PARALLEL ANALYSIS OF ASIAN SOYBEAN RUST VARIABILITY, CANDIDATE SOYBEAN RESISTANCE GENES AND SOYBEAN RUST EFFECTORS

by

PENG TIAN

(Under the Direction of Shavannor M. Smith)

ABSTRACT

Obligate basdiomycete *Phakopsora pachyrhizi*, is the causal agent of Asian soybean rust (ASR) and is one of the most destructive foliar soybean diseases. It is an exotic pathogen that first arrived in the United States mainland in 2004, and has the potential to cause severe yield loss. ASR is currently managed by fungicides. However, fungicide-resistant *P. pachyrhizi* are inevitable and the economic cost of continued fungicide use is not practical. Host resistance is therefore, considered to be the most economical and practical method for controlling ASR. To date, traditional breeding strategies generate soybean cultivars that are resistant to particular races of ASR. Once deployed, resistance is soon overcome by the ever changing pathogen. It is therefore necessary to characterize ASR pathogen variability and to define the genetics of resistance to ASR in soybean. In this study, eight *P. pachyrhizi* Georgia field isolates were characterized and grouped into three different isolates based on phenotype when inoculated on to a set of soybean differentials. Population genetic analysis of each ASR field isolate was also performed and demonstrated that the Georgia field isolates were diverse and were a mixture of genotypes from different countries and states boarding Georgia. One predominate isolate was identified in the Georgia fields isolates and one resistant Japanese cultivar was resistant to all of

the Georgia field isolates. RNA sequence analysis was performed on the predominate ASR isolate and a resistant cultivar to simultaneously identify soybean defense-related genes and ASR effector genes expressed during an incompatible interaction in the host and the pathogen. Nineteen NBS-LRR resistance genes, 24 *GmWRKY* genes, and 11 *MYB* genes were differentially expressed (> 2 fold change) in the compatible and incompatible interactions. A total of 1,786 putative fungal effector genes were identified in the pathogen. Twenty-six of these genes were similar to known secreted protein genes in *Melampsora Larici-populina*. These findings will have major implications for the characterization of resistance mechanisms that may be more durable and provide new sources of resistance to assist with the development of novel strategies to control ASR.

INDEX WORDS: Asian soybean rust, soybean, *Phakopsora pachyrhizi*, disease resistance gene, population structure, field isolates, defense-related gene, host-pathogen interactions, incompatible interaction, dual RNA sequencing (RNA-seq), fosmid, resistance gene homolog (RGH), nucleotide-binding site leucine-rich repeat (NBS-LRR) resistance gene, *GmWRKY*, *MYB* gene, effector, wild relative

PARALLEL ANALYSIS OF ASIAN SOYBEAN RUST VARIABILITY, CANDIDATE SOYBEAN RESISTANCE GENES AND SOYBEAN RUST EFFECTORS

by

PENG TIAN

B.Eng., Shandong Agricultural University, China, 2006

M.Sc., Chinese Academy of Agricultural Sciences, China, 2009

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

PARALLEL ANALYSIS OF ASIAN SOYBEAN RUST VARIABILITY, CANDIDATE SOYBEAN RESISTANCE GENES AND SOYBEAN RUST EFFECTORS

by

PENG TIAN

Major Professor:	Shavannor M. Smith
Committee:	H. Roger Boerma
	Robert C. Kemerait
	Xiaoyu Zhang

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

DEDICATION

To my beloved wife, Haoyu Xia and my dear son, Chris Tian.

ACKNOWLEDGEMENTS

I would like to express my great appreciation to my major professor Dr. Shavannor Smith for always being such a wonderful, encouraging, caring, patient, and trusting mentor! I am thankful for her unselfish love, enormous support, deep understanding and huge dedications of her time and efforts in guiding me through all kinds of obstacles and success. She truly helped me grow into a better student, scientist and leader. I will be always thankful and proud to once have her as my boss for my graduate school life. I also thank my committee members Dr. Roger Boerma, Dr. Bob Kemerait and Dr. Xiaoyu Zhang for their guidance and research advice.

Special thanks go to my lovely former lab-mates, Judy Gray, Dr. Suchitra Chavan and Courtney Wright for their immeasurable support and help in almost every single aspect of my research and my life at UGA.

I also would like to thank Dr. Philip Brannen, Dr. Harald Scherm, Dr. John Sherwood, Dr. Ron Walcott, Dr. Jean Woodward, Ansuya Jogi, Amy Savelle, Matt Seader, and Kisha Shelton for all their assistance in departmental business, coursework, greenhouse and lab experiments, as well as for their friendship during these years. Sincere thanks to all the graduate students in the Plant Pathology Department, past and present, especially Gong Chen, Abraham Fulmer, Dr. Imana Power, Dr. Sydney Everhart, Dr. Sara Thomas, and Dr. Tom Baldwin for their friendship.

Finally, I would like to thank all of my family, especially my parents and my sister for their indescribable love and support. At last, I give thanks to my wife and my son with all my heart, for making it today together with me.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES xii

LIST OF FIGURES xiv

CHAPTER

1 **INTRODUCTION AND LITERATURE REVIEW** 1

Purpose and Significance of Study 1

Literature Review 2

Project Objectives 33

Literature cited 34

2 **CHARACTERIZATION OF HOST RESISTANCE OF SOYBEAN AND PERENNIAL GLYCINE SPECIES AND GENETIC DIVERSITY OF ASIAN SOYBEAN RUST IN GEORGIA** 61

Abstract 62

Introduction 64

Materials and Methods 67

Results 72

Discussion 77

Literature cited 82

3 **IDENTIFICATION OF DIFFERENTIALLY EXPRESSED SOYBEAN DEFENSE-RELATED GENES AND ASIAN SOYBEAN RUST EFFECTOR GENES IN RESPONSE TO *P. PACHYRHIZI* INFECTION** 99

Abstract 100

Introduction 101

Materials and Methods 105

Results 111

Discussion 116

Literature cited 124

4 **STRUCTURE AND EVOLUTION OF NBS-LRR RESISTANCE GENES IN SOYBEAN AND PERENNIAL *GLYCINE* SPECIES** 151

Abstract 152

Introduction 153

Materials and Methods 156

Results 161

Discussion 167

Literature cited 173

5 **CONCLUSIONS** 193

Literature cited 196

LIST OF TABLES

Page

Table 2.1 Reaction types for nine soybean lines inoculated with eight ASR isolates collected from different locations in Georgia.....94

