003518	4	5	22.96743708
Glucosidase activity	0.000422238	21	202	2.984629819
Glucosyltransferase activity	0.000947237	31	390	2.282020992
UDP-glucosyltransferase activity	0.001549745	26	309	2.415668949
Cellulose synthase (UDP-forming) activity	0.001559635	10	60	4.784882726
Glycerol-3-phosphate dehydrogenase [NAD+] activity	0.002692137	5	14	10.25332013
Glucan endo-1,3-beta-D-glucosidase activity	0.003453673	12	94	3.665016556
Phospholipase A1 activity	0.005097929	5	16	8.97165511
Microtubule motor activity	0.006677144	16	165	2.783931768
Cytoskeletal motor activity	0.006677144	19	216	2.525354772
Microtubule binding	0.007335576	25	330	2.174946693
Acyltransferase activity	0.007335576	49	828	1.698980098
Aspartic-type endopeptidase activity	0.007665279	18	205	2.520816265
Aspartic-type peptidase activity	0.007665279	18	205	2.520816265
Tubulin binding	0.007665279	26	353	2.114565737
Oxidoreductase activity, acting on CH-OH group of donors	0.007665279	29	413	2.015907008
Acyltransferase activity, transferring groups other than amino-acyl groups	0.007665279	45	745	1.734118572
Single-stranded DNA binding	0.008149364	15	155	2.778319002
L-ascorbate oxidase activity	0.008310789	3	5	17.22557781
Single-stranded 3 -5 DNA helicase activity	0.008310789	3	5	17.22557781
Squalene monooxygenase activity	0.008310789	4	11	10.43974413
Beta-glucosidase activity	0.008963757	16	175	2.624849952
Efflux transmembrane transporter activity	0.011167047	4	12	9.569765451
Borate efflux transmembrane transporter activity	0.011167047	4	12	9.569765451

Sucrose alpha-glucosidase activity	0.012225688	5	21	6.835546751
DNA primase activity	0.014307038	3	6	14.35464818
Beta-fructofuranosidase activity	0.014671166	5	22	6.52484008
Structural constituent of cytoskeleton	0.01524311	8	58	3.959902945
Alpha-glucosidase activity	0.017419981	5	23	6.241151381
Inorganic anion exchanger activity	0.018629865	4	14	8.202656101
Cytoskeletal protein binding	0.022901331	35	581	1.729475684
Xylan 1,4-beta-xylosidase activity	0.028957416	4	16	7.177324088
Xylanase activity	0.028957416	4	16	7.177324088
3 -5 DNA helicase activity	0.028957416	6	38	4.533046793
Acetyltransferase activity	0.028957416	21	296	2.036808187
Polygalacturonate 4-alpha-galacturonosyltransferase activity	0.036671569	6	40	4.306394453
Transmembrane receptor protein serine/threonine kinase activity	0.042305628	23	349	1.892016665
Lipid binding	0.047565549	30	503	1.712284077
Unique DEGs upregulated in Council				
Biological process				
Defense response (GO:0006952)	6.22E-09	57	796	2.795163843
Hormone-mediated signaling pathway (GO:0009755)	6.22E-09	67	1040	2.514704421
Auxin-activated signaling pathway (GO:0009734)	8.02E-09	47	613	2.992835628
Cellular response to auxin stimulus (GO:0071365)	8.02E-09	47	613	2.992835628
Response to auxin (GO:0009733)	8.02E-09	49	652	2.933553184
Cellular response to hormone stimulus (GO:0032870)	8.02E-09	67	1068	2.448775841
Cellular response to organic substance (GO:0071310)	1.78E-08	73	1252	2.275956793
Response to endogenous stimulus (GO:0009719)	4.06E-08	73	1279	2.227910794
Response to hormone (GO:0009725)	0.000000043	72	1258	2.234072883
Response to organic substance (GO:0010033)	0.000000888	80	1573	1.985211335
Cellular response to chemical stimulus (GO:0070887)	0.00000175	82	1657	1.931687306
Response to biotic stimulus (GO:0009607)	0.0000896	30	428	2.736043309
Defense response to bacterium (GO:0042742)	0.000174095	16	151	4.136076066
Response to bacterium (GO:0009617)	0.000407962	16	162	3.855231395
Glucosamine-containing compound	0.001098066	7	32	8.538735161

metabolic proc. (GO:1901071)				
Aminoglycan catabolic proc. (GO:0006026)	0.001185788	6	23	10.18283945
Chitin metabolic proc. (GO:0006030)	0.001185788	6	23	10.18283945
Chitin catabolic proc. (GO:0006032)	0.001185788	6	23	10.18283945
Amino sugar catabolic proc. (GO:0046348)	0.001185788	6	23	10.18283945
Glucosamine-containing compound catabolic proc. (GO:1901072)	0.001185788	6	23	10.18283945
Defense response to other organism (GO:0098542)	0.001591921	23	340	2.640550033
Immune system proc. (GO:0002376)	0.001591921	23	341	2.632806484
Immune response (GO:0006955)	0.001591921	23	341	2.632806484
Biological proc. involved in interspecies interaction between organisms (GO:0044419)	0.003200604	24	382	2.452411594
Response to external biotic stimulus (GO:0043207)	0.004508498	23	368	2.439638617
Response to other organism (GO:0051707)	0.004508498	23	368	2.439638617
Salicylic acid metabolic proc. (GO:0009696)	0.00469998	7	43	6.354407561
Aminoglycan metabolic proc. (GO:0006022)	0.010265381	6	35	6.691580208
Phenol-containing compound metabolic proc. (GO:0018958)	0.010265381	7	49	5.57631684
Recognition of pollen (GO:0048544)	0.011221365	13	161	3.151831257
Cell recognition (GO:0008037)	0.011221365	13	161	3.151831257
Protein autophosphorylation (GO:0046777)	0.013293808	24	428	2.188834647
Pollen-pistil interaction (GO:0009875)	0.013398934	13	165	3.075423227
Response to external stimulus (GO:0009605)	0.015146634	30	595	1.968111826
Purine nucleobase salvage (GO:0043096)	0.015756454	4	15	10.40912477
AMP salvage (GO:0044209)	0.015756454	4	15	10.40912477
Adenine salvage (GO:0006168)	0.015756454	4	15	10.40912477
Pollination (GO:0009856)	0.015917422	14	193	2.83149767
Multi-multicellular organism proc. (GO:0044706)	0.015917422	14	193	2.83149767
Oligopeptide transmembrane transport (GO:0035672)	0.019077241	4	16	9.758554469
Nucleoside salvage (GO:0043174)	0.02178582	4	17	9.184521853
Adenine metabolic proc. (GO:0046083)	0.02178582	4	17	9.184521853
Adenine biosynthetic proc. (GO:0046084)	0.02178582	4	17	9.184521853
Purine ribonucleoside salvage (GO:0006166)	0.02178582	4	17	9.184521853
Amino sugar metabolic proc. (GO:0006040)	0.02178582	7	59	4.631178392
Benzene-containing compound metabolic	0.021858343	8	77	4.055503156

proc. (GO:0042537)				
Induced systemic resistance (GO:0009682)	0.030365683	3	9	13.01140596
Induced systemic resistance, jasmonic acid mediated signaling pathway (GO:0009864)	0.030365683	3	9	13.01140596
Immune effector proc. (GO:0002252)	0.030365683	3	9	13.01140596
Salicylic acid biosynthetic proc. (GO:0009697)	0.030365683	4	19	8.217730079
Reg. of salicylic acid metabolic proc. (GO:0010337)	0.030365683	4	19	8.217730079
Activation of innate immune response (GO:0002218)	0.038885792	3	10	11.71026536
Nitrate metabolic proc. (GO:0042126)	0.038885792	3	10	11.71026536
Nitrate assimilation (GO:0042128)	0.038885792	3	10	11.71026536
Reactive nitrogen species metabolic proc. (GO:2001057)	0.038885792	3	10	11.71026536
AMP biosynthetic proc. (GO:0006167)	0.039369148	4	21	7.435089119
Phenol-containing compound biosynthetic proc. (GO:0046189)	0.046363364	4	22	7.097130523
Hormone metabolic proc. (GO:0042445)	0.049751043	12	179	2.616819076
Molecular functions				
Protein serine/threonine kinase activity (GO:0004674)	2.82E-10	99	1764	2.190695901
Chitinase activity (GO:0004568)	0.0000495	10	45	8.674270639
Oxidoreductase activity, acting on paired donors, with incorporation or reductio (GO:0016705)	0.002872474	38	697	2.128120917
Chitin binding (GO:0008061)	0.005681428	6	26	9.007896433
Protein serine kinase activity (GO:0106310)	0.005681428	22	328	2.61814876
Monooxygenase activity (GO:0004497)	0.011919124	30	553	2.117588673
Iron ion binding (GO:0005506)	0.015704434	32	621	2.011425076
Adenine phosphoribosyltransferase activity (GO:0003999)	0.025999917	4	14	11.15263368
Oligopeptide transmembrane transporter activity (GO:0035673)	0.030233552	10	109	3.581120906
Purine phosphoribosyltransferase activity (GO:0106130)	0.030251514	4	16	9.758554469
ADP binding (GO:0043531)	0.030251514	25	471	2.07187993
Heme binding (GO:0020037)	0.036349267	35	766	1.783547814
Unique DEGs downregulated in Council				
Biological process				
Nucleosome assembly (GO:0006334)	5.08E-20	29	118	12.11410785
Chromatin assembly (GO:0031497)	7.07E-20	29	122	11.71692399
DNA packaging (GO:0006323)	9.96E-19	30	147	10.0595688

Chromatin assembly or disassembly (GO:0006333)	1.46E-18	29	138	10.35844005
Nucleosome organization (GO:0034728)	2.93E-17	29	154	9.282238485
Protein-DNA complex assembly (GO:0065004)	1.28E-15	33	238	6.834589391
Chromatin remodeling (GO:0006338)	2.49E-15	30	197	7.50637875
Protein-DNA complex subunit organization (GO:0071824)	1.31E-13	33	280	5.809400983
Chromatin organization (GO:0006325)	4.25E-13	36	350	5.070022676
Chromosome organization (GO:0051276)	3.18E-10	57	977	2.875780518
Cell cycle (GO:0007049)	8.32E-10	56	975	2.831123773
Cell cycle proc. (GO:0022402)	2.72E-09	45	702	3.159736354
Nuclear division (GO:0000280)	3.02E-09	31	362	4.221128455
Response to water deprivation (GO:0009414)	3.11E-09	13	52	12.32297178
Meiotic nuclear division (GO:0140013)	3.33E-09	19	133	7.041698161
Meiotic cell cycle (GO:0051321)	3.57E-09	23	203	5.584795093
Chromosome segregation (GO:0007059)	4.07E-09	23	205	5.530309287
Meiotic cell cycle proc. (GO:1903046)	4.49E-09	20	154	6.401543782
Response to inorganic substance (GO:0010035)	4.49E-09	26	266	4.818004005
DNA conformation change (GO:0071103)	2.25E-08	37	548	3.328101868
Response to acid chemical (GO:0001101)	2.65E-08	13	63	10.17134179
Organelle fission (GO:0048285)	0.000000205	31	440	3.472837502
Meiotic chromosome segregation (GO:0045132)	0.000000305	12	63	9.388930881
Nuclear chromosome segregation (GO:0098813)	0.00000254	18	182	4.875021804
Reproductive proc. (GO:0022414)	0.00000574	39	741	2.594309849
Reproduction (GO:0000003)	0.00000704	39	748	2.570031548
Mitotic cell cycle (GO:0000278)	0.000037	30	531	2.78485238
Organic acid biosynthetic proc. (GO:0016053)	0.0000701	36	729	2.434167265
Organic acid metabolic proc. (GO:0006082)	0.0000718	62	1603	1.906485965
Fatty acid biosynthetic proc. (GO:0006633)	0.0000775	17	209	4.009387948
Mitotic cell cycle proc. (GO:1903047)	0.0000918	26	444	2.886461859
Monocarboxylic acid metabolic proc. (GO:0032787)	0.000104169	34	684	2.450181524
Carboxylic acid metabolic proc. (GO:0019752)	0.000112397	57	1454	1.932350458
Monocarboxylic acid biosynthetic proc. (GO:0072330)	0.000251088	19	280	3.344806626
Carboxylic acid biosynthetic proc. (GO:0046394)	0.000265726	33	687	2.367732569

Protein-containing complex assembly (GO:0065003)	0.000265726	44	1048	2.069506711
Oxoacid metabolic proc. (GO:0043436)	0.000265726	58	1537	1.860071212
Response to oxygen-containing compound (GO:1901700)	0.000424978	35	769	2.2434539
Cellular protein-containing complex assembly (GO:0034622)	0.000584231	40	948	2.079826461
Mitotic nuclear division (GO:0140014)	0.000670062	17	251	3.338494347
Response to abiotic stimulus (GO:0009628)	0.000828628	35	797	2.164637452
Reg. of cell cycle proc. (GO:0010564)	0.001273307	17	265	3.162121061
Fatty acid metabolic proc. (GO:0006631)	0.001518797	19	324	2.890573628
Response to reactive oxygen species (GO:0000302)	0.001910784	11	126	4.303259987
Sequestering of metal ion (GO:0051238)	0.002514174	6	36	8.215314521
Intracellular sequestering of iron ion (GO:0006880)	0.002514174	6	36	8.215314521
Cellular cation homeostasis (GO:0030003)	0.002514174	14	203	3.399440491
Sexual reproduction (GO:0019953)	0.002542642	7	52	6.635446344
Multi-organism reproductive proc. (GO:0044703)	0.002542642	7	52	6.635446344
Glutamine family amino acid metabolic proc. (GO:0009064)	0.002651402	12	156	3.791683625
Cellular chemical homeostasis (GO:0055082)	0.002751637	15	233	3.173297454
Reg. of cell cycle (GO:0051726)	0.002751637	22	433	2.504437683
Recombinational repair (GO:0000725)	0.003064867	11	136	3.986843812
DNA-dependent DNA replication (GO:0006261)	0.003064867	15	236	3.132958927
Double-strand break repair via break-induced replication (GO:0000727)	0.003422053	5	25	9.858377425
Cellular metal ion homeostasis (GO:0006875)	0.003422053	11	138	3.929063467
Cellular ion homeostasis (GO:0006873)	0.003422053	14	213	3.239842346
DNA replication initiation (GO:0006270)	0.003476489	7	56	6.161485891
Reg. of mitotic cell cycle (GO:0007346)	0.003476489	14	214	3.224702896
Cell cycle phase transition (GO:0044770)	0.003591457	14	215	3.209704278
DUMP metabolic proc. (GO:0046078)	0.003786608	3	6	24.64594356
DUMP biosynthetic proc. (GO:0006226)	0.003786608	3	6	24.64594356
Proline metabolic proc. (GO:0006560)	0.003786608	5	26	9.479209063
Reg. of nuclear division (GO:0051783)	0.004058561	11	143	3.791683625
Nucleosome positioning (GO:0016584)	0.004379754	5	27	9.128127245
Double-strand break repair via homologous recombination (GO:0000724)	0.004590468	10	122	4.040318617
Mitotic cell cycle phase transition	0.005493856	12	174	3.399440491