Table 2.2 Reaction types for candidate soybean cultivars inoculated with eight ASR isolates collected from different locations in Georgia. 95

Table 2.3 Reaction types for eight soybean wild relatives inoculated with eight ASR isolates collected from different locations in Georgia. 96

Table 2.4 Number of genotypes in ASR Georgia field isolates collected from different locations in Georgia based on sequence analysis of the Internal Transcribed Space (ITS) region……. 97

Table 2.5 Genetic diversity of Internal Transcribed Space (ITS) and the ADP-Ribosylation Factor gene isolated from 8 field isolates collected from different locations in Georgia..98

Table 3.1 Statistical results from RNA sequencing analysis 138

Table 3.2 Statistical results from RNA sequencing analysis for potential rust reads 139

Table 3.3 Statistical results of comparison of differentially expressed genes (DEGs) between inoculated resistant line (IR) vs mock resistant line (MR) and inoculated susceptible line (IS) vs mock susceptible line (MS).....140

Table 3.4 Identification of all up-regulated and down-regulated DEGs in inoculated resistant line (IR) and inoculated susceptible line (IS). 141

Table 3.5 NBS-LRR disease resistance genes identified in compatible (IS) and incompatible interaction (IR). 142
Table 3.6 Similarity between the amino acid sequences of NBS regions of NBS-LRR R-genes using paisewise alignment by BioEdit. 144
Table 3.7 Identification of soybean transcription factors WRKY gene from up-regulated and down-regulated DEGs in IR and IS. 145
Table 3.8 Identification of soybean transcription factor MYB genes from up-regulated and down-regulated DEGs in IR and IS. 146
Table 3.9 Identification of *P. pachyrhizi* genes encode for candidate secreted effector proteins (CSEPs). 148
Table 4.1 Degenerate and common primers used to screen a soybean fosmid library for NBS RGHs. 179
Table 4.2 Summary statistics for 16 soybean fosmids containing RGHs 182
Table 4.3 Structural comparison of the TIR regions of RHG proteins from fosmids and resistance proteins encoded by reference R-genes. 184
Table 4.4 Structural comparison of the NBS regions of RGH proteins from soybean fosmids and resistance proteins encoded by reference R-genes. 185
Table 4.5 Number of NBS and LRR nucleotide and amino acid sequences cloned from each soybean wild relative. 190

LIST OF FIGURES

Page

Figure 2.1 Disease rating system used to assess the resistance reactions of soybean lines and soybean wild relatives inoculated with ASR. 91

Figure 2.2 Phylogenetic tree of ITS and ARF genes present in Georgia ASR field isolates…….....93

Figure 3.1 Inoculation of Hyuuga (resistant line) and Williams 82 (susceptible line) with *P. pachyrhizi* 10 dpi (days post-inoculation).....136

Figure 3.2 Flow chart of experimental procedure to identify potential soybean defense-related genes and Asian Soybean Rust effector genes.....137

Figure 3.3 Phylogenetic tree of NBS regions of IR up-regulated NBS-LRR R-genes and resistance gene homologs (RGHs) characterized from fosmid analysis from Chapter 4……......143

Figure 3.4 Schematic of bioinformatics pipeline used to identify *P. pachyrhizi* putative secreted effector genes. 147

Figure 4.1 Location on the NBS-LRR class resistance genes that degenerate primers were designed from…......180

Figure 4.2 Gene structures of 12 fosmids containing NBS RGHs in soybean Bingnan cultivar……. 181

Figure 4.3 Structure of R-proteins encoded by soybean RGHs from fosmid 7, fosmid 14, fosmid 15 and fosmid 16.....183

Figure 4.4 Conserved motifs identified in the TIR, NBS and LRR domains in soybean RGH proteins encoded by RGHs from fosmid 7, 14, 15 and 16 and reference proteins N, SR1 and KR1. …. 187

Figure 4.5 Neighbor-Joining tree of the different domains from RGHs trees generated based on the alignment of amino acid sequences TIR regions (A), NBS regions (B) and LRR regions (C) of soybean RGHs and reference R-genes. 189

Figure 4.6 Two phylogenetic trees were generated based on the alignment of amino acid sequences of NBS regions (A) and LRR regions (B) of RGH clones of PI 505151……..192

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Purpose and Significance of Study

Disease susceptibility to Asian soybean rust (ASR), caused by the obligate fungal pathogen *Phakopsora pachyrhizi*, is one of most notable problems affecting soybean production. Host resistance is considered as the most economical and practical method for controlling this disease. However, none of disease resistance genes that have been identified are effective against all of the current soybean rust isolates. It has been shown that resistance is likely to be durable when the corresponding core effector protein has an important virulence function in the pathogen and is therefore conserved in the pathogen population.

In this study, we will characterize the population structure of *P. pachyrhizi* collected from different locations in Georgia, analyze the structure of NBS-LRR resistance gene loci in soybean and analyze the expression pattern of soybean defense-related genes and pathogen effector genes that may be responsible for the resistance against ASR.

The overall goal of this work was to characterize the population structure and variability of *P. pachyrhizi* isolates from Georgia, identify and characterize differentially expressed soybean defense-related genes, and ASR secreted proteins with the aim of identifying soybean R-genes that can potentially recognize conserved, expressed and functional ASR core effector genes. By integrating the host and pathogen dataset, the goal was to characterize soybean resistance against ASR that is potentially durable. This work will have major implications for the characterization

of resistance mechanisms that may be more durable and the development of novel strategies to control ASR.

Literature Review

History of Soybean

Cultivated soybean (*Glycine max* L. Merr.) originated from East Asia and belongs to the Fabaceae family. It is the world's most important oil seed crop and accounts for about 56% of the global oilseed production (Wilson 2008). Soybean has become a major crop in the United States, Brazil, China, and Argentina and is very important in many other countries. The demand for soybean continues to grow worldwide due to its use as an ingredient in various food, feed, and industrial products (Wilcox 2004).

It has been suggested that soybean was domesticated in the northeast areas of China. However, the history of domestication itself and how the domesticated soybean was disseminated is not fully understood due to longstanding errors and misconceptions in the historical records that were written in Chinese (Hymowitz and Shurtleff 2005; Hymowitz 2008). Based on the current information regarding the history of soybean, Hymowitz (2008) suggested that the domestication process took place during the Shang Dynasty (ca. 1600 BC to 1046 BC) and cultivated soybean was domesticated in the eastern half of Northern China during the Zhou Dynasty (ca. 1046 BC to 256 BC). The domesticated soybean was then disseminated into central and south China as well as the Korean peninsula by the first century AC (Hymowitz 1970; Hymowitz and Newell 1980). From the first century to the age of Discovery (15^{th} to 17^{th} century AD), soybean was introduced to many Asian countries such as Indonesia, Malaysia, Nepal, North India, Japan, Thailand, Vietnam, and the Philippines (Hymowitz 2008). As the farmers

continued to grow this crop in East Asia, they selected soybeans with certain traits and developed specific land races.

In terms of dissemination of the soybean into the western world, Hymowitz (2008) stated that soybean reached Europe quite late and it was found to be planted in the Jardin des Plantes in 1740 in Paris, France and in 1790 in the Royal Botanic Garden in Kew, England. In 1765, Samuel Bowen brought soybean from China to North America to manufacture soy sauce. In 1770, Benjamin Franklin sent soybean seeds to John Bertram who was a botanist. He planted the seed in his garden in Philadelphia (Smyth 1907). Dr. James Mease was the first person who used the word soybean in American literature in 1804. About 50 years later, John H. Lea planted soybean in 1851 in the Midwest near Alton, Illinois, where soybeans were disseminated and spread through the "Corn Belt" (Hymowitz 1987). However, the U.S. soybean production was not commercially established until 1915 when soybeans were first crushed for oil in Elizabeth City, North Carolina (Wilson 1987).

Soybean Uses and Production

Prior to 1941, soybean was grown primarily as a forage crop in the USA. However, the discovery of soybean as an important oilseed (38-44% protein and 18-23% oil) permanently changed the perception of soybean from forage to a seed crop. This significantly increased the need for more productive or agronomic types of soybean (Wilson 2004). Traditionally, soybean has been used to produce various soy foods such as soymilk, tofu, miso (fermented soybean paste), natto (soybean paste), tempeh (fermented soybean), and soy sauce. Currently, soybean seed is used for producing an even wider array of products, which facilitated the development of the U.S. soybean industry. Soybean oil is used mostly for human consumption as cooking oil as

well as margarine, shortenings, and salad dressing. Another important product of soybean seed is protein-containing meal. This product is used in poultry and swine production as a protein supplement, and more recently in the aquaculture industry. Both soybean oil and meal products are also used in industrial products including soap, cosmetics, plastics, inks, coatings, adhesives, and building materials (Wilson 2004).

During the past half century, soybean has emerged as the dominant oilseed in world trade and has been grown in about 50 countries. In 2013, the USDA (USDA, FAS 2013) estimated world soybean production at 284 million metric tons (MMT), accounting for about 56% of the total oilseed production. The USA is the largest producer of soybean in the world and produced 89.5 MMT of soybean, or 31.7% of the world total in 2013. Brazil and Argentina are the second (30.7%) and third (18.8%) largest producers of soybean in the world, respectively. As the use of soybean oil for human consumption and use of high-protein meal for animal feed continue to increase, soybean production has continued to expand worldwide.

In North America, the major areas of soybean production extend from the eastern part of the country, including the coastal areas of the Gulf of Mexico and north to southern Canada. Soybean is grown in 31 states in the USA The leading soybean production states are Illinois (12.6 MMT, or 14%), Iowa (11.2 MMT, or 13%), Minnesota (7.4 MMT, or 8%), Indiana (7.2 MMT, or 8%), Nebraska (6.9 MMT, or 8%), Ohio (5.9 MMT, or 7%), Missouri (5.4 MMT, or 6%), and South Dakota (5 MMT, or 6%) (Soy stats 2013, http://soystats.com). Much of the soybean production has shifted from the south to the northern and western states. The north central states produced over 77% of the total soybeans in the USA in 2013, while the southern states (Arkansas, Mississippi, Tennessee, Louisiana, South Carolina, Georgia, and Alabama) produced 12% (Soy stats 2013, http://soystats.com). The availability of agricultural fields to

plant soybean in the north-central states and the increasing challenges of disease and insects in the southeast USA are the two factors contributing to the geographical shift of soybean production (Orf 2010).

Taxonomy

There are two subgenera in the genus *Glycine*, *Glycine* Willd (perennial) and *Soja* (Moench) (Ratnaparkhe *et al.*, 2011). The subgenus *Soja* includes soybean and its wild annual progenitor, *G. soja* Sieb. and Zucc. Both species are diploid (2n=40) and are cross compatible. The subgenus *Glycine* contains 26 wild perennial species that are distributed in Australia and various South and West Pacific Islands. They grow in a wide range of agroclimatic regions of Australia and are very diverse genomically, morphologically, and cytologically (Chung and Singh 2008). For example, *G. tomentella* Hyata constitutes four cytotypes (2n=38, 40, 78, 80), and *G. hirticaulis* Tindale and Craven, and *G. tabacina* (Labill.) Benth. have accessions with 2n=40 and 80 chromosomes, respectively (Ratnaparkhe *et al*., 2011).

Soybean Diseases

Similar to other crop plants, there are various constraints affecting soybean production including pathogens, insects, adaption, and nodulation (Dita *et al*., 2006). Different pathogens, including fungi, nematodes, viruses, bacteria, and phytoplasmae can cause serious diseases in soybean limiting production (Grau *et al*., 2004). Currently, soybean cyst nematode (*Heterodera glycines* Ichinohe) is considered to be the most damaging soybean pathogen (Niblack *et al*., 2004).