(GO:0044772)				
Microtubule-based movement (GO:0007018)	0.005493856	12	174	3.399440491
Maintenance of DNA methylation (GO:0010216)	0.005974787	3	7	21.12509448
Protein-containing complex subunit organization (GO:0043933)	0.006039434	44	1235	1.756148205
Movement of cell or subcellular component (GO:0006928)	0.006673583	14	232	2.97451043
Cellular component assembly (GO:0022607)	0.006826747	48	1394	1.697281623
Chromosome condensation (GO:0030261)	0.007492495	6	47	6.292581335
Mitotic sister chromatid segregation (GO:0000070)	0.008134189	10	133	3.706156927
Reg. of mitotic nuclear division (GO:0007088)	0.008134189	10	133	3.706156927
Small molecule biosynthetic proc. (GO:0044283)	0.008252606	37	1000	1.823799824
Neg. reg. of DNA recombination (GO:0045910)	0.008413649	5	32	7.701857363
Neg. reg. of organelle organization (GO:0010639)	0.008413649	8	88	4.481080648
Removal of superoxide radicals (GO:0019430)	0.009001656	4	19	10.37723939
Cellular response to oxygen radical (GO:0071450)	0.009001656	4	19	10.37723939
Cellular response to superoxide (GO:0071451)	0.009001656	4	19	10.37723939
Response to superoxide (GO:0000303)	0.009001656	4	19	10.37723939
Response to oxygen radical (GO:0000305)	0.009001656	4	19	10.37723939
Superoxide metabolic proc. (GO:0006801)	0.009001656	4	19	10.37723939
Cellular transition metal ion homeostasis (GO:0046916)	0.009001656	7	69	5.00062623
Meiosis I cell cycle proc. (GO:0061982)	0.009366299	8	91	4.333352714
Cation homeostasis (GO:0055080)	0.00940297	15	274	2.698460974
Microtubule-based proc. (GO:0007017)	0.010958206	21	465	2.226085225
Metal ion homeostasis (GO:0055065)	0.010981841	11	167	3.246771008
Response to temperature stimulus (GO:0009266)	0.011011301	14	250	2.760345679
Inorganic ion homeostasis (GO:0098771)	0.01124349	15	280	2.64063681
Response to hydrogen peroxide (GO:0042542)	0.011509241	7	73	4.726619313
Cell division (GO:0051301)	0.012018073	17	343	2.443038137
Iron ion transport (GO:0006826)	0.012107778	6	54	5.476876347
Cellular iron ion homeostasis (GO:0006879)	0.012107778	6	54	5.476876347

Chromosome separation (GO:0051304)	0.012107778	7	74	4.662746079
Cellular homeostasis (GO:0019725)	0.012360803	22	506	2.143125527
Neg. reg. of nuclear division (GO:0051784)	0.013095021	6	55	5.377296777
Pyrimidine deoxyribonucleoside monophosphate metabolic proc. (GO:0009176)	0.013490715	3	10	14.78756614
Reg. of DNA recombination (GO:0000018)	0.014591739	5	38	6.485774622
Ion homeostasis (GO:0050801)	0.015369993	15	292	2.532117489
DNA replication (GO:0006260)	0.015857128	16	324	2.434167265
Double-strand break repair (GO:0006302)	0.01651787	11	179	3.029110382
Deoxyribose phosphate biosynthetic proc. (GO:0046385)	0.016698471	3	11	13.44324194
Pyrimidine deoxyribonucleotide biosynthetic proc. (GO:0009221)	0.016698471	3	11	13.44324194
2 -deoxyribonucleotide biosynthetic proc. (GO:0009265)	0.016698471	3	11	13.44324194
Maintenance of location in cell (GO:0051651)	0.016698471	7	80	4.313040123
Sister chromatid segregation (GO:0000819)	0.016698471	10	153	3.221691969
Pre-replicative complex assembly (GO:0036388)	0.017161018	4	24	8.215314521
Pre-replicative complex assembly involved in nuclear cell cycle DNA replication (GO:0006267)	0.017161018	4	24	8.215314521
Maintenance of location (GO:0051235)	0.017161018	8	104	3.791683625
Disaccharide metabolic proc. (GO:0005984)	0.01809692	8	105	3.755572352
Reg. of organelle organization (GO:0033043)	0.018102053	16	332	2.375512633
Post-embryonic development (GO:0009791)	0.01919391	19	431	2.172960221
Spindle checkpoint signaling (GO:0031577)	0.020071249	4	26	7.58336725
Neg. reg. of sister chromatid segregation (GO:0033046)	0.020071249	4	26	7.58336725
Reg. of mitotic sister chromatid segregation (GO:0033047)	0.020071249	4	26	7.58336725
Neg. reg. of mitotic metaphase/anaphase transition (GO:0045841)	0.020071249	4	26	7.58336725
Meiosis II cell cycle proc. (GO:0061983)	0.020071249	4	26	7.58336725
Mitotic spindle checkpoint signaling (GO:0071174)	0.020071249	4	26	7.58336725
Neg. reg. of chromosome separation (GO:1905819)	0.020071249	4	26	7.58336725
Meiosis II (GO:0007135)	0.020071249	4	26	7.58336725
Neg. reg. of cell cycle (GO:0045786)	0.020071249	9	133	3.335541234
Cell cycle checkpoint signaling	0.02098853	7	86	4.012130347

(GO:0000075)				
Disaccharide biosynthetic proc. (GO:0046351)	0.02301803	6	65	4.55002035
Transition metal ion homeostasis (GO:0055076)	0.024823873	7	89	3.876889999
Neg. reg. of cell cycle phase transition (GO:1901988)	0.024823873	7	89	3.876889999
Cellular carbohydrate metabolic proc. (GO:0044262)	0.024823873	27	727	1.830647802
Pyrimidine deoxyribonucleotide salvage (GO:0010139)	0.025034175	2	4	24.64594356
DNA methylation on cytosine within a CG sequence (GO:0010424)	0.025034175	2	4	24.64594356
Reg. of oxidoreductase activity (GO:0051341)	0.025034175	2	4	24.64594356
Pos. reg. of oxidoreductase activity (GO:0051353)	0.025034175	2	4	24.64594356
Reg. of superoxide dismutase activity (GO:1901668)	0.025034175	2	4	24.64594356
Male meiotic nuclear division (GO:0007140)	0.025034175	4	28	7.041698161
Reg. of cell population proliferation (GO:0042127)	0.025034175	7	90	3.833813443
Neg. reg. of cell cycle proc. (GO:0010948)	0.02532482	8	115	3.429000843
Very long-chain fatty acid biosynthetic proc. (GO:0042761)	0.026098573	3	14	10.56254724
Male meiosis II (GO:0007142)	0.026098573	3	14	10.56254724
Cell population proliferation (GO:0008283)	0.029129515	8	118	3.341822856
Cellular proc. involved in reproduction in multicellular organism (GO:0022412)	0.030082882	4	30	6.572251617
Male gamete generation (GO:0048232)	0.030082882	4	30	6.572251617
Gamete generation (GO:0007276)	0.030082882	4	30	6.572251617
Neg. reg. of mitotic cell cycle phase transition (GO:1901991)	0.030576399	5	49	5.029784401
Iron ion homeostasis (GO:0055072)	0.033756048	6	72	4.10765726
Oligosaccharide metabolic proc. (GO:0009311)	0.034297251	9	149	2.97736231
Cell wall macromolecule catabolic proc. (GO:0016998)	0.035177351	5	51	4.832537953
Nuclear DNA replication (GO:0033260)	0.035177351	5	51	4.832537953
Neg. reg. of mitotic nuclear division (GO:0045839)	0.035177351	5	51	4.832537953
Neg. reg. of cellular proc. (GO:0048523)	0.035177351	28	792	1.742642474
L-proline biosynthetic proc. (GO:0055129)	0.03567097	3	16	9.242228836

Proline biosynthetic proc. (GO:0006561)	0.03567097	3	16	9.242228836
Homologous chromosome pairing at meiosis (GO:0007129)	0.03567097	3	16	9.242228836
Post-embryonic plant morphogenesis (GO:0090698)	0.040243011	5	53	4.650178031
Calcium import into the mitochondrion (GO:0036444)	0.040471978	3	17	8.698568316
Mitochondrial calcium ion homeostasis (GO:0051560)	0.040471978	3	17	8.698568316
Cytosolic calcium ion transport (GO:0060401)	0.040471978	3	17	8.698568316
Calcium ion import (GO:0070509)	0.040471978	3	17	8.698568316
Mitochondrial calcium ion transmembrane transport (GO:0006851)	0.040471978	3	17	8.698568316
Deoxyribonucleoside monophosphate biosynthetic proc. (GO:0009157)	0.040471978	3	17	8.698568316
Mitotic cell cycle checkpoint signaling (GO:0007093)	0.041638596	5	54	4.564063623
Deoxyribonucleotide biosynthetic proc. (GO:0009263)	0.047033728	3	18	8.215314521
Oligosaccharide biosynthetic proc. (GO:0009312)	0.048976167	7	105	3.286125808
DNA methylation on cytosine (GO:0032776)	0.049662694	2	6	16.43062904
C-5 methylation of cytosine (GO:0090116)	0.049662694	2	6	16.43062904
Meiosis I (GO:0007127)	0.049662694	6	80	3.696891534
Cellular components				
Nucleosome (GO:0000786)	1.93E-21	29	115	12.43012806
DNA packaging complex (GO:0044815)	1.93E-21	30	124	11.92545656
Protein-DNA complex (GO:0032993)	4.31E-19	33	193	8.428146503
Chromosome (GO:0005694)	4.31E-19	59	669	4.347117101
Chromatin (GO:0000785)	6.07E-11	30	301	4.91281267
Chromosomal region (GO:0098687)	3.85E-10	16	82	9.617929195
Chromosome, centromeric region (GO:0000775)	0.000000119	12	61	9.69676468
Kinetochore (GO:0000776)	0.000901702	7	44	7.841891134
CAF-1 complex (GO:0033186)	0.000951401	3	4	36.96891534
Origin recognition complex (GO:0000808)	0.001079314	4	11	17.92432259
Nuclear origin of replication recognition complex (GO:0005664)	0.001079314	4	11	17.92432259
Condensed chromosome, centromeric region (GO:0000779)	0.001079314	7	48	7.188400206
Condensed chromosome (GO:0000793)	0.002183224	9	92	4.822032436
Supramolecular complex (GO:0099080)	0.002242002	25	534	2.307672618

Mitotic checkpoint complex (GO:0033597)	0.002595338	3	6	24.64594356
Bub1-bub3 complex (GO:1990298)	0.002595338	3	6	24.64594356
Microtubule cytoskeleton (GO:0015630)	0.005466292	22	476	2.278196464
Microtubule (GO:0005874)	0.007609676	19	394	2.377019938
Cytoskeleton (GO:0005856)	0.007633897	27	660	2.016486291
Phragmoplast (GO:0009524)	0.008071523	3	9	16.43062904
Chromosome, telomeric region (GO:0000781)	0.008071523	4	21	9.388930881
Supramolecular polymer (GO:0099081)	0.008071523	19	407	2.301095468
Supramolecular fiber (GO:0099512)	0.008071523	19	407	2.301095468
Polymeric cytoskeletal fiber (GO:0099513)	0.008071523	19	407	2.301095468
Thylakoid (GO:0009579)	0.008071523	23	535	2.119090475
Spindle (GO:0005819)	0.015979365	8	108	3.651250898
Kinesin complex (GO:0005871)	0.015979365	9	133	3.335541234
Cell-cell junction (GO:0005911)	0.016251461	24	613	1.929861812
Plasmodesma (GO:0009506)	0.016251461	24	613	1.929861812
Symplast (GO:0055044)	0.016251461	24	613	1.929861812
Anchoring junction (GO:0070161)	0.016251461	24	613	1.929861812
Cell junction (GO:0030054)	0.016251461	24	614	1.926718715
Microtubule associated complex (GO:0005875)	0.020227612	11	198	2.738438174
Mitotic spindle (GO:0072686)	0.029594265	4	32	6.161485891
Ion channel complex (GO:0034702)	0.031445542	3	17	8.698568316
Cation channel complex (GO:0034703)	0.031445542	3	17	8.698568316
Calcium channel complex (GO:0034704)	0.031445542	3	17	8.698568316
Uniplex complex (GO:1990246)	0.031445542	3	17	8.698568316
Outer kinetochore (GO:0000940)	0.035463647	2	6	16.43062904
ATP-independent citrate lyase complex (GO:0009346)	0.035463647	2	6	16.43062904
Ndc80 complex (GO:0031262)	0.035463647	2	6	16.43062904
NMS complex (GO:0031617)	0.035463647	2	6	16.43062904
Nuclear chromosome (GO:0000228)	0.035463647	9	161	2.755447106
Spindle microtubule (GO:0005876)	0.043518543	3	20	7.393783069
Molecular functions				
Protein heterodimerization activity (GO:0046982)	9.59E-13	27	185	7.193951094
Protein dimerization activity (GO:0046983)	0.000384904	47	1077	2.151085139
Superoxide dismutase copper chaperone activity (GO:0016532)	0.001271249	4	7	28.16679264
DNA replication origin binding (GO:0003688)	0.002056607	7	39	8.847261792
Copper chaperone activity (GO:0016531)	0.002658	4	9	21.90750539

Nucleosomal DNA binding (GO:0031492)	0.002676893	7	43	8.024260695
Metallochaperone activity (GO:0016530)	0.004813491	4	11	17.92432259
Ferroxidase activity (GO:0004322)	0.008584163	4	14	14.08339632
Oxidoreductase activity, acting on superoxide radicals as acceptor (GO:0016721)	0.008584163	4	14	14.08339632
Oxidoreductase activity, acting on metal ions, oxygen as acceptor (GO:0016724)	0.008584163	4	14	14.08339632
Microtubule motor activity (GO:0003777)	0.008584163	12	165	3.584864518
Phosphoenolpyruvate carboxylase activity (GO:0008964)	0.013182048	4	16	12.32297178
Cytoskeletal motor activity (GO:0003774)	0.025284842	13	216	2.966641355
Double-stranded DNA binding (GO:0003690)	0.026254662	39	1092	1.76042454
Nucleosome binding (GO:0031491)	0.035284898	7	75	4.600576132
ATP citrate synthase activity (GO:0003878)	0.043471768	3	11	13.44324194
Phosphoglycerate mutase activity (GO:0004619)	0.043471768	3	11	13.44324194
Phosphoenolpyruvate carboxykinase activity (GO:0004611)	0.043471768	4	23	8.572502109
Unique DEGs upregulated in Henderson				
Biological process				
Ribosome biogenesis (GO:0042254)	1.87E-31	90	703	4.777094922
RRNA processing (GO:0006364)	9.91E-31	72	462	5.815234165
RRNA metabolic proc. (GO:0016072)	9.91E-31	73	475	5.73463706
Ribonucleoprotein complex biogenesis (GO:0022613)	2.36E-25	90	862	3.895937042
NcRNA processing (GO:0034470)	4.36E-25	81	716	4.221323963
RNA modification (GO:0009451)	7.46E-25	77	657	4.373227215
NcRNA metabolic proc. (GO:0034660)	8.17E-22	83	843	3.673898927
RNA processing (GO:0006396)	4.33E-11	90	1445	2.324081474
RRNA modification (GO:0000154)	1.6E-09	17	74	8.572231444
Ribosomal large subunit biogenesis (GO:0042273)	3.75E-09	28	227	4.60265964
Maturation of SSU-rRNA (GO:0030490)	4.64E-09	20	115	6.489464213
Ribosomal small subunit biogenesis (GO:0042274)	3.65E-08	23	172	4.98971885
Maturation of LSU-rRNA (GO:0000470)	4.32E-08	17	92	6.895055726
Protein import into mitochondrial intermembrane space (GO:0045041)	0.00000107	6	8	27.98581442
Mitochondrial gene expression (GO:0140053)	0.00000145	16	102	5.853242231

Mitochondrion organization (GO:0007005)	0.00000903	27	308	3.271069218
Maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, L (GO:0000462)	0.0000384	13	85	5.706911176
RNA methylation (GO:0001510)	0.0000384	16	130	4.592543905
Reg. of vacuole fusion, non-autophagic (GO:0032889)	0.0000409	6	13	17.22203964
Reg. of vacuole organization (GO:0044088)	0.0000409	6	13	17.22203964
RRNA methylation (GO:0031167)	0.0000523	10	50	7.462883845
Protein transmembrane import into intracellular organelle (GO:0044743)	0.0000582	12	76	5.891750404
Maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, L (GO:0000463)	0.000114714	9	43	7.809994722
Vacuole fusion, non-autophagic (GO:0042144)	0.000142529	6	16	13.99290721
Vacuole fusion (GO:0097576)	0.000142529	6	16	13.99290721
Reg. of mitochondrial gene expression (GO:0062125)	0.000162966	5	10	18.65720961
Reg. of mitochondrial translation (GO:0070129)	0.000162966	5	10	18.65720961
Pos. reg. of mitochondrial translation (GO:0070131)	0.000162966	5	10	18.65720961
Mitochondrial translation (GO:0032543)	0.000162966	11	71	5.781107204
Protein refolding (GO:0042026)	0.001266554	10	73	5.111564277
Mitochondrial transmembrane transport (GO:1990542)	0.001266554	13	122	3.976126639
Transcription by RNA polymerase I (GO:0006360)	0.001814065	8	48	6.219069871
Protein targeting to mitochondrion (GO:0006626)	0.00190865	12	110	4.070663916
Mitochondrial RNA metabolic proc. (GO:0000959)	0.002015315	9	63	5.330631318
Establishment of protein localization to mitochondrion (GO:0072655)	0.002423985	12	114	3.927833603
Macromolecule methylation (GO:0043414)	0.002423985	21	295	2.656280691
Box C/D RNA 3 -end processing (GO:0000494)	0.002806221	3	4	27.98581442
Box C/D RNA metabolic proc. (GO:0033967)	0.002806221	3	4	27.98581442
Histone glutamine methylation (GO:1990258)	0.002806221	3	4	27.98581442
Sno(s)RNA 3 -end processing (GO:0031126)	0.004561472	6	30	7.462883845
de novo protein folding (GO:0006458)	0.005085691	12	125	3.582184246

RRNA pseudouridine synthesis (GO:0031118)	0.005283591	5	20	9.328604806
Pseudouridine synthesis (GO:0001522)	0.005350299	7	44	5.936384877
Small nucleolar ribonucleoprotein complex assembly (GO:0000491)	0.006951589	4	12	12.43813974
Sno(s)RNA metabolic proc. (GO:0016074)	0.006951589	6	33	6.784439859
Sno(s)RNA processing (GO:0043144)	0.006951589	6	33	6.784439859
Cytochrome complex assembly (GO:0017004)	0.007805737	7	48	5.441686137
Intracellular protein transmembrane transport (GO:0065002)	0.007805737	12	134	3.341589781
ADP metabolic proc. (GO:0046031)	0.007805737	13	153	3.170506209
Glycolytic proc. (GO:0006096)	0.007805737	13	153	3.170506209
ATP generation from ADP (GO:0006757)	0.007805737	13	153	3.170506209
Ribosome assembly (GO:0042255)	0.007805737	14	174	3.002309593
Peptidyl-glutamine methylation (GO:0018364)	0.009865577	3	6	18.65720961
GMP biosynthetic proc. (GO:0006177)	0.009865577	3	6	18.65720961
Seed dormancy proc. (GO:0010162)	0.010480452	4	14	10.66126264
Dormancy proc. (GO:0022611)	0.010480452	4	14	10.66126264
RRNA base methylation (GO:0070475)	0.010480452	6	37	6.050986901
Purine nucleoside diphosphate metabolic proc. (GO:0009135)	0.010480452	13	160	3.031796562
Cellular transition metal ion homeostasis (GO:0046916)	0.012640164	8	69	4.326309475
Seed maturation (GO:0010431)	0.013425346	4	15	9.950511794
Cellular iron ion homeostasis (GO:0006879)	0.013782654	7	54	4.837054344
Pos. reg. of transcription by RNA polymerase I (GO:0045943)	0.01393398	3	7	15.99189395
Endonucleolytic cleavage to generate mature 5 -end of SSU-rRNA from (SSU-rRNA, 5 (GO:0000472)	0.01393398	3	7	15.99189395
Endonucleolytic cleavage in 5 -ETS of tricistronic rRNA transcript (SSU-rRNA, 5. (GO:0000480)	0.01393398	3	7	15.99189395
Transition metal ion homeostasis (GO:0055076)	0.014145398	9	89	3.773368236
Nucleoside diphosphate phosphorylation (GO:0006165)	0.014145398	13	168	2.887425297
Iron ion homeostasis (GO:0055072)	0.014520536	8	72	4.146046581
Ribonucleoside diphosphate metabolic proc. (GO:0009185)	0.014520536	13	169	2.87033994
Nucleotide phosphorylation (GO:0046939)	0.015009898	13	170	2.853455588
Protein import into mitochondrial matrix	0.015255258	6	41	5.460646716

(GO:0030150)				
SnRNA pseudouridine synthesis (GO:0031120)	0.019810276	3	8	13.99290721
Histone arginine methylation (GO:0034969)	0.023398751	4	18	8.292093161
NcRNA 3 -end processing (GO:0043628)	0.024000759	7	61	4.281982534
Cellular metal ion homeostasis (GO:0006875)	0.025113068	11	138	2.974337764
Recognition of pollen (GO:0048544)	0.026592628	12	161	2.781198948
Protein transmembrane transport (GO:0071806)	0.026592628	12	161	2.781198948
Cell recognition (GO:0008037)	0.026592628	12	161	2.781198948
SnRNA modification (GO:0040031)	0.026637235	3	9	12.43813974
Pollen-pistil interaction (GO:0009875)	0.031958417	12	165	2.713775944
Cleavage involved in rRNA processing (GO:0000469)	0.032012667	7	65	4.018475917
Metal ion homeostasis (GO:0055065)	0.034130849	12	167	2.681275633
RRNA 5 -end processing (GO:0000967)	0.034364198	3	10	11.19432577
Endonucleolytic cleavage involved in rRNA processing (GO:0000478)	0.034364198	5	34	5.487414592
Endonucleolytic cleavage of tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, L (GO:0000479)	0.034364198	5	34	5.487414592
TRNA acetylation (GO:0051391)	0.035904378	2	3	24.87627948
Nucleoside monophosphate metabolic proc. (GO:0009123)	0.041752821	11	150	2.736390743
Sequestering of metal ion (GO:0051238)	0.042179246	5	36	5.182558226
Intracellular sequestering of iron ion (GO:0006880)	0.042179246	5	36	5.182558226
Protein-tetrapyrrole linkage (GO:0017006)	0.043385122	3	11	10.17665979
Peptidyl-arginine N-methylation (GO:0035246)	0.04819136	4	23	6.489464213
Pyruvate metabolic proc. (GO:0006090)	0.048421136	13	200	2.42543725
Iron ion transport (GO:0006826)	0.048922425	6	54	4.146046581
Porphyrin-containing compound biosynthetic proc. (GO:0006779)	0.048922425	7	72	3.627790758
Nucleoside monophosphate biosynthetic proc. (GO:0009124)	0.049687873	9	112	2.998480116
Ribosomal large subunit assembly (GO:0000027)	0.04982758	8	92	3.244732107
Response to oxidative stress (GO:0006979)	0.04982758	22	431	1.904680332
Cellular components				
Nucleolus (GO:0005730)	4.43E-42	84	457	6.858667866
Preribosome (GO:0030684)	3.46E-32	48	176	10.17665979
Small-subunit processome (GO:0032040)	1.15E-25	32	89	13.4164204

Membrane-enclosed lumen (GO:0031974)	1.33E-22	117	1519	2.87411919
Organelle lumen (GO:0043233)	1.33E-22	117	1519	2.87411919
Intracellular organelle lumen (GO:0070013)	1.33E-22	117	1519	2.87411919
Nuclear lumen (GO:0031981)	4.85E-20	98	1225	2.985153538
Mitochondrion (GO:0005739)	3.36E-18	116	1707	2.535719174
90S preribosome (GO:0030686)	2.01E-15	21	69	11.35656237
Sno(s)RNA-containing ribonucleoprotein complex (GO:0005732)	2.53E-14	17	45	14.09655837
Ribonucleoprotein complex (GO:1990904)	7.41E-13	87	1324	2.45192936
Box C/D RNP complex (GO:0031428)	1.01E-10	9	13	25.83305946
Pwp2p-containing subcomplex of 90S preribosome (GO:0034388)	6.06E-08	7	11	23.74553951
Phosphopyruvate hydratase complex (GO:0000015)	0.00000565	6	12	18.65720961
Preribosome, large subunit precursor (GO:0030687)	0.0000226	10	53	7.040456458
Box H/ACA snoRNP complex (GO:0031429)	0.0000488	5	10	18.65720961
Mitochondrial envelope (GO:0005740)	0.003159672	29	521	2.077002222
RNA polymerase I complex (GO:0005736)	0.003655855	6	34	6.58489751
Cajal body (GO:0015030)	0.00484498	3	6	18.65720961
Preribosome, small subunit precursor (GO:0030688)	0.00720858	5	26	7.175849851
Mitochondrial membrane (GO:0031966)	0.008454956	27	508	1.983246691
Mitochondrial inner membrane (GO:0005743)	0.009546806	23	410	2.093247908
TIM23 mitochondrial import inner membrane translocase complex (GO:0005744)	0.010654209	5	29	6.433520556
Mitochondrial matrix (GO:0005759)	0.014049569	15	227	2.465710522
Organelle inner membrane (GO:0019866)	0.016323709	25	484	1.927397687
Mitochondrial intermembrane space (GO:0005758)	0.024453885	4	22	6.784439859
Organelle envelope lumen (GO:0031970)	0.024453885	4	22	6.784439859
Noc complex (GO:0030689)	0.031491279	3	12	9.328604806
Mitochondrial DNA-directed RNA polymerase complex (GO:0034245)	0.031491279	3	12	9.328604806
NSL complex (GO:0044545)	0.034792057	2	4	18.65720961
Pre-snoRNP complex (GO:0070761)	0.034792057	2	4	18.65720961
Molecular functions				
SnoRNA binding (GO:0030515)	1.9E-16	20	48	15.54767468
Phosphopyruvate hydratase activity (GO:0004634)	0.000192241	6	13	17.22203964