However, more than 40 fungal pathogens have been reported to cause significant disease problems worldwide in soybean (Hartman *et al*., 1999). Different fungi that are pathogenic on

soybean may be restricted to a specific plant part, or infect several or all parts of the plant, resulting in different soybean diseases with various symptoms and signs. For examples, soilborne pathogens infecting soybean seeds and the root of stem of soybean seedlings include *Rhizoctonia solani* Kühn, *Phytophthora sojae* Kaufmann, *Macrophomina phaseolina* (Tassi) Goidanich, *Mycoleptodiscus terrestrus* (Gerd.) Ostazeski as well as several species of *Pythium* and *Fusarium*. Pathogens that cause foliar diseases include *Septoria glycines* Hemmi, *Cercospora kikuchii* (Matsumoto & Tomoyasu) M.W. Gardner, *Cercospora sojina* K. Hara., *Peronospora manshurica* (Naumov) Syd., *Microsphaera diffusa* Cooke & Peck., and *Phakospora pachyrhizi* Sydow. In addition, *Diaporthe phaseolorum* and *Phomopsis phaseoli* cause Pod and stem blight and Phomopsis seed decay. Sclerotinia stem rot of soybean (commonly called white mold) is caused by *Sclerotinia sclerotiorum* (Lib.) deBary (Grau *et al*., 2004). The fungal pathogens mentioned above vary greatly in their prevalence, frequency of occurrence, and agronomic loss which is highly influenced by disease management strategies, environmental conditions, and the resistance of soybean cultivars against the pathogen.

Asian soybean rust (ASR), caused by obligate fungal pathogen *Phakopsora pachyrhizi*, is a devastating disease of soybean, causing soybean yield losses ranging from 10% to 80% world-wide (Miles *et al*., 2011). This disease was present in Asia for many decades before it was first identified in Africa in 1996 and in South America in 2001. The first report of ASR in the continental United States was in November 2004 in Louisiana (Schneider *et al*., 2005). ASR rapidly spread to several states boarding Louisiana in subsequent years increasing the concern for soybean production in the United States. The deployment of resistant soybean cultivars is considered to be the most effective and economic method in controlling this disease. However, race specific resistance controlled by a single dominate resistance gene is often overcome by the

pathogen in the field, resulting in susceptibility (Burdon and Thrall 2006). In addition, very little is known about the genetic diversity of current prevalent *P. pachyrhizi* isolates in Georgia and the surrounding states. Thus, greater knowledge of the genetics of resistance to ASR and the variability of ASR populations is essential in order to dissect the relationship between the host and this pathogen and effectively breed for durable resistance.

Plant-Pathogen Interactions

Plants have developed a complicated and efficient defense system to detect the presence of pathogens and induce defense responses. These defense responses help the host plant to avoid further infection. Induction of the plant defense signaling involves direct or indirect recognition of specific pathogen effectors by the products of specialized host resistance genes (R-genes). This results in a hypersensitive reaction (HR) with plant localized cell death. More than 70 plant R-genes have been cloned and characterized in several plant species (Liu *et al*., 2007). Many of the R-genes are being effectively used in crop improvement research programs. Expression analysis, cloning, characterization, and genetic transformation of these genes will improve the ability to deploy efficient and sustained disease control methods against various pathogens.

To fight off pathogens in an efficient, rapid, and localized manner, plant innate immunity relies on a two-tier perception system associated with plasma membrane-localized and intracellular immune receptors (Dodds and Rathjen 2010; Zipfel 2014). The first tier of the plant immune system is mediated by extracellular surface pattern recognition receptors (PRRs) that are activated by microbe- or pathogen-associated molecular patterns (MAMPs/PAMPs), or self-molecules (damage-associated molecular patterns, DAMPs), resulting in PRR-triggered immunity (PTI) (Boller and Felix 2009; Macho and Zipfel 2014). A series of PTI events are

triggered with the perception of signals by the PRRs, including the rapid generation of reactive oxygen species (ROS), the expression of defense-related genes, and the activation of mitogen-activated protein kinases (MAPKs) (Boller and Felix 2009).

Due to the conserved nature of PAMPs (e.g., bacterial flagellin, fungal chitin), PTI is sufficient to ward off most non-adapted pathogens (Zipfel 2014). However, pathogens adapted to specific host plants are able to avoid and/or suppress PTI through the action of effector proteins secreted into the plant cells during the initial stage of infection. These effectors can alter the physiological state of the host plant to facilitate pathogen colonization or inhibit the activation of host plant defenses (Collmer 1998). Host plants have evolved a second tier of the defense system known as effector triggered immunity (ETI). ETI relies on intracellular plant receptors such as nucleotide-binding site leucine-rich repeat (NBS-LRR) proteins, to recognize and interact with the pathogen secreted virulence effectors. The interaction between the plant and pathogen proteins can be direct or occur as the result of a modification of a host cellular target that is able to activate the NBS-LRR protein (indirect). The perception of effector proteins by plant resistance proteins typically evokes a strong selection process acting on the effectors while plants acquire new R-genes to recognize the new pathogen effectors, leading to an arms race between plants and their pathogens (Dangl *et al*., 2013).

Once the recognition is initiated, a series of cellular events occurs, leading to the activation of plant defenses. These cellular events include a rapid influx of Ca^{2+} into the plant cell and K^{+} and H^{+} out of the plant cell, a burst of ROS, activation of mitogen-activated protein kinases (MAPKs), expression of pathogenesis related (PR) proteins, induction of salicylic acid (SA), ethylene (ET), and Nitric oxide (NO), suppression of jasmonic acid (JA), deposition of callosic cell wall appositions at infection sites, and programmed cell death (PCD). Many PR

proteins have been characterized in rice, soybean, wheat, maize, barley, sorghum, tomato, and many other plant species. They are normally expressed in pathological or related conditions with biotic and/or abiotic stress. Currently 17 families of PR proteins have been categorized, including PR-1 through PR-17, which play an important role in host defense mainly in incompatible interactions that impede the pathogen progress (van Loon *et al*., 1999). For examples, PR-1 family members have antifungal activity, PR-3, -4, -8, and -11 function as chitinases, while other PR proteins are proteinase inhibitors, peroxidases, defensins, or oxalate like proteins (Sels *et al*., 2008). Therefore, analysis of the expression of well characterized PR genes is a method that is readily utilized as a marker for determining activation of plant defense (De Coninck *et al*., 2015).

Causes of Asian Soybean Rust

Rust fungi are obligate plant parasites that belong to the order Pucciniales in the subphylum Pucciniomycotina, which is considered the earliest diverging lineage in Basidiomycota (Kirk *et al*., 2008). There are about 7,800 species of rust fungi that are categorized into 166 genera and 14 families (Kirk *et al*., 2008). Some of rust pathogens may cause severe damage to crops such as wheat, oat, barley, cotton, and soybean as well as various fruits and vegetables, ornamentals, and timber and pulp trees (Agrios 1997). Therefore, rust fungi are important pathogens in agricultural, horticultural, and industrial forest habitats.