RNA methyltransferase activity (GO:0008173)	0.000371681	14	110	4.749107901
Translation activator activity (GO:0008494)	0.000444147	5	10	18.65720961
Mitochondrial ribosome binding (GO:0097177)	0.000444147	5	10	18.65720961
RRNA methyltransferase activity (GO:0008649)	0.001120115	8	40	7.462883845
Catalytic activity, acting on a rRNA (GO:0140102)	0.002012234	8	44	6.784439859
U3 snoRNA binding (GO:0034511)	0.002016179	5	14	13.32657829
Box H/ACA snoRNA binding (GO:0034513)	0.002368767	4	8	18.65720961
Magnesium ion binding (GO:0000287)	0.002368767	19	240	2.954058189
Protein-glutamine N-methyltransferase activity (GO:0036009)	0.004505319	3	4	27.98581442
Histone-glutamine methyltransferase activity (GO:1990259)	0.004505319	3	4	27.98581442
NAD+ nucleosidase activity (GO:0003953)	0.007880868	13	146	3.32251678
NAD(P)+ nucleosidase activity (GO:0050135)	0.007880868	13	146	3.32251678
NAD+ nucleotidase, cyclic ADP-ribose generating (GO:0061809)	0.007880868	13	146	3.32251678
Zinc ion binding (GO:0008270)	0.02567225	53	1229	1.609165353
Catalytic activity, acting on RNA (GO:0140098)	0.02609958	39	830	1.753328132
Histone-arginine N-methyltransferase activity (GO:0008469)	0.036886228	4	17	8.779863347
DNA-directed 5 -3 RNA polymerase activity (GO:0003899)	0.038427151	10	114	3.273194669
ADP binding (GO:0043531)	0.03920241	25	471	1.9805955
Unique DEGs downregulated in Henderson				
Biological process				
Cell wall organization or biogenesis (GO:0071554)	3.23E-10	76	799	2.535460694
External encapsulating structure organization (GO:0045229)	2.21E-08	61	633	2.568716612
Cell wall organization (GO:0071555)	4.82E-08	58	603	2.563898043
Polysaccharide metabolic proc. (GO:0005976)	0.0000446	60	770	2.077067425
Auxin-activated signaling pathway (GO:0009734)	0.000201348	49	613	2.130716529
Cellular response to auxin stimulus (GO:0071365)	0.000201348	49	613	2.130716529

Cell wall polysaccharide metabolic proc. (GO:0010383)	0.000303485	25	224	2.974966364
Response to auxin (GO:0009733)	0.000751624	49	652	2.003265694
Xyloglucan metabolic proc. (GO:0010411)	0.001254517	15	104	3.844571916
Fatty acid derivative metabolic proc. (GO:1901568)	0.001406217	8	30	7.108186298
Cellular polysaccharide metabolic proc. (GO:0044264)	0.001406217	37	453	2.177176267
Neg. reg. of catalytic activity (GO:0043086)	0.002541859	32	381	2.238798834
Neg. reg. of molecular function (GO:0044092)	0.002541859	32	381	2.238798834
Cutin biosynthetic proc. (GO:0010143)	0.003827531	8	35	6.092731112
Plant-type cell wall organization (GO:0009664)	0.004390742	17	149	3.041254205
Cellular developmental proc. (GO:0048869)	0.004390742	38	504	2.009755054
Cell differentiation (GO:0030154)	0.004815211	37	491	2.0086779
Lipid metabolic proc. (GO:0006629)	0.004815211	91	1594	1.521749419
Hormone-mediated signaling pathway (GO:0009755)	0.007272801	64	1040	1.640350684
Galactose metabolic proc. (GO:0006012)	0.007561589	8	40	5.331139723
Response to hormone (GO:0009725)	0.008071044	74	1258	1.567982272
Hemicellulose metabolic proc. (GO:0010410)	0.008266843	18	176	2.726150995
Phenylpropanoid biosynthetic proc. (GO:0009699)	0.0089961	9	53	4.526439388
Glucan metabolic proc. (GO:0044042)	0.0089961	29	365	2.117850027
Cellular glucan metabolic proc. (GO:0006073)	0.009805492	28	350	2.132455889
Neg. reg. of peptidase activity (GO:0010466)	0.010180736	12	94	3.402855143
Neg. reg. of proteolysis (GO:0045861)	0.010180736	12	94	3.402855143
Plant-type cell wall organization or biogenesis (GO:0071669)	0.010180736	26	317	2.186271811
Cellular response to hormone stimulus (GO:0032870)	0.010180736	64	1068	1.597345235
Response to endogenous stimulus (GO:0009719)	0.010180736	74	1279	1.542237449
Reg. of endopeptidase activity (GO:0052548)	0.012056648	12	96	3.331962327
Polysaccharide catabolic proc. (GO:0000272)	0.013125678	25	308	2.163611901
Neg. reg. of hydrolase activity (GO:0051346)	0.016514415	16	161	2.649013527
Wax biosynthetic proc. (GO:0010025)	0.020556557	4	11	9.692981315

Wax metabolic proc. (GO:0010166)	0.020556557	4	11	9.692981315
NADH oxidation (GO:0006116)	0.020556557	4	11	9.692981315
NADH metabolic proc. (GO:0006734)	0.020556557	5	19	7.014657531
Inorganic anion transport (GO:0015698)	0.020556557	18	199	2.411068217
Suberin biosynthetic proc. (GO:0010345)	0.027210169	4	12	8.885232872
Long-chain fatty-acyl-CoA metabolic proc. (GO:0035336)	0.027210169	4	12	8.885232872
Cellular carbohydrate metabolic proc. (GO:0044262)	0.029222385	45	727	1.649940079
Cell wall biogenesis (GO:0042546)	0.030144877	23	296	2.071219825
Glycerol-3-phosphate catabolic proc. (GO:0046168)	0.035704047	4	13	8.201753421
Phenylpropanoid metabolic proc. (GO:0009698)	0.036640209	12	113	2.830693658
Chloroplast relocation (GO:0009902)	0.040367076	7	45	4.146442007
Chloroplast localization (GO:0019750)	0.040367076	7	45	4.146442007
Plastid localization (GO:0051644)	0.040367076	7	45	4.146442007
Establishment of plastid localization (GO:0051667)	0.040367076	7	45	4.146442007
Fatty-acyl-CoA metabolic proc. (GO:0035337)	0.043014612	4	14	7.615913891
Stomatal complex development (GO:0010374)	0.044382091	7	46	4.056301963
Cellular components				
Cell wall (GO:0005618)	5.61E-12	74	749	2.633540317
External encapsulating structure (GO:0030312)	5.61E-12	76	776	2.610609658
Extracellular region (GO:0005576)	5.61E-12	112	1395	2.1400991
Anchored component of membrane (GO:0031225)	0.00000202	36	341	2.81409135
Anchored component of plasma membrane (GO:0046658)	0.000318932	27	275	2.617104955
Apoplast (GO:0048046)	0.00236893	27	311	2.31416033
Cell-cell junction (GO:0005911)	0.014415012	40	613	1.739360432
Plasmodesma (GO:0009506)	0.014415012	40	613	1.739360432
Symplast (GO:0055044)	0.014415012	40	613	1.739360432
Anchoring junction (GO:0070161)	0.014415012	40	613	1.739360432
Cell junction (GO:0030054)	0.014415012	40	614	1.736527597
Glycerol-3-phosphate dehydrogenase complex (GO:0009331)	0.016628436	4	12	8.885232872
Molecular functions				
Hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553)	0.000845919	58	788	1.961967665
Hydrolase activity, acting on glycosyl bonds	0.000990099	68	1004	1.805366042

(GO:0016798)				
Inorganic anion transmembrane transporter activity (GO:0015103)	0.002879588	19	162	3.12628564
Monooxygenase activity (GO:0004497)	0.006219612	41	553	1.976281453
Oxidoreductase activity, acting on paired donors, with incorporation or reductio (GO:0016705)	0.007283985	48	697	1.835686562
Hexosyltransferase activity (GO:0016758)	0.013145323	53	820	1.722868325
Fatty acid ligase activity (GO:0015645)	0.016584291	7	34	5.487937951
Racemase and epimerase activity, acting on carbohydrates and derivatives (GO:0016857)	0.016584291	10	66	4.038742215
Endopeptidase inhibitor activity (GO:0004866)	0.016584291	12	94	3.402855143
Peptidase inhibitor activity (GO:0030414)	0.016584291	12	94	3.402855143
Endopeptidase regulator activity (GO:0061135)	0.016832935	12	96	3.331962327
Oxidoreductase activity, acting on metal ions (GO:0016722)	0.017293456	12	98	3.263963096
Peptidase regulator activity (GO:0061134)	0.017293456	12	98	3.263963096
Oxidoreductase activity, acting on paired donors, with oxidation of a pair of do (GO:0016717)	0.017300327	6	26	6.151315065
Alkylbase DNA N-glycosylase activity (GO:0003905)	0.017969172	5	18	7.404360727
DNA-3-methyladenine glycosylase activity (GO:0008725)	0.017969172	5	18	7.404360727
DNA-3-methylbase glycosylase activity (GO:0043733)	0.017969172	5	18	7.404360727
Alcohol-forming fatty acyl-CoA reductase activity (GO:0102965)	0.018229413	4	11	9.692981315
Nitrate transmembrane transporter activity (GO:0015112)	0.018229413	6	28	5.711935418
Xyloglucan:xyloglucosyl transferase activity (GO:0016762)	0.018229413	9	62	3.869375606
O-acyltransferase activity (GO:0008374)	0.018229413	20	230	2.317886836
Voltage-gated anion channel activity (GO:0008308)	0.018300989	7	39	4.784356162
Glycosyltransferase activity (GO:0016757)	0.018721786	68	1201	1.509231895
Fatty-acyl-CoA reductase (alcohol-forming) activity (GO:0080019)	0.022727196	4	12	8.885232872
Acyltransferase activity, transferring groups other than amino-acyl groups (GO:0016747)	0.022727196	46	745	1.645855217
Racemase and epimerase activity (GO:0016854)	0.030390378	10	83	3.211529954

Anion channel activity (GO:0005253)	0.031563261	7	44	4.240679325
Glycerol-3-phosphate dehydrogenase [NAD+] activity (GO:0004367)	0.033155373	4	14	7.615913891
Acid-thiol ligase activity (GO:0016878)	0.033155373	7	45	4.146442007
Acyltransferase activity (GO:0016746)	0.033155373	49	828	1.577450764
Enzyme regulator activity (GO:0030234)	0.033155373	52	894	1.55044332
Pectin acetylesterase activity (GO:0052793)	0.037312965	5	24	5.553270545
Carbon-oxygen lyase activity, acting on polysaccharides (GO:0016837)	0.037312965	7	47	3.969997666
Glucosyltransferase activity (GO:0046527)	0.038860644	27	390	1.84539452
Aspartic-type endopeptidase activity (GO:0004190)	0.038886746	17	205	2.210472568
Aspartic-type peptidase activity (GO:0070001)	0.038886746	17	205	2.210472568
Molecular function regulator (GO:0098772)	0.038886746	56	997	1.497210755
Oxidoreductase activity, acting on CH-OH group of donors (GO:0016614)	0.039360201	28	413	1.807166008
Rhamnogalacturonan endolyase activity (GO:0102210)	0.045365618	3	8	9.995886981
Phosphoenolpyruvate carboxylase activity (GO:0008964)	0.045365618	4	16	6.663924654
Protein homodimerization activity (GO:0042803)	0.048838115	7	51	3.6586253
Unique DEGs upregulated in Pembina				
Biological process				
RNA modification (GO:0009451)	6.76E-12	45	657	4.01737843
Phenylpropanoid metabolic proc. (GO:0009698)	6.48E-08	16	113	8.304952224
Lignin metabolic proc. (GO:0009808)	0.000000154	13	74	10.30403278
Secondary metabolic proc. (GO:0019748)	0.000000222	21	232	5.309173391
Phenylpropanoid catabolic proc. (GO:0046271)	0.00000271	10	51	11.50073041
Lignin catabolic proc. (GO:0046274)	0.00000271	10	51	11.50073041
Hydrogen peroxide metabolic proc. (GO:0042743)	0.0000584	16	199	4.715877393
Secondary metabolite biosynthetic proc. (GO:0044550)	0.0000615	11	91	7.090010724
Reactive oxygen species metabolic proc. (GO:0072593)	0.0000671	17	230	4.335275332
Cellular oxidant detoxification (GO:0098869)	0.000101221	20	322	3.643088514
Cellular response to toxic substance (GO:0097237)	0.000126489	21	360	3.421467296
Detoxification (GO:0098754)	0.000128243	21	362	3.402564162

Response to toxic substance (GO:0009636)	0.000246011	21	379	3.249942551
Plant-type cell wall biogenesis (GO:0009832)	0.000723851	14	199	4.126392719
Cell wall biogenesis (GO:0042546)	0.001166721	17	296	3.368626102
Plant-type secondary cell wall biogenesis (GO:0009834)	0.001937228	9	92	5.73786441
Detection of chemical stimulus (GO:0009593)	0.001962123	4	12	19.55124169
Phenylpropanoid biosynthetic proc. (GO:0009699)	0.001962123	7	53	7.746718407
Lignin biosynthetic proc. (GO:0009809)	0.002878309	5	25	11.73074502
Response to oxidative stress (GO:0006979)	0.00314851	20	431	2.721750584
Response to nitric oxide (GO:0071731)	0.003289257	4	14	16.75820717
Cellular response to nitric oxide (GO:0071732)	0.003289257	4	14	16.75820717
Pos. reg. of seed germination (GO:0010030)	0.004266743	4	15	15.64099335
Ethylene metabolic proc. (GO:0009692)	0.004621811	4	16	14.66343127
Ethylene biosynthetic proc. (GO:0009693)	0.004621811	4	16	14.66343127
Cellular alkene metabolic proc. (GO:0043449)	0.004621811	4	16	14.66343127
Alkene biosynthetic proc. (GO:0043450)	0.004621811	4	16	14.66343127
Olefin metabolic proc. (GO:1900673)	0.004621811	4	16	14.66343127
Olefin biosynthetic proc. (GO:1900674)	0.004621811	4	16	14.66343127
Olefinic compound biosynthetic proc. (GO:0120255)	0.004933345	5	30	9.775620846
Detection of stimulus (GO:0051606)	0.005638671	5	31	9.460278238
Cellular response to reactive nitrogen species (GO:1902170)	0.01049486	4	20	11.73074502
Seed germination (GO:0009845)	0.01049486	5	36	8.146350705
Pos. reg. of post-embryonic development (GO:0048582)	0.01049486	5	36	8.146350705
Flavonoid biosynthetic proc. (GO:0009813)	0.01049486	6	55	6.39858819
Pos. reg. of developmental proc. (GO:0051094)	0.010635876	5	37	7.926179065
Seedling development (GO:0090351)	0.010635876	5	37	7.926179065
Flavonoid metabolic proc. (GO:0009812)	0.010635876	6	56	6.284327687
Reg. of mitochondrial gene expression (GO:0062125)	0.013675735	3	10	17.59611752
Reg. of mitochondrial translation (GO:0070129)	0.013675735	3	10	17.59611752
Pos. reg. of mitochondrial translation (GO:0070131)	0.013675735	3	10	17.59611752
Polyketide metabolic proc. (GO:0030638)	0.013906586	4	23	10.20064784
Polyketide biosynthetic proc. (GO:0030639)	0.013906586	4	23	10.20064784

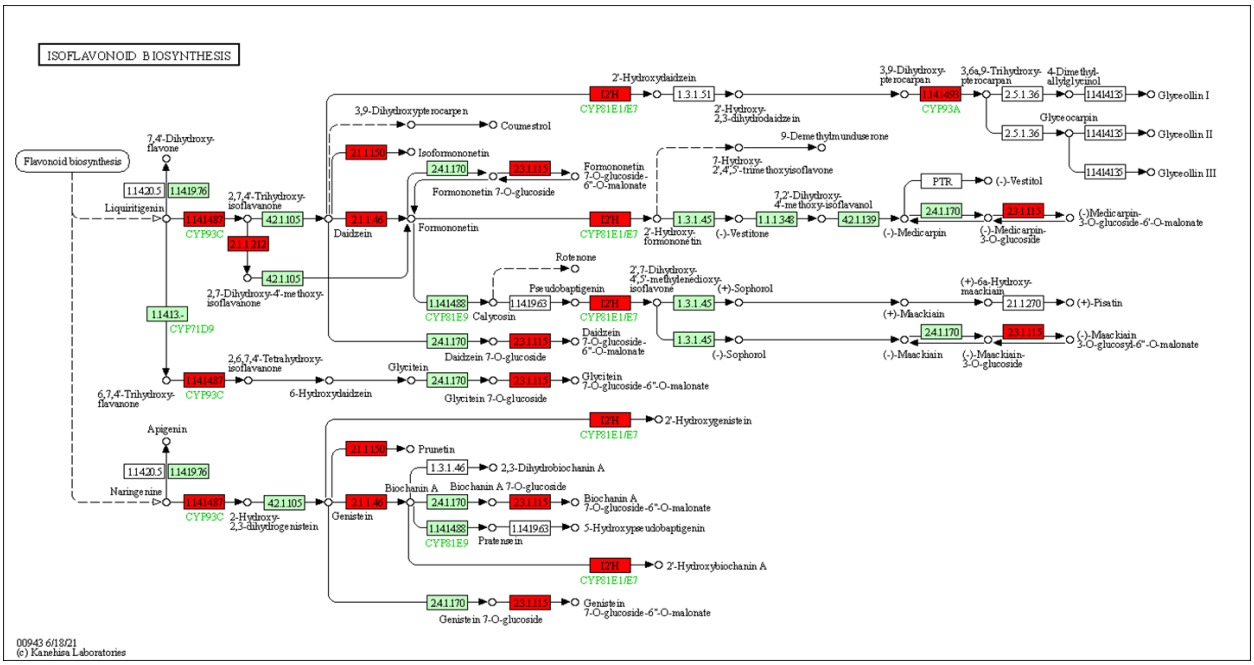
Aromatic compound catabolic proc. (GO:0019439)	0.014147813	18	436	2.421484063
Cellular manganese ion homeostasis (GO:0030026)	0.015359542	4	24	9.775620846
Cellular response to iron ion (GO:0071281)	0.015359542	4	24	9.775620846
Serine family amino acid biosynthetic proc. (GO:0009070)	0.015359542	7	86	4.774140413
Response to iron ion (GO:0010039)	0.015831673	5	42	6.982586319
Manganese ion homeostasis (GO:0055071)	0.017264721	4	25	9.384596013
Organic cyclic compound catabolic proc. (GO:1901361)	0.018161394	18	450	2.346149003
Reg. of seed germination (GO:0010029)	0.019106412	4	26	9.023650012
Reg. of seedling development (GO:1900140)	0.019106412	4	26	9.023650012
Response to cytokinin (GO:0009735)	0.027683328	7	97	4.232743047
Xylem development (GO:0010089)	0.030520865	3	14	12.56865537
Hydrocarbon biosynthetic proc. (GO:0120251)	0.03520426	4	31	7.568222591
Iron ion transmembrane transport (GO:0034755)	0.039075286	4	32	7.331715635
Iron ion transport (GO:0006826)	0.041472902	5	54	5.43090047
Cellular iron ion homeostasis (GO:0006879)	0.041472902	5	54	5.43090047
Biological proc. involved in interspecies interaction between organisms (GO:0044419)	0.04785361	15	382	2.303156744
Threonine catabolic proc. (GO:0006567)	0.048953307	2	5	23.46149003
Stomatal movement (GO:0010118)	0.048953307	5	57	5.145063603
Reg. of stomatal movement (GO:0010119)	0.048953307	5	57	5.145063603
Cellular response to chemical stimulus (GO:0070887)	0.049020139	44	1657	1.557491794
Sequestering of metal ion (GO:0051238)	0.049140728	4	36	6.517080564
Intracellular sequestering of iron ion (GO:0006880)	0.049140728	4	36	6.517080564
Response to fatty acid (GO:0070542)	0.049140728	6	83	4.240028319
Water-soluble vitamin metabolic proc. (GO:0006767)	0.049140728	7	112	3.665857817
Alpha-amino acid biosynthetic proc. (GO:1901607)	0.049140728	12	277	2.540955599
Cellular components				
Extracellular region (GO:0005576)	0.00388283	46	1395	1.934101329
Molecular functions				
Oxidoreductase activity, acting on diphenols and related substances as donors, o (GO:0016682)	5.06E-09	15	82	10.72933995

Copper ion binding (GO:0005507)	0.00000012	18	166	6.360042478
Hydroquinone:oxygen oxidoreductase activity (GO:0052716)	0.00000184	10	51	11.50073041
Heme binding (GO:0020037)	0.00000766	36	766	2.756571936
Peroxidase activity (GO:0004601)	0.0000254	19	274	4.067229111
Oxidoreductase activity, acting on peroxide as acceptor (GO:0016684)	0.0000254	19	275	4.052439187
Tetrapyrrole binding (GO:0046906)	0.0000254	36	821	2.571905119
Antioxidant activity (GO:0016209)	0.0000559	20	322	3.643088514
Oxidoreductase activity, acting on metal ions (GO:0016722)	0.000392186	10	98	5.985073988
Monooxygenase activity (GO:0004497)	0.000650336	25	553	2.651615058
Oxidoreductase activity, acting on the CH-NH2 group of donors, oxygen as accepto (GO:0016641)	0.00068383	6	32	10.99757345
1-aminocyclopropane-1-carboxylate oxidase activity (GO:0009815)	0.001568029	4	12	19.55124169
Oxidoreductase activity, acting on the CH-NH2 group of donors (GO:0016638)	0.001593564	7	55	7.465019555
Aldehyde dehydrogenase (NAD+) activity (GO:0004029)	0.004119768	5	29	10.11271122
Naringenin-chalcone synthase activity (GO:0016210)	0.015469928	4	22	10.66431365
Translation activator activity (GO:0008494)	0.015605637	3	10	17.59611752
Mitochondrial ribosome binding (GO:0097177)	0.015605637	3	10	17.59611752
Zinc ion binding (GO:0008270)	0.020456453	37	1229	1.765815971
Iron ion binding (GO:0005506)	0.031189436	22	621	2.077909745
Unique DEGs downregulated in Pembina				
Biological process				
External encapsulating structure organization (GO:0045229)	0.00000453	46	633	2.593885873
Cell wall organization or biogenesis (GO:0071554)	0.00000453	54	799	2.412368909
Cell wall organization (GO:0071555)	0.00000567	44	603	2.604546446
Xyloglucan metabolic proc. (GO:0010411)	0.0000123	16	104	5.491403871
Polysaccharide metabolic proc. (GO:0005976)	0.0000372	49	770	2.271444328
Glucan metabolic proc. (GO:0044042)	0.000150743	29	365	2.835971588
Cellular glucan metabolic proc. (GO:0006073)	0.000174615	28	350	2.855530013
Cellular polysaccharide metabolic proc. (GO:0044264)	0.001070806	31	453	2.442644327

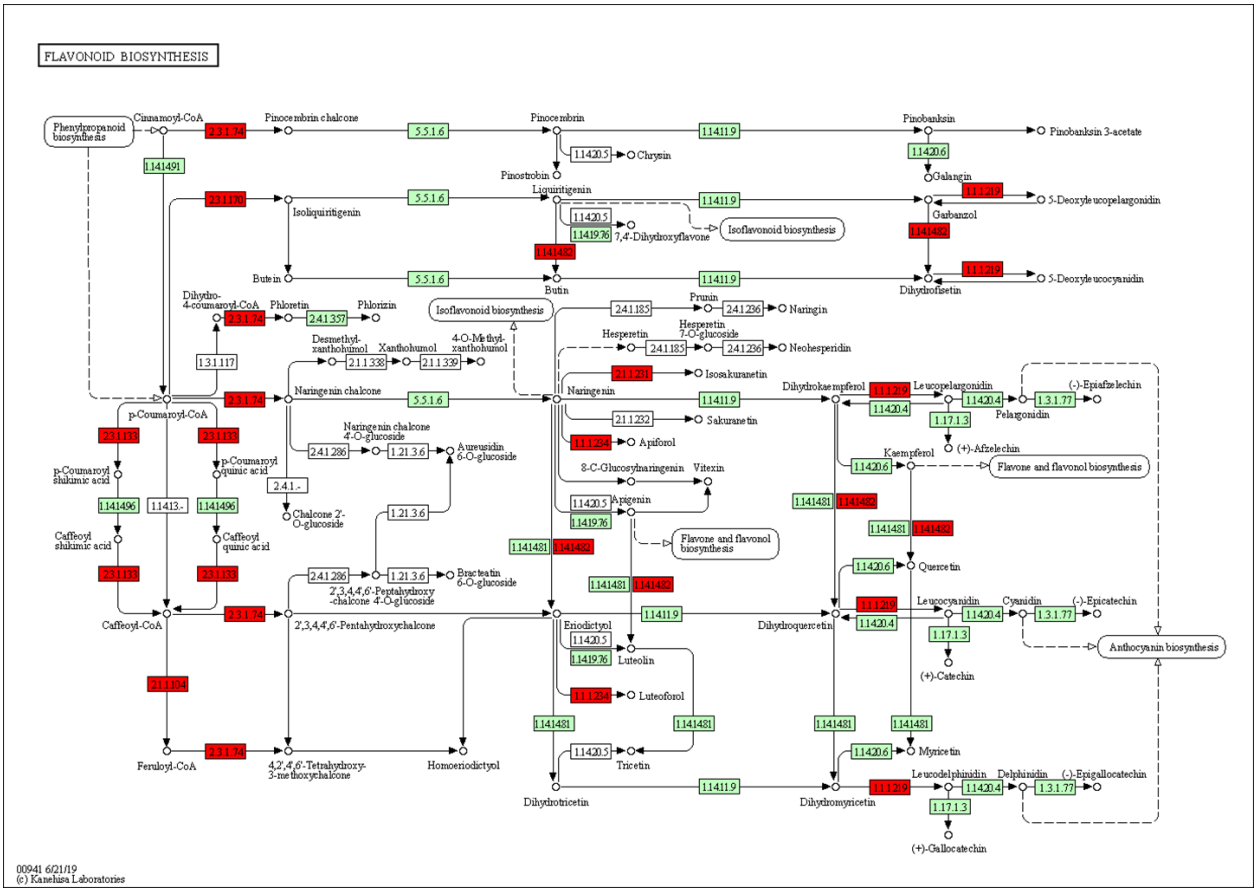
Cell wall macromolecule metabolic proc. (GO:0044036)	0.001202599	21	247	3.034723192
Hemicellulose metabolic proc. (GO:0010410)	0.001487777	17	176	3.447727998
Cell wall polysaccharide metabolic proc. (GO:0010383)	0.002703567	19	224	3.027626688
Polysaccharide catabolic proc. (GO:0000272)	0.002810283	23	308	2.665470385
Auxin-activated signaling pathway (GO:0009734)	0.007263396	35	613	2.038000621
Cellular response to auxin stimulus (GO:0071365)	0.007263396	35	613	2.038000621
Photosynthetic electron transport in photosystem II (GO:0009772)	0.016201408	4	11	12.97968188
Response to auxin (GO:0009733)	0.020639922	35	652	1.916095676
Plant-type cell wall organization (GO:0009664)	0.025485251	13	149	3.114252531
Pectin catabolic proc. (GO:0045490)	0.037041851	14	176	2.83930541
Neg. reg. of catalytic activity (GO:0043086)	0.038653784	23	381	2.154763461
Neg. reg. of molecular function (GO:0044092)	0.038653784	23	381	2.154763461
Hormone-mediated signaling pathway (GO:0009755)	0.041393141	48	1040	1.647421161
Cell wall biogenesis (GO:0042546)	0.049853648	19	296	2.291176953
Cellular components				
Cell wall (GO:0005618)	8.86E-13	64	749	3.049965301
External encapsulating structure (GO:0030312)	8.86E-13	65	776	2.989842958
Extracellular region (GO:0005576)	1.22E-08	83	1395	2.123736479
Anchored component of membrane (GO:0031225)	0.000000678	31	341	3.244920469
Anchored component of plasma membrane (GO:0046658)	0.00000361	26	275	3.374717288
Apoplast (GO:0048046)	0.000329634	24	311	2.754530559
Intrinsic component of plasma membrane (GO:0031226)	0.046644899	28	560	1.784706258
Cell-cell junction (GO:0005911)	0.046644899	30	613	1.746857675
Plasmodesma (GO:0009506)	0.046644899	30	613	1.746857675
Symplast (GO:0055044)	0.046644899	30	613	1.746857675
Anchoring junction (GO:0070161)	0.046644899	30	613	1.746857675
Cell junction (GO:0030054)	0.046644899	30	614	1.74401263
Molecular functions				
Xyloglucan:xyloglucosyl transferase activity (GO:0016762)	0.00000953	13	62	7.48425205