Soybean rust is a devastating foliar disease in tropical and subtropical regions in the Eastern and Western hemisphere, where it was reported to cause yield losses ranging from 10 to 80% (Miles *et al*., 2011). Soybean rust can have a significant financial impact on the world soybean production when severe epidemics develop. Therefore, understanding the biology and phytopathology of ASR is a primary goal for soybean breeders and farmers.

Soybean rust is caused by two closely related pathogens, *P. meibomiae* Author and *P. pachyrhizi* Sydow. They that are distinguished based on a precise morphological comparison, isozyme analysis, as well as a sequence comparison in the internal transcribed spacer (ITS) region of its rDNA (Bond *et al*., 1988; Ono *et al*., 1992; Frederick *et al*., 2002). *P. meibomiae* is native to South America and is described as the 'American' form of soybean rust. This species has been identified on both wild and cultivated plants belonging to the leguminosae family but is less virulent than *P. pachyrhizi* with no reports of severe losses in soybean. In addition, *P. meibomiae* has not been identified in the USA.

P. pachyrhizi is native to Asia, thus designated the causal agent of Asian soybean rust (ASR), and is referred to as soybean rust. Unlike most rust pathogens, *P. pachyrhizi* has a wide host range, naturally infecting more than 90 plant species in 42 genera including 31 legume species in 17 different genera (Bromfield 1984), with *Glycine max* (cultivated soybean), *G. soja* (wild soybean progenitor), *Pachyrhizus erosus* (Jicama or Yam bean), *Pueraria lobata* (kudzu), and *Vigna unguiculata* (Cowpea, Black-eyed pea, Southern pea) being the primary hosts (Vakili 1979; Rytter *et al*., 1984; Ono *et al*., 1992). Severe soybean rust infections were also observed on Kudzu growing in close proximity to soybean fields in Paraguay and in the United States (Morel 2001). The broad host range undoubtedly contributes to the pathogens aggressiveness and ability to survive and overwinter.

Disease History

For almost a century, soybean rust has caused significant yield loss in virtually all soybean producing tropical countries in the Eastern Hemisphere (Ogle *et al*., 1979; Sinclair 1989). *P. pachyrhizi* was considered one of the most important exotic pathogens until its

introduction into the USA (McGregor 1973; Public Health Security and Bioterrorism Preparedness and Response Act of 2002, Select Agents 7 CRF part 331). *P. pachyrhizi* was first reported on *Pachyrhizus erosus* (yam bean, jicama) in Japan in 1902 (Hennings 1903). More recently, soybean rust has been reported in China (Kitani *et al*., 1960; Yeh *et al*., 1975), Africa (Akinsanmi and Ladipo 2001; Pretorius *et al*., 2001), Argentina (Yorinori *et al*., 2002), Australia (Bromfield 1984), Bolivia (Yorinori *et al*., 2005), Brazil (Rossi 2003), Hawaii (Killgore and Heu 1994) and Paraguay (Morel 2001) with associated yield losses ranging from 60 to 100% in some areas costing growers more than 1 billion dollars. The detection of rust in the southern region of South America in 2000 increased the concern of this invasive pathogen entering the USA due to the ability of rust spores to travel over long distances. For this to occur, the density of urediniospores must first increase on living plant tissue for survival and reproduction. Rust spores can then be disseminated in wind currents and storms across and even between continents. Caldwell and Laing (2002) indicated that *P. pachyrhizi* urediniospores identified in Africa were believed to be wind-borne from Asia.

November 6, 2004 marked the first observation of *P. pachyrhizi* rust spores in the continental USA (Schneider *et al*., 2005). Rust spores were initially detected in a soybean production field near Baton Rouge, LA. However, with the unusually wet and moderately cool evening weather conditions experienced in the southeasten USA that is favorable for rust development, spores had spread to other states in the Gulf region within 2 years. In 2007, the United States Soybean Rust Commentary (http://www.sbrusa.net/) reported detection of soybean rust spores in soybean fields in Alabama (five counties), Florida (ten counties), Georgia (five counties), Louisiana (five parishes), and Texas (six counties). Although soybean rust is not seed-borne, spores could travel as contaminants in seed lots from infested areas. These findings

demonstrate that *P. pachyrhizi* is not only spreading to different geographic regions but has the potential to significantly impact yield losses effecting both soybean production and production costs in the USA.

Disease Symptoms and Losses

The disease symptoms of ASR disease can be identified by grayish brown, tan to dark brown or reddish brown lesions with one to many erumpent, globose uredinia on leaves, petioles and stems (Hartman *et al*., 1999). The symptoms begin as small dots of 1 to 2 mm in diameter with greenish gray coloration. The small dots then develop into water-soaked yellowish lesions that quickly expand to 2-3 mm in size and become chlorotic. As the pathogen continues to develop inside the leaf tissue, urediniospores are produced by the volcano-shaped uredinia which are found on the underside of the soybean leaf and develop a rust color (Bromfield 1980).

The repeated infection and colonization of soybean rust on the soybean plant can result in early maturation and defoliation, which will eventually lead to fewer pods, fewer seeds per pod, and reduced seed weight (Hartman *et al*., 1991; Melching *et al*., 1989). In severe cases when the disease reaches the stage of the soybean pod formation, it can cause abortion and drop of the pods, resulting a total loss of income (Soare 2008). Yield losses caused by soybean rust in commercial soybean fields have reported up to 90% yield losses in Asia (Sinclair and Hartman 1999), up to 80% in Africa (Levy 2005), and 30%-80% in North and South America (Yorinori *et al*., 2005; Rupe and Sconyers 2008). In the first year of the arrival of ASR in Brazil, the total economic loss due to soybean rust was estimated at $125.5 million, and resulted in a $10 billion dollar loss through 2007 (Yorinori 2008; Soares 2008). According to a report from the USDA

Economic Research Service (ERS), from $640 million to $1.3 billion in net economic losses were expected during the first year of the appearance of soybean rust in the USA.

Epidemiology and Disease Cycle

Soybean rust epidemics depend upon several factors including host availability, weather conditions and an alternative host with foliar vegetation between cropping cycles. Asian soybean rust cannot survive in most of the soybean production regions in the USA because all of the known hosts of this disease die out during the winter each year and there is no known alternate host for it to overwinter on. Although it only survives in warmer, more frost-free southern and coastal areas of the USA with tropical or subtropical climate, it can infect kudzu (*Pueraria lobata*). ASR spores can potentially survive the winter on kudzu and re-enter the soybean fields during the next spring, causing more damage in production fields. The variations of the frost line in the southern and coastal U.S. areas and unpredictable air from hurricanes or storms make it difficult to detect the dissemination of the disease.

Unlike other rust fungi with five different stages in their life cycles (spermogonium, aecium, uredinium, telium, and basidium), there is a lack of information on sexual reproduction for *P. pachyrhizi* such as spermatia and aeciospores. Although teliospores have been found in nature in Asia, the germination of teliospores and the formation of basidiospores were only induced under laboratory conditions (Bromfield 1984; Saksirirat and Hoppe 1991). The main spore type for *P. pachyrhizi* is asexual urediniospores which are generated on short stalks within a uredium 5 to 8 days after inoculation on colonized leaves (Goellner *et al*., 2010). The germination of urediniospores requires moderate temperatures (15-30°C) and free moisture that can be obtained by mist, fog, or dew. Once germination occurs, the urediniospore produce a

single germ tube that develops across the surface of the leaf until it reaches an appropriate surface where an appressorium forms. Rather than through natural openings or through wounds in the leaf tissue, a penetration hyphae forms from the appressorial cone and directly penetrates through the cuticle and epidermal cells of the leaf between 6 to12 hours after the spores germinate. A septum is then formed to produce the primary hyphae in the intercellular space (Koch *et al*., 1983). The primary hyphae continues to develop to form secondary hyphae and eventually haustoria mother cells form in close contact with mesophyll cells, which leads to the formation of a haustorium between spongy mesophyll cells (Tremblay *et al*., 2010). Besides serving as specialized organ to obtain nutrients from the host, the haustorium is an infection structure associated with multiple functions such as suppression of host defense responses and redirection or reprogramming of the host metabolic flow (Voegele *et al*., 2004). After this stage, more haustoria form as additional hyphae emerge and spread through the apoplast. Four days after a spore lands on the leaf surface, slight necrosis of epidermal and mesophyll cells occurs, that appear at the upper surface of the leaves as yellow mosaic discolorations. As uredinia develop about 5 to 8 days after infection, the first urediniospores are produced from the uredinia 9 to 10 days after infection. The development of uredinia can extend up to 4 weeks, while the spore production can be observed for up to 3 weeks. Urediniospores are finally released by rupture of the epidermis, and uredinia can actively disseminate urediniospores for 4 weeks (Koch *et al*., 1983). The susceptible cultivars are distinguished from resistant cultivars based on yielding high rates of sporulation with tan lesions on the upper surface of the leaf, while resistant cultivars develop dark, reddish-brown lesions with few or no spores (Tremblay *et al*., 2010).

Disease Control

Disease management strategies for soybean rust include application of fungicides with assistance of disease monitoring system, deployment of disease resistant cultivars, and crop cultural practice. Currently soybean rust has been effectively controlled by timely application of fungicide due to the lack of commercial soybean lines that are resistant to the current rust population.

To monitor distribution and severity of soybean rust and provide a "warning" network for tracking the spread of the disease in North America, the USDA established a coordinated framework for soybean rust surveillance, reporting, prediction, management, and outreach (Isard *et al*., 2006; USDA 2005). This framework is called Soybean Rust-Pest Information Platform for Extension and Education (SBR-PIPE). This system obtains data generated from soybean sentinel plots established in multiple locations within cooperating states and provinces (Giesler *et al*., 2007; Hershman 2009; Hershman *et al*., 2011). With the assistance of this monitoring and predication system, state-specific commentary on disease risk and management can be obtained to improve regional suggestions for timely fungicide applications. Strobilurin and triazole are two classes of chemistries that growers currently use to manage soybean rust as well as other diseases of soybean (Kemerait 2014). However, controlling soybean rust with fungicides requires multiple applications at the right stage of plant development as well as fungicide penetration of the canopy of the plant, making this strategy less cost effective. Moreover, health and environmental issues caused by fungicide residue left on the crop raises consumer concerns. Therefore, the deployment of resistant soybean cultivars is considered as the most effective and economic method for controlling soybean rust.

Three levels of host resistance against ASR are defined based on three natural responses of soybean to ASR infection: tan sporulating lesions (TAN), red-brown (RB) lesions, or no visible lesions (IMMUNE). Susceptible genotypes are characterized by tan-colored lesions due to the production of numerous urediniospores within the uredinia, whereas RB lesions typically show limited fungal growth with limited sporulation as a result of a HR. The absence of lesions indicates that the soybean plant is immune to ASR. At present, none of the soybean cultivars utilized in commercial production are resistant to all *P. pachyhrizi* isolates (Hartman *et al*., 2005). Six loci have been identified in soybean with dominant disease resistance genes that confer resistance against ASR: (1) *Rpp1* (Resistance to *P. pachyhrizi*), (2) *Rpp2*, (3) *Rpp3* (4) *Rpp4*, (5) *Rpp5* and (6) *Rpp6* (McLean and Byth 1980; Bromfield and Hartwig 1980; Hartwig and Bromfield 1983; Hartwig 1986; McDonald and Linde 2002; Garcia *et al*., 2008; Li *et al*., 2012).

The majority of crop plants rely on single-gene resistance to various diseases. However, this type of resistance has not been durable. The utility of three of the *Rpp* sources became ineffective soon after deployment according to previous reports. The original cultivar carrying *Rpp2* was highly resistant in field trials only to lose its effectiveness within three years (Bromfield 1984). The same is true for *Rpp1* and *Rpp3* (Singh 1975; Kochman 1977). Recently, both *Rpp1* and *Rpp3* were reported ineffective in conferring resistance in Brazil, leaving only *Rpp2*, *Rpp4*, and *Rpp5* resistance in some regions of that country (Alves Pereira Morales *et al*., 2012). Therefore, pyramiding *Rpp* genes and characterizing new sources of resistance are important to create cultivars with durable resistance.