Electron transfer activity (GO:0009055)	0.00218671	28	389	2.569242942
Hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553)	0.003480424	44	788	1.993072978
Aspartic-type endopeptidase activity (GO:0004190)	0.009827024	17	205	2.960000623
Aspartic-type peptidase activity (GO:0070001)	0.009827024	17	205	2.960000623
Voltage-gated anion channel activity (GO:0008308)	0.010294727	7	39	6.406637849
Alanine-glyoxylate transaminase activity (GO:0008453)	0.013864008	3	5	21.4164751
Anion channel activity (GO:0005253)	0.013864008	7	44	5.678610821
Voltage-gated ion channel activity (GO:0005244)	0.013864008	9	72	4.461765645
Enzyme inhibitor activity (GO:0004857)	0.016650354	15	190	2.817957249
Molecular function inhibitor activity (GO:0140678)	0.016650354	15	190	2.817957249
Glucosyltransferase activity (GO:0046527)	0.016650354	24	390	2.196561548
O-acetyltransferase activity (GO:0016413)	0.025127333	12	138	3.10383697
Chlorophyll binding (GO:0016168)	0.02936372	7	53	4.714318417
Electron transporter, transferring electrons within the cyclic electron transpor (GO:0045156)	0.044918028	4	17	8.398617685

Appendix 1C: Isoflavonoid biosynthesis



Appendix 1D: Flavonoid biosynthesis



Appendix 1E: List of primer used for qRT-PCR

Genes	Forward sequence	Reverse sequence
Glyma.01g106700	TCATTGCCGAGGTTGTCTCC	GCCAGCTACAGCAGCTATGA
Glyma.03g142100	GCTCTCACTCCGCTATGGAC	TTTCGACGGCGACAGTGTTA
Glyma.08g046500	ACGGGCACATTGGAACTGAT	TGAAATTGCAGCGTGAAGGC
Glyma.09g269600	TGAACGAGCCTGAGGAAGTG	ACGGTGGAGCCACTAGAAGT
Glyma.10g161500	ATGTGGTAAGGGAAGCTGCG	TCGTCATCTTGGGTTCGCTT
Glyma.13g067200	TCACCTTGTTGGGGCTTCAT	AAGACACGGTTCTCCCTTCG
Glyma.17g147500	GCACATACACCGGCAATGTC	GCCACTTGGTGGGTTAGGAA
Glyma.20g188800	GGTTGCACTGTGGGTACCAT	GGGTCAACCAAAGCACTCCT
GmUBC2 (Control)	TCCCCTCACACCCTTCCTC	CCATCCCAAGGGGTGTCAT

Appendix 2

Appendix 2A: Distribution of population structure

Subpopulation	Number of genotypes	Origin distribution
Group 1	12	South Korea (100%)
Group 2	17	China (50%)
	1	South Korea
	2	Nepal
	9	Vietnam
	1	Japan
	3	Taiwan
	1	United States
Group 3	13	China (68.42%)
	1	Japan
	2	South Korea
	1	Unknown
	2	United States
Group4	76	South Korea (87.35%)
	2	North Korea
	5	Japan
	1	China
	2	United States
	1	Unknown
Group 5	31	United States (75.60%)
	1	South Korea
	2	Nepal
	2	Pakistan
	1	Georgia
	1	Unknown
	1	Australia
	1	Japan
	1	China
Group 6	47	Japan (88.67%)
	2	South Korea
	2	United States
	1	Taiwan
	1	North Korea

Appendix 2B: List of candidate genes integration of data from GWAS and RNA-Seq

Isolates	Gene id_RNA_Seq_GWAS	Chr	Start	End	Putative function	Associtate marker
LIM01 & SSTA	Glyma.04G123800	4	16009828	16014505	alternative oxidase 1A	ss715587169
LIM01 & SSTA	Glyma.15G051400	15	4035140	4043645	UDP-Glycosyltransferase superfamily protein	ss715621861
LIM01 & SSTA	Glyma.15G052000	15	4099684	4107040	sulfate transporter 1;3	ss715621861
LIM01 & SSTA	Glyma.15G054000	15	4244811	4247783	metacaspase 1	ss715621861
LIM01 & SSTA	Glyma.18G075000	18	7106464	7112929	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	ss715632608
LIM01 & SSTA	Glyma.18G075300	18	7144678	7150764	RNA polymerase I subunit 43	ss715632608
LIM01 & SSTA	Glyma.18G075500	18	7159568	7163740	Sucrase/ferredoxin-like family protein	ss715632608
LIM01 & SSTA	Glyma.20G129400	20	37061238	37064473	D-3-phosphoglycerate dehydrogenase	ss715637816
LIM01 & SSTA	Glyma.20G130200	20	37106093	37108629	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	ss715637816
LIM01 & SSTA	Glyma.13G312900	13	40803230	40808367	BEL1-like homeodomain 1	ss715616244
LIM01 & SSTA	Glyma.13G315200	13	41035905	41037849	lycopene cyclase	ss715616244
LIM01 & SSTA	Glyma.15G051300	15	4032576	4034552	UDP-glucosyl transferase 85A2	ss715621861
LIM01 & SSTA	Glyma.15G055400	15	4338136	4338342	RPM1-interacting protein 4 (RIN4) family protein	ss715621861
LIM01 & SSTA	Glyma.18G074600	18	7067323	7071282	RING/U-box superfamily protein	ss715632608
LIM01 & SSTA	Glyma.18G074800	18	7088533	7092750	ABC-2 type transporter family protein	ss715632608
LIM01 & SSTA	Glyma.20G130600	20	37142286	37144278	zinc finger protein 4	ss715637816
LIM01 & SSTA	Glyma.20G132400	20	37266529	37269324	Leucine-rich repeat (LRR) family protein	ss715637816
LIM01 & SSTA	Glyma.04G123700	4	16002630	16006606	solute:sodium symporters;urea	ss715587169

					transmembrane transporters	
LIM01 & SSTA	Glyma.10G161100	10	39528649	39529764	translocase of inner mitochondrial membrane 23	ss715606800
LIM01 & SSTA	Glyma.10G161500	10	39568701	39571015	CONTAINS InterPro DOMAIN/s: EGF-like	ss715606800
LIM01 & SSTA	Glyma.13G312800	13	40797465	40799085	plant U-box 23	ss715616244
LIM01 & SSTA	Glyma.13G313200	13	40863551	40865337	Tetratricopeptide repeat (TPR)-like superfamily protein	ss715616244
LIM01 & SSTA	Glyma.13G314400	13	40977943	40978908	unknown protein	ss715616244
LIM01 & SSTA	Glyma.13G315300	13	41041851	41044605	NAC domain containing protein 90	ss715616244
LIM01 & SSTA	Glyma.15G050800	15	3997913	4001197	Peroxidase superfamily protein	ss715621861
LIM01 & SSTA	Glyma.15G051500	15	4044021	4048207	beta-D-xylosidase 4	ss715621861
LIM01 & SSTA	Glyma.15G052600	15	4143223	4145151	Peroxidase superfamily protein	ss715621861
LIM01 & SSTA	Glyma.15G052700	15	4147574	4149546	Peroxidase superfamily protein	ss715621861
LIM01 & SSTA	Glyma.15G053100	15	4173451	4179144	pumilio 7	ss715621861
LIM01 & SSTA	Glyma.15G053700	15	4224743	4229800	Protein phosphatase 2C family protein	ss715621861
LIM01 & SSTA	Glyma.15G054500	15	4274887	4276923	UDP-glucosyl transferase 85A2	ss715621861
LIM01 & SSTA	Glyma.15G054600	15	4295527	4296772	EXORDIUM like 2	ss715621861
LIM01 & SSTA	Glyma.20G128600	20	37015598	37017590	GroES-like zinc-binding alcohol dehydrogenase family protein	ss715637816
LIM01 & SSTA	Glyma.20G129900	20	37091399	37092697	Plant basic secretory protein (BSP) family protein	ss715637816
LIM01 & SSTA	Glyma.20G130000	20	37095050	37095962	Plant basic secretory protein (BSP) family protein	ss715637816
LIM01 & SSTA	Glyma.10G162400	10	39661091	39664486	indoleacetic acid-induced protein 16	ss715606800
LIM01 & SSTA	Glyma.13G312700	13	40788194	40790353	plant U-box 23	ss715616244
LIM01 & SSTA	Glyma.15G052100	15	4114405	4115605	Unknown Protein	ss715621861
LIM01 & SSTA	Glyma.20G129600	20	37075297	37078272	stress enhanced protein 1	ss715637816
LIM01 & SSTA	Glyma.20G132300	20	37263193	37263944	Uncharacterised protein family	ss715637816

					SERF	
LIM01 & SSTA	Glyma.10G163300	10	39725270	39726261	winged-helix DNA-binding transcription factor family protein	ss715606800
LIM01 & SSTA	Glyma.10G163400	10	39727331	39731534	TPX2 (targeting protein for Xklp2) protein family	ss715606800
LIM01 & SSTA	Glyma.13G314200	13	40949752	40951896	unknown protein	ss715616244
LIM01 & SSTA	Glyma.13G316000	13	41077356	41081233	Rho GTPase activating protein with PAK-box/P21-Rho-binding domain	ss715616244
LIM01 & SSTA	Glyma.15G053900	15	4241010	4243926	alpha/beta-Hydrolases superfamily protein	ss715621861
LIM01 & SSTA	Glyma.18G076200	18	7212910	7215902	Concanavalin A-like lectin protein kinase family protein	ss715632608
LIM01 & SSTA	Glyma.20G132000	20	37245101	37248599	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	ss715637816
LIM01 & SSTA	Glyma.20G133100	20	37310238	37316817	Phototropic-responsive NPH3 family protein	ss715637816
LIM01 & SSTA	Glyma.10G161000	10	39522164	39526626	N-MYC downregulated-like 2	ss715606800
LIM01 & SSTA	Glyma.10G161400	10	39557553	39563940	NIMA-related serine/threonine kinase 1	ss715606800
LIM01 & SSTA	Glyma.10G161900	10	39612743	39613906	calmodulin 8	ss715606800
LIM01 & SSTA	Glyma.10G163500	10	39742349	39744332	Alba DNA/RNA-binding protein	ss715606800
LIM01 & SSTA	Glyma.13G313700	13	40902956	40904226	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	ss715616244
LIM01 & SSTA	Glyma.13G315400	13	41047454	41049093	Pollen Ole e 1 allergen and extensin family protein	ss715616244
LIM01 & SSTA	Glyma.15G050300	15	3968865	3971243	cytochrome P450, family 71, subfamily B, polypeptide 37	ss715621861
LIM01 & SSTA	Glyma.15G051900	15	4089702	4092390	Chloroplast-targeted copper chaperone protein	ss715621861
LIM01 & SSTA	Glyma.15G052200	15	4115777	4117976	Bifunctional inhibitor/lipid-transfer	ss715621861

					protein/seed storage 2S albumin superfamily protein	
LIM01 & SSTA	Glyma.18G073500	18	6971929	6982758	kinesin-like protein 1	ss715632608
LIM01 & SSTA	Glyma.18G073600	18	6992311	6996296	Leucine-rich receptor-like protein kinase family protein	ss715632608
LIM01 & SSTA	Glyma.18G076600	18	7283307	7285852	FASCICLIN-like arabinogalactan 1	ss715632608
LIM01 & SSTA	Glyma.18G077000	18	7308150	7308773	Unknown Protein	ss715632608
LIM01 & SSTA	Glyma.20G129200	20	37049072	37052666	Leucine-rich repeat protein kinase family protein	ss715637816
LIM01 & SSTA	Glyma.20G130300	20	37112131	37113717	unknown protein	ss715637816
LIM01 & SSTA	Glyma.20G130400	20	37120058	37135456	1,2-alpha-L-fucosidases	ss715637816
LIM01 & SSTA	Glyma.20G131100	20	37178766	37179518	Gibberellin-regulated family protein	ss715637816
LIM01 & SSTA	Glyma.20G132200	20	37257212	37261238	unknown protein	ss715637816
LIM01	Glyma.03G019500	3	2000006	2006209	glutaredoxin-related	ss715584729
LIM01	Glyma.03G197900	3	40721164	40723468	NAC domain containing protein 90	ss715586174
LIM01	Glyma.03G199200	3	40816015	40816757	Unknown Protein	ss715586174
LIM01	Glyma.05G022600	5	1980136	1985661	UDP-glucose 6-dehydrogenase family protein	ss715590348
LIM01	Glyma.05G023500	5	2045417	2048973	Plant calmodulin-binding protein-related	ss715590348
LIM01	Glyma.05G024000	5	2085852	2090902	Leucine-rich repeat protein kinase family protein	ss715590348
LIM01	Glyma.06G272800	6	46354602	46356277	senescence-associated gene 12	ss715594534
LIM01	Glyma.11G145200	11	11158223	11163437	glycosyltransferase family protein 2	ss715609438
LIM01	Glyma.15G254100	15	48310048	48312110	Ribosomal protein L1p/L10e family	ss715622335
LIM01	Glyma.15G256200	15	48540267	48541595	unknown protein	ss715622335
LIM01	Glyma.15G256300	15	48546007	48547401	unknown protein	ss715622335
LIM01	Glyma.16G203400	16	36449021	36451535	HIT-type Zinc finger family protein	ss715624869
LIM01	Glyma.16G204600	16	36552157	36556546	Enolase	ss715624869
LIM01	Glyma.16G205500	16	36627840	36631950	Arabidopsis thaliana DUF794)	ss715624869