The six *Rpp* loci have been genetically mapped to different chromosomes in soybean. *Rpp1*, *Rpp4*, and *Rpp6* are located in three different regions on chromosome 18 (Silva *et al*.,

2008; Chakraborty *et al*., 2009; Li *et al*., 2012). *Rpp2*, *Rpp3*, and *Rpp5* are located on chromosome 16, 6, and 3, respectively (Hyten *et al*., 2007; Garcia *et al*., 2008; Silva *et al*., 2008). Soybean cultivar Hyuuga carries the *Rpp?* resistance allele that was mapped to the *Rpp3* locus on chromosome 6. Based on the additional screening of Hyuuga-derived recombinant inbred lines, it was recently reported that Hyuuga also carries another R-gene located on chromosome 3 (Monteros *et al*., 2007; Kendrick *et al*., 2011). The above described R-genes have been identified genetically but to date have not been cloned. This indicates that *Rpp* genes may be disbursed throughout the soybean genome similar to other well characterized disease resistance genes in maize, rice and wheat (Li *et al*., 2012).

More than 16,000 soybean accessions have been screened, resulting in the identification of at least twenty-two additional accessions conferring dominate resistance that is not governed by *Rpp2* or *Rpp4*, indicating that the twenty-two accessions are likely new sources of resistance to ASR (Hartman *et al*., 2005; Goellner *et al*., 2010). Potential resistance to ASR was also evaluated among other legume plants, such as common bean (*Phaseolus vulgaris*) and kudzu (*Pueraria* spp.). For example, a set of forty-four *P. vulgaris* genotypes were tested for resistance. Twenty-four of the *P. vulgaris* genotypes were resistant and showed the classic reddish brown lesion phenotype. In regards to kudzu, a total of 139 kudzu sites were evaluated to determine the natural resistance against ASR. At the evaluated sites, 18% of the kudzu plants were found to be free of soybean rust infection, while 23% had reduced sporulation. In addition, ten kudzu accessions from north-central Florida were inoculated with ASR. Six of the accessions were susceptible, three were immune, and one was resistant (Jordan *et al*., 2010; Souza *et al*., 2014). This indicates that kudzu carries resistance genes that are potentially effective against some rust field isolates.

Soybean perennial wild relatives were also screened for resistance to ASR (Burdon and Marshall 1981; Hartman *et al*., 1992). For example, screening the resistance of *G. tomentella* and *G. tabacina* wild relatives indicated that 33% and 32% of the total number of accessions, respectively, were highly resistant. In addition, the inheritance of resistance associated with specific wild species was also studied. Seven *G. canescen*s accessions were reported to carry single dominant resistant genes at more than four loci, while *G. argyrea* has one dominate resistance gene (Burdon 1988; Jarosz and Burdon 1990). This indicates that some soybean wild relatives may carry disease resistance genes that can be potentially effective against the current *P. pachyrhizi* population. However, only *G. tomentella* and *G. tabacina* soybean wild relatives have been successfully crossed with soybean and required a significant amount of work to obtain viable progeny (Burdon and Marshall 1981; Patzoldt *et al*., 2007). Therefore, although many soybean wild relatives have demonstrated resistance against ASR, introgression of resistance genes into cultivated soybean from soybean wild relatives is not realistic. Resistant wild soybean wild relatives can however, be used to analyze the genetics of resistance to ASR (Kuang *et al*., 2005; Wicker *et al*., 2007; Lee *et al*., 2009).

The mechanism and basis of race specificity is often mediated by a process where an R-gene product detects the presence of a specific pathogen avirulence (Avr) protein. This detection activates many plant defense responses including the HR (Flor 1955; Keen 1990; Staskawicz *et al*., 1995). HR involves cell death at the site of infection and the accumulation of antimicrobial compounds (Scofield *et al*., 1996; Tang *et al*., 1996). Therefore, host resistance and pathogen avirulence (Avr) is specific and complementary, referred to as the gene-for-gene interaction (Flor 1955). This type of resistance is often broken when the pathogen population loses or mutates the Avr proteins so that they are no longer recognized by the corresponding plant R-gene protein and

results in the loss of plant resistance (McDonald and Linde 2002). This is especially true for soybean given its cultivation on large acreages which in-turn increases the probability that large pathogen populations will mutate to new forms. It is therefore easily understood that instability in pathogen effector proteins has been the biggest problem in controlling plant disease, requiring breeders to go through the lengthy process of repeatedly developing new cultivars with different R-genes often introgressed from related species. Other factors thought to influence the durability of race-specific resistance include genetic plasticity of pathogen populations, effects of gene combinations or pyramids, the physiological cost of losing the corresponding *Avr* gene for the pathogen, and the level of selection pressure conferred by the host resistance gene (Schafer and Roelfs 1985; Leonard 1997; Bai *et al*., 2000; Leach *et al*., 2001; McDonald and Linde 2002).

The risks associated with the use of single-gene resistance are well known. However, this method is still utilized due to the lack of other viable options. Consequently, the efforts of most breeding programs have been directed toward seeking more durable resistance strategies to protect against rust diseases. Other strategies proposed to prolong the usefulness of R-genes are gene pyramiding and the use of slow-rusting or partial resistance. Combining different R-genes into a single cultivar frequently results in resistance to multiple pathotypes carrying the corresponding *Avr* gene which can potentially decrease selection pressure against the pathogens and slow the rapid evolutionary changes of the pathogen isolates. Examples of this multi-gene directed type of nonspecific resistance to rusts can be found in some cereal crops, most notably maize (Hu *et al*., 1994; Richter *et al*., 1995; Smith *et al*., 2004). Facilitated by marker assisted selection and molecular engineering, this approach is being investigated in soybean and is thought to have the potential for management of soybean rust.

The ability of the pathogen to overcome host resistance is of particular concern, as genetic resistance and chemical protectants are the primary management tools for ASR. Although improving resistance to ASR is a primary breeding objective of all private and public breeding efforts, little is known about the genetics of resistance to ASR. In addition, very little is known about the specificity or genetic diversity of prevalent *P. pachyrhizi* isolates in Georgia and the surrounding states. Therefore, greater knowledge of the genetics of resistance to ASR and the variability of ASR populations is essential in order to dissect the relationship between the host and this pathogen and effectively breed for resistance.

Pathogen Variability and Population Genetics of *P. pachyrhizi*

Pathogen variability has been described by terms such as races, pathotypes, and virulence types. A physiological race is defined as a population where all individuals have the same combination of virulence genes and can be determined based on the response to a set of differentials with resistance genes (Parlevliet 1985). Pathotypes are normally used to demonstrate pathogen genotypes where pathogens express the same pattern of host-specific virulence towards individual clones or species within a specified set, while virulence is just the genetic ability to overcome host resistance and cause a compatible interaction (Shaner *et al*., 1992; Ramstedt *et al*., 2002).

By inoculating different isolates of *P. pachyrhizi* onto various sets of soybean differentials, different races were identified based on the reaction types, although a standard set of differentials and the methods of evaluating reaction types in soybean have yet to be established (Lin 1966; McLean and Byth 1976; Bromfield 1984; Burdon and Speer 1984; Kato and Yorinori 2008; Twizeyimana *et al*., 2009; Yeh 1983; Pham *et al*., 2009; Twizeyimana and

Hartman 2012). The pathogenicity of three ASR populations from Japan and Brazil was compared by using 13 selected cultivars. It was found that only two resistance genes, *Rpp4* from PI459025 and *Rpp5* from Shiranui, commonly conferred resistance to all three rust populations. However, the phenotypes of cultivars carrying *Rpp1*, *Rpp2*, *Rpp3* and unknown *Rpp* genes were different when inoculated with the three rust populations. This indicated that the South American population varied and pathogenicity was affected geographically and temporally. Additionally, the populations obtained from Japan seemed to be different from the South American populations.

Pathogen variability can also be detected by analyzing the DNA sequences of Internal Transcribed Space (ITS) and the ADP Ribosylation Factor (ARF) gene in different isolates. Zhang *et al*. (2012) identified twenty-nine polymorphisms in the ITS and ARF gene sequences of fifty-nine isolates and demonstrated that there were at least five different *P. pachyrhizi* isolates residing within the USA, suggesting a unique origin or rapid diversification of *P. pachyrhizi* isolates. As the use of single dominate resistance genes and fungicides increase, this may put more selection pressure on the ASR population to evolve and generate new isolates, resulting in the breakdown of resistance. Therefore, knowledge of the changing population structure of ASR is important for developing disease management strategies.

Therefore, inoculation of a core set of soybean differentials with ASR field isolates and sequence analysis of regions that differentiate fungal isolates can be used to assess the diversity and genetic relatedness of ASR field isolates in Georgia. This will also test our ability to predict the performance of new field isolates.

Molecular Aspects of Soybean-rust Interaction

Currently little is known about the molecular interaction of soybean and *P. pachyrhizi* and the defense pathways triggered by pathogen recognition. Therefore, understanding the molecular mechanisms involved in plant defense responses is critical to better develop strategies to control ASR. With the development of large scale sequencing and gene expression techniques, soybean's response to different pathogens has been analyzed using large scale transcriptional profiling methods including microarray and RNA sequencing (RNA-seq) analyses. For example, by using RNA-seq, a comparative transcriptomic analysis of differentially expressed genes (DEG) was performed on a susceptible soybean cultivar (Williams 82) and ten resistant soybean near isogenic lines (NILs). The eleven soybean lines were inoculated with *Phytophthora sojae*, the causal agent of Phytophthora root and stem rot (PRR) of soybean (Lin *et al*., 2014). A total of 4,330 DEGs were identified in the susceptible Williams cultivar, while from 2,014 to 5,499 DEGs were identified in the individual NILs. Many of the DEGs were associated with *Rps*-mediated resistance responses that are involved in multiple signaling pathways in the soybean NILs, including ethylene (ET), jasmonic acid (JA), (reactive oxygen species) ROS, and (MAP kinase) MAPK signaling. This indicated that genes that are upregulated in the resistant lines may play an important role in differentiating molecular defense responses.

Choi *et al*. (2008) utilized a similar approach to analyze DEGs in soybean in response to ASR inoculation. This group performed a large-scale transcript profiling analysis on soybean plants that were homozygous for the *Rpp1* genes (Resistance to *P. pachyrhizi*). Genes that were up-regulated greater than two-fold in the resistant plants when compared to the susceptible plants were studied. Lipoxygenease and peroxidase genes demonstrated the greatest differential expression indicating that these genes have important functions in *Rpp1*-mediated resistance.

Additionally, a bi-phasic response of soybean to the pathogen was found in *Rpp1*- and *Rpp2*-mediated resistance (van de Mortel *et al*., 2007; Panthee *et al*., 2007; Choi *et al*., 2008). The first wave of differential gene expression was detected between 6 and 36 hours post-inoculation (hpi), in compatible and incompatible interactions. This period of time corresponds to the stage when the germination and penetration of ASR spore occurs. The highest level of differential expression was detected 12 hpi in the resistant plants when comparing the inoculated plants to the mock inoculated plants. Gene expression returned to basal levels by 24 hpi. The second wave of defense-related gene expression was observed 72 hpi and corresponded to haustoria mother cell and haustoria formation. Interestingly, the increase in gene expression of defense-related genes was more rapid in the resistant plants indicating that the timing of gene expression is also important in plant defenses. The changes in the pattern of gene expression during the early stages of infection might represent a non-specific recognition of the pathogen and activation of basal defense response in both resistant and susceptible plants, while the second wave of differential gene expression suggests recognition of pathogen effector proteins by plant resistance genes, triggering the rapid defense response in the resistant lines.

Many genes are expressed in the plant and the pathogen during infection. Simultaneous analysis of genes differentially expressed in the host plant and pathogen during infection has provided more insight into the putative function of these genes during the infection process. Using laser capture microdissection, Tremblay *et al*. (2009, 2010) isolated palisade and mesophyll cells from susceptible soybean showing signs of infection. A small set of ASR ESTs (expressed sequence tags) were generated from the infected susceptible soybean showing signs of infection and was used to performed transcriptome profiling. About 80% of the identified ESTs corresponded to genes that shared no homology to previously described *Phakopsora* genes.

This group also performed RNA-seq on susceptible soybean leaves inoculated with *P. pachyrhizi*, to analyze pathogen transcript abundance at 15 seconds, 7 hours, 48 hours, and 10 days post-inoculation. The results indicated that energy, nucleotide metabolism, and protein synthesis are major priorities for the fungus during infection and development. These results are similar