LIM01	Glyma.03G021400	3	2202325	2234931	WD-40 repeat family protein / beige-related	ss715584729
LIM01	Glyma.12G044700	12	3247145	3249958	polyol/monosaccharide transporter 5	ss715612322
LIM01	Glyma.12G046800	12	3402230	3403241	unknown protein	ss715612322
LIM01	Glyma.12G047600	12	3446040	3446807	Unknown Protein	ss715612322
LIM01	Glyma.15G253400	15	48201791	48205539	blue-copper-binding protein	ss715622335
LIM01	Glyma.15G254000	15	48295488	48300322	NAC domain containing protein 1	ss715622335
LIM01	Glyma.03G019900	3	2050178	2051816	D-isomer specific 2-hydroxyacid dehydrogenase family protein	ss715584729
LIM01	Glyma.03G020400	3	2078755	2084370	cytochrome P450, family 704, subfamily A, polypeptide 2	ss715584729
LIM01	Glyma.03G021200	3	2170567	2172891	cytochrome P450, family 76, subfamily C, polypeptide 4	ss715584729
LIM01	Glyma.03G021600	3	2240686	2246512	cytochrome P450, family 704, subfamily A, polypeptide 2	ss715584729
LIM01	Glyma.03G022000	3	2296902	2299177	Pentatricopeptide repeat (PPR) superfamily protein	ss715584729
LIM01	Glyma.03G197200	3	40664580	40665418	SAUR-like auxin-responsive protein family	ss715586174
LIM01	Glyma.03G201100	3	40972049	40972976	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	ss715586174
LIM01	Glyma.03G201600	3	40993013	40994536	NDR1/HIN1-like 1	ss715586174
LIM01	Glyma.05G021900	5	1917507	1920054	Cytochrome P450 superfamily protein	ss715590348
LIM01	Glyma.05G022100	5	1929176	1933096	Cytochrome P450 superfamily protein	ss715590348
LIM01	Glyma.05G023100	5	2020400	2023401	soluble N-ethylmaleimide-sensitive factor adaptor protein 33	ss715590348
LIM01	Glyma.06G271600	6	46147876	46153766	high-affinity K+ transporter 1	ss715594534
LIM01	Glyma.06G272900	6	46369984	46370970	Cysteine proteinases superfamily protein	ss715594534

LIM01	Glyma.06G273600	6	46449321	46451017	senescence-associated gene 12	ss715594534
LIM01	Glyma.11G144500	11	11040836	11042646	Chloroplast-targeted copper chaperone protein	ss715609438
LIM01	Glyma.12G044600	12	3238665	3241765	unknown seed protein like 1	ss715612322
LIM01	Glyma.12G045300	12	3289163	3291323	Quinone reductase family protein	ss715612322
LIM01	Glyma.12G047900	12	3473688	3479089	extra-large G-protein 1	ss715612322
LIM01	Glyma.12G048000	12	3480508	3481344	unknown protein	ss715612322
LIM01	Glyma.12G049100	12	3528884	3530440	homolog of carrot EP3-3 chitinase	ss715612322
LIM01	Glyma.12G049300	12	3538428	3539170	unknown protein	ss715612322
LIM01	Glyma.12G049700	12	3564430	3573107	P-loop containing nucleoside triphosphate hydrolases superfamily protein	ss715612322
LIM01	Glyma.15G253500	15	48215975	48220366	Protein kinase superfamily protein	ss715622335
LIM01	Glyma.15G254800	15	48396968	48399125	Pentatricopeptide repeat (PPR) superfamily protein	ss715622335
LIM01	Glyma.15G255800	15	48508359	48513281	Lojap-related protein	ss715622335
LIM01	Glyma.16G204800	16	36583127	36586048	Tetratricopeptide repeat (TPR)-like superfamily protein	ss715624869
LIM01	Glyma.16G206000	16	36656248	36658633	Pentatricopeptide repeat (PPR) superfamily protein	ss715624869
LIM01	Glyma.05G020400	5	1821664	1825334	ABI five binding protein 3	ss715590348
LIM01	Glyma.05G021100	5	1871904	1878678	respiratory burst oxidase protein F	ss715590348
LIM01	Glyma.16G203100	16	36416105	36417800	2 iron, 2 sulfur cluster binding	ss715624869
LIM01	Glyma.03G021300	3	2196188	2197852	Family of unknown function (DUF716)	ss715584729
LIM01	Glyma.03G197800	3	40709426	40711214	DUF584	ss715586174
LIM01	Glyma.12G048400	12	3497445	3499204	Family of unknown function (DUF716)	ss715612322
LIM01	Glyma.16G206400	16	36670972	36674446	plastid movement impaired1	ss715624869
LIM01	Glyma.03G019700	3	2010835	2018418	galacturonosyltransferase 1	ss715584729

LIM01	Glyma.03G021700	3	2248429	2251737	ataurora3	ss715584729
LIM01	Glyma.03G201200	3	40977183	40977806	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	ss715586174
LIM01	Glyma.03G201300	3	40980745	40990121	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	ss715586174
LIM01	Glyma.03G201800	3	41008486	41009267	damaged DNA binding;DNA-directed DNA polymerases	ss715586174
LIM01	Glyma.06G272200	6	46282101	46287146	fatty alcohol oxidase 3	ss715594534
LIM01	Glyma.11G145400	11	11173794	11178803	GATA transcription factor 11	ss715609438
LIM01	Glyma.16G203300	16	36437330	36445131	ERECTA-like 1	ss715624869
LIM01	Glyma.03G019000	3	1895988	1898974	subtilase family protein	ss715584729
LIM01	Glyma.03G197300	3	40669086	40675486	P-loop containing nucleoside triphosphate hydrolases superfamily protein	ss715586174
LIM01	Glyma.03G198600	3	40779817	40780824	F-box family protein	ss715586174
LIM01	Glyma.03G199800	3	40872786	40874231	unknown protein	ss715586174
LIM01	Glyma.03G200100	3	40897609	40898945	ovate family protein 13	ss715586174
LIM01	Glyma.05G020600	5	1836095	1841454	RPA70-kDa subunit B	ss715590348
LIM01	Glyma.05G023700	5	2057228	2060045	Flavin-binding monooxygenase family protein	ss715590348
LIM01	Glyma.05G023800	5	2066114	2068799	Flavin-binding monooxygenase family protein	ss715590348
LIM01	Glyma.05G023900	5	2079928	2080918	CONTAINS InterPro DOMAIN/s: NTP Pyrophosphohydrolase MazG-related, RS21-C6	ss715590348
LIM01	Glyma.06G271400	6	46134359	46138237	Unknown Protein	ss715594534
LIM01	Glyma.11G143100	11	10924084	10925772	tonoplast intrinsic protein 1;3	ss715609438
LIM01	Glyma.11G144600	11	11057446	11059997	exocyst subunit exo70 family protein H7	ss715609438

LIM01	Glyma.11G144800	11	11096950	11105343	Protein kinase superfamily protein	ss715609438
LIM01	Glyma.11G145100	11	11144477	11150703	Putative lysine decarboxylase family protein	ss715609438
LIM01	Glyma.12G043800	12	3175262	3180077	O-Glycosyl hydrolases family 17 protein	ss715612322
LIM01	Glyma.12G047400	12	3430200	3438358	long-chain acyl-CoA synthetase 2	ss715612322
LIM01	Glyma.12G048800	12	3515939	3520048	unknown protein	ss715612322
LIM01	Glyma.12G049200	12	3533167	3535182	homolog of carrot EP3-3 chitinase	ss715612322
LIM01	Glyma.15G256000	15	48531279	48532198	unknown protein	ss715622335
LIM01	Glyma.15G256400	15	48553974	48559510	ATP binding microtubule motor family protein	ss715622335
LIM01	Glyma.16G206300	16	36666141	36667442	Unknown Protein	ss715624869
LIM01	Glyma.16G206900	16	36686943	36688802	ENTH/ANTH/VHS superfamily protein	ss715624869
SSTA	Glyma.06G155100	6	12675063	12679263	Copine (Calcium-dependent phospholipid-binding protein) family	ss715592934
SSTA	Glyma.10G153600	10	38863200	38865413	Unknown Protein	ss715606653
SSTA	Glyma.10G154500	10	38951834	38954460	rRNA processing protein-related	ss715606653
SSTA	Glyma.12G173700	12	33046449	33054373	beta-galactosidase 8	ss715612259
SSTA	Glyma.16G152800	16	31347028	31354847	Pentatricopeptide repeat (PPR) superfamily protein	ss715624393
SSTA	Glyma.16G153400	16	31393061	31398041	glutamate dehydrogenase 1	ss715624393
SSTA	Glyma.20G103600	20	34635685	34636903	Unknown Protein	ss715637471-ss715637482
SSTA	Glyma.20G104600	20	34682410	34684475	nodulin MtN21 /EamA-like transporter family protein	ss715637471-ss715637482
SSTA	Glyma.06G154100	6	12565656	12566840	plant U-box 25	ss715592934
SSTA	Glyma.06G155000	6	12642936	12647549	gibberellin 20-oxidase 3	ss715592934
SSTA	Glyma.10G153300	10	38809874	38814159	unknown protein	ss715606653
SSTA	Glyma.16G151300	16	31181365	31186281	Aquaporin-like superfamily protein	ss715624393

SSTA	Glyma.16G151800	16	31225865	31230698	Dynein light chain type 1 family protein	ss715624393
SSTA	Glyma.16G154900	16	31508874	31509401	Expressed protein	ss715624393
SSTA	Glyma.16G193900	16	35617977	35621903	disease resistance family protein / LRR family protein	ss715624785
SSTA	Glyma.16G197300	16	35870607	35871520	Unknown Protein	ss715624785
SSTA	Glyma.17G218500	17	36908440	36910940	Concanavalin A-like lectin protein kinase family protein	ss715627202
SSTA	Glyma.20G103800	20	34643077	34647277	Tetratricopeptide repeat (TPR)-like superfamily protein	ss715637471-ss715637482
SSTA	Glyma.06G153600	6	12533507	12538895	golgi nucleotide sugar transporter 1	ss715592934
SSTA	Glyma.06G154600	6	12602378	12605254	Pentatricopeptide repeat (PPR-like) superfamily protein	ss715592934
SSTA	Glyma.06G154700	6	12606286	12607961	embryo defective 2170	ss715592934
SSTA	Glyma.10G150400	10	38573748	38575248	phloem protein 2-B15	ss715606653
SSTA	Glyma.10G152200	10	38696013	38703000	respiratory burst oxidase homolog B	ss715606653
SSTA	Glyma.10G153900	10	38898358	38904608	lipoxygenase 1	ss715606653
SSTA	Glyma.10G163800	10	39770124	39770911	Unknown Protein	ss715606833
SSTA	Glyma.10G164000	10	39773910	39774726	unknown protein	ss715606833
SSTA	Glyma.12G173800	12	33069515	33070786	FASCICLIN-like arabinogalactan-protein 11	ss715612259
SSTA	Glyma.12G175800	12	33348849	33352707	alpha carbonic anhydrase 7	ss715612259
SSTA	Glyma.13G079200	13	18577742	18580774	unknown protein	ss715615362
SSTA	Glyma.16G151500	16	31198908	31201131	NAC domain containing protein 47	ss715624393
SSTA	Glyma.16G153000	16	31363737	31366379	NDH-dependent cyclic electron flow 5	ss715624393
SSTA	Glyma.16G153700	16	31431020	31433072	unknown protein	ss715624393
SSTA	Glyma.16G195600	16	35719546	35722126	cytochrome P450, family 71, subfamily A, polypeptide 26	ss715624785
SSTA	Glyma.16G196000	16	35747398	35749949	Unknown Protein	ss715624785

SSTA	Glyma.16G196100	16	35749107	35749458	Unknown Protein	ss715624785
SSTA	Glyma.17G219200	17	37082064	37085069	Protein kinase superfamily protein	ss715627202
SSTA	Glyma.17G220100	17	37197689	37204991	Pentatricopeptide repeat (PPR) superfamily protein	ss715627202
SSTA	Glyma.20G113100	20	35508342	35512020	Major facilitator superfamily protein	ss715637615
SSTA	Glyma.20G113700	20	35561605	35563401	Unknown Protein	ss715637615
SSTA	Glyma.20G114200	20	35592130	35594408	cinnamate-4-hydroxylase	ss715637615
SSTA	Glyma.20G114400	20	35614355	35616709	unknown protein	ss715637615
SSTA	Glyma.10G164100	10	39822153	39825768	SSXT family protein	ss715606833
SSTA	Glyma.20G103500	20	34630769	34632042	Unknown Protein	ss715637471-ss715637482
SSTA	Glyma.16G193600	16	35576270	35580207	disease resistance family protein / LRR family protein	ss715624785
SSTA	Glyma.10G151100	10	38607998	38608397	Unknown Protein	ss715606653
SSTA	Glyma.20G115600	20	35746760	35748974	B-box type zinc finger protein with CCT domain	ss715637615
SSTA	Glyma.06G153000	6	12486501	12489369	PRLI-interacting factor, putative	ss715592934
SSTA	Glyma.10G153100	10	38772208	38774371	Photosystem II reaction center PsbP family protein	ss715606653
SSTA	Glyma.12G174100	12	33132698	33136308	auxin response factor 16	ss715612259
SSTA	Glyma.16G152000	16	31235374	31237966	Unknown Protein	ss715624393
SSTA	Glyma.16G194500	16	35666818	35671212	Chromosome transmission fidelity protein 8	ss715624785
SSTA	Glyma.20G105300	20	34757381	34771672	ACT-like protein tyrosine kinase family protein	ss715637471-ss715637482
SSTA	Glyma.20G114100	20	35586177	35589284	Aluminium activated malate transporter family protein	ss715637615
SSTA	Glyma.06G153800	6	12543567	12544601	RmlC-like cupins superfamily protein	ss715592934
SSTA	Glyma.10G159000	10	39295428	39305050	Transducin/WD40 repeat-like superfamily protein	ss715606779

SSTA	Glyma.10G164700	10	39884140	39887915	Leucine-rich receptor-like protein kinase family protein	ss715606833
SSTA	Glyma.16G153600	16	31420465	31421069	arabinogalactan protein 14	ss715624393
SSTA	Glyma.16G154200	16	31463230	31469750	syntaxin of plants 131	ss715624393
SSTA	Glyma.16G192800	16	35494129	35497008	disease resistance family protein / LRR family protein	ss715624785
SSTA	Glyma.16G193000	16	35515610	35518426	disease resistance family protein / LRR family protein	ss715624785
SSTA	Glyma.16G195900	16	35737872	35741786	Tetratricopeptide repeat (TPR)-like superfamily protein	ss715624785
SSTA	Glyma.16G197800	16	35897985	35901946	phloem protein 2-A13	ss715624785
SSTA	Glyma.17G218900	17	37041468	37047674	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain	ss715627202
SSTA	Glyma.20G102900	20	34579855	34593297	myosin 1	ss715637471-ss715637482
SSTA	Glyma.20G112600	20	35468911	35474203	Pectin lyase-like superfamily protein	ss715637615
SSTA	Glyma.20G115500	20	35732477	35735054	3-ketoacyl-CoA synthase 6	ss715637615
SSTA	Glyma.06G152000	6	12384370	12387657	Plant invertase/pectin methylesterase inhibitor superfamily	ss715592934
SSTA	Glyma.06G154400	6	12584515	12585942	xylem NAC domain 1	ss715592934
SSTA	Glyma.06G154500	6	12600307	12602169	ABC-2 type transporter family protein	ss715592934
SSTA	Glyma.10G151000	10	38606877	38609056	Dormancy/auxin associated family protein	ss715606653
SSTA	Glyma.10G152000	10	38682900	38684549	rapid alkalinization factor 1	ss715606653
SSTA	Glyma.10G152300	10	38716290	38721057	Actin-binding FH2 (formin homology 2) family protein	ss715606653
SSTA	Glyma.10G154100	10	38926113	38928738	cyclin family	ss715606653
SSTA	Glyma.10G154200	10	38930043	38934295	P-loop containing nucleoside triphosphate hydrolases superfamily protein	ss715606653

SSTA	Glyma.13G079900	13	18635427	18644868	long chain acyl-CoA synthetase 9	ss715615362
SSTA	Glyma.13G080000	13	18636522	18637286	Unknown Protein	ss715615362
SSTA	Glyma.13G080500	13	18710722	18713639	Adenine nucleotide alpha hydrolases-like superfamily protein	ss715615362
SSTA	Glyma.16G151200	16	31173115	31175417	syntaxin of plants 111	ss715624393
SSTA	Glyma.16G151900	16	31232074	31238858	O-Glycosyl hydrolases family 17 protein	ss715624393
SSTA	Glyma.16G152600	16	31324673	31328127	O-Glycosyl hydrolases family 17 protein	ss715624393
SSTA	Glyma.16G152700	16	31330128	31332137	GATA transcription factor 9	ss715624393
SSTA	Glyma.16G155000	16	31513389	31517035	plasma membrane intrinsic protein 2	ss715624393
SSTA	Glyma.16G155100	16	31522889	31524889	plasma membrane intrinsic protein 2	ss715624393
SSTA	Glyma.17G219000	17	37048856	37050044	Unknown Protein	ss715627202
SSTA	Glyma.17G219700	17	37144552	37145322	Integrase-type DNA-binding superfamily protein	ss715627202
SSTA	Glyma.20G112400	20	35458398	35459304	ovate family protein 13	ss715637615
SSTA	Glyma.20G114900	20	35664672	35665579	Unknown Protein	ss715637615
SSTA	Glyma.20G115900	20	35771953	35773591	unknown protein	ss715637615

Appendix 2C: List of colocalized regions resistance to multiple diseases and abiotic stress by comparing GWAS

SNP Marker	Chr	Position (bp)	Disease	Published research papers
ss715590348	GM05	1,952,710	Targetspot-LIM01	present study
Iron deficiency chlorosis 3-g2	Gm05	2,612,793	Iron Deficiency Chlorosis	(Mamidi et al. 2011)
ss715592934	GM06	12,584,572	Targetspot-SSTA	present study
	Gm06	13,578,142	sudden death syndrome	(Bao *et al.* 2015)
Iron deficiency chlorosis 2-g5	Gm06	45,851,263	Iron Deficiency Chlorosis	(Mamidi et al. 2011)
ss715594534	GM06	46,295,298	Targetspot-LIM01	present study
SCN 5-g14	Gm06	46,820,673	*Heterodera glycines*	(Li *et al.* 2016)
ss715606653	GM10	38,753,114	Targetspot-SSTA	present study
Phytoph 2-g47	Gm10	39,347,529	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g48	Gm10	39,354,992	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g49	Gm10	39,358,965	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g50	Gm10	39,364,212	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g51	Gm10	39,365,693	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g52	Gm10	39,376,495	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g53	Gm10	39,394,332	*Phytophthora sojae*	(QIN *et al.* 2017)
ss715606779	GM10	39,456,010	Targetspot-SSTA	present study
ss715606793	GM10	39,500,426	Targetspot-SSTA	present study
ss715606796	GM10	39,521,125	Targetspot-SSTA	present study
ss715606800	GM10	39,534,982	Targetspot-both	present study
ss715606801	GM10	39,543,852	Targetspot-SSTA	present study
ss715606817	GM10	39,666,999	Targetspot-SSTA	present study
ss715606822	GM10	39,670,594	Targetspot-SSTA	present study
ss715606823	GM10	39,682,612	Targetspot-SSTA	present study
ss715606824	GM10	39,687,502	Targetspot-SSTA	present study
ss715606829	GM10	39,709,662	Targetspot-SSTA	present study
ss715606832	GM10	39,714,825	Targetspot-SSTA	present study
ss715606833	GM10	39,716,038	Targetspot-SSTA	present study
Phytoph 2-g18	Gm10	39,904,244	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g19.1	Gm10	39,911,607	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g19.2	Gm10	39,915,580	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g19.3	Gm10	39,920,827	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g20.1	Gm10	39,922,308	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g20.2	Gm10	39,933,110	*Phytophthora sojae*	(QIN *et al.* 2017)
Phytoph 2-g21	Gm10	39,950,947	*Phytophthora sojae*	(QIN *et al.* 2017)

Reniform nematode 1-g1	Gm10	40,672,699	*Reniform nematode*	(Chang *et al.* 2016)
SCN 1-g8	Gm10	40,672,699	*Heterodera glycines*	(Chang *et al.* 2016)
SCN 4-g8	Gm11	10,206,223	*Heterodera glycines*	(VUONG *et al.* 2015)
ss715609438	GM11	11,087,536	Targetspot-LIM01	present study
SCN 5-g23	Gm11	11,368,576	*Heterodera glycines*	(Li *et al.* 2016)
Sclero 3-g5	Gm12	32,434,240	*Sclerotinia sclerotiorum*	(Moellers *et al.* 2017)
Sclero 3-g23	Gm12	32,727,465	*Sclerotinia sclerotiorum*	(Moellers *et al.* 2017)
ss715612259	GM12	33,175,894	Targetspot-SSTA	present study
SDS 1-g36	Gm13	17,285,679	sudden death syndrome	(Wen *et al.* 2014)
Sclero 4-g6	Gm13	17,625,098	*Sclerotinia sclerotiorum*	(ZHAO *et al.* 2015)
Sclero 4-g7	Gm13	18,422,904	*Sclerotinia sclerotiorum*	(ZHAO *et al.* 2015)
Sclero 4-g4	Gm13	18,572,456	*Sclerotinia sclerotiorum*	(ZHAO *et al.* 2015)
Sclero 4-g5	Gm13	18,572,787	*Sclerotinia sclerotiorum*	(Zhao *et al.* 2015)
Sclero 3-g54	Gm13	18,644,670	*Sclerotinia sclerotiorum*	(Moellers *et al.* 2017)
ss715615362	GM13	18,693,392	Targetspot-SSTA	present study
Sclero 3-g35	Gm16	35,186,332	*Sclerotinia sclerotiorum*	(Moellers *et al.* 2017)
ss715624785	GM16	35,789,070	Targetspot-SSTA	present study
SCN 5-g38	Gm16	36,221,174	*Heterodera glycines*	(Li *et al.* 2016)
ss715624869	GM16	36,571,566	Targetspot-LIM01	present study
Iron deficiency chlorosis 3-g10	Gm16	37,046,875	Iron Deficiency Chlorosis	(Mamidi et al. 2011)
Phytoph 1-g3	Gm17	36,718,722	*Phytophthora sojae*	(Sun *et al.* 2014)
Yellow Mosaic Virus 1-g1	Gm17	36,718,722	*yellow mosaic virus (YMV)*	(Kumar *et al.* 2015)
SCN 5-g42	Gm17	36,977,039	*Heterodera glycines*	(Li *et al.* 2016)
ss715627202	GM17	37,097,907	Target spot-SSTA	present study
Yellow Mosaic Virus 1-g2	Gm17	37,799,168	*yellow mosaic virus (YMV)*	(Kumar *et al.* 2015)
ss715632608	GM18	7,161,394	Target spot-both	present study
SCN 2-g2	Gm18	7,450,433	*Heterodera glycines*	(Mamidi et al. 2011)
Sclero 3-g28	Gm18	7,929,040	Sclerotinia sclerotiorum	(Moellers *et al.* 2017)
SCN 4-g16	Gm20	34,357,030	*Heterodera glycines*	(Vuong *et al.* 2015)
ss715637471	GM20	34,699,114	Targetspot-SSTA	present study

ss715637481	GM20	34,776,314	Targetspot-SSTA	present study
ss715637482	GM20	34,795,864	Targetspot-SSTA	present study
Sclero 3-g7	Gm20	35,150,084	*Sclerotinia sclerotiorum*	(Moellers *et al.* 2017)
ss715637615	GM20	35,651,832	Targetspot-SSTA	present study

References

Bao, Y., J. E. Kurle, G. Anderson and N. D. Young, 2015 Association mapping and genomic prediction for resistance to sudden death syndrome in early maturing soybean germplasm. Molecular Breeding 35**:** 1-14.

Chang, H.-X., A. E. Lipka, L. L. Domier and G. L. Hartman, 2016 Characterization of disease resistance loci in the USDA soybean germplasm collection using genome-wide association studies. Phytopathology 106**:** 1139-1151.

Kumar, B., A. Talukdar, K. Verma, I. Bala, G. Harish *et al.*, 2015 Mapping of yellow mosaic virus (*YMV*) resistance in soybean (*Glycine max* L. Merr.) through association mapping approach. Genetica 143**:** 1-10.

Li, Y. h., X. h. Shi, H. h. Li, J. C. Reif, J. j. Wang *et al.*, 2016 Dissecting the genetic basis of resistance to soybean cyst nematode combining linkage and association mapping. The Plant Genome 9**:** Plantgenome2015.2004.0020.

Mamidi, S., S. Chikara, R. J. Goos, D. L. Hyten, D. Annam *et al.*, 2011 Genome-wide association analysis identifies candidate genes associated with iron deficiency chlorosis in soybean. The Plant Genome 4.

Moellers, T. C., A. Singh, J. Zhang, J. Brungardt, M. Kabbage *et al.*, 2017 Main and epistatic loci studies in soybean for *Sclerotinia sclerotiorum* resistance reveal multiple modes of resistance in multi-environments. Scientific Reports 7**:** 1-13.

Qin, J., Q. Song, A. Shi, S. Li, M. Zhang *et al.*, 2017 Genome-wide association mapping of resistance to *Phytophthora sojae* in a soybean [*Glycine max* (L.) Merr.] germplasm panel from maturity groups IV and V. PLOS One 12**:** e0184613.

Sun, J., N. Guo, J. Lei, L. Li, G. Hu *et al.*, 2014 Association mapping for partial resistance to *Phytophthora sojae* in soybean (*Glycine max* (L.) Merr.). Journal of Genetics 93**:** 355-363.

Vuong, T., H. Sonah, C. Meinhardt, R. Deshmukh, S. Kadam *et al.*, 2015 Genetic architecture of cyst nematode resistance revealed by genome-wide association study in soybean. BMC Genomics 16**:** 1-13.

Wen, Z., R. Tan, J. Yuan, C. Bales, W. Du *et al.*, 2014 Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. BMC Genomics 15**:** 1-11.

Zhao, X., Y. Han, Y. Li, D. Liu, M. Sun *et al.*, 2015 Loci and candidate gene identification for resistance to *Sclerotinia sclerotiorum* in soybean (*Glycine max* L. Merr.) via association and linkage maps. The Plant Journal 82**:** 245-255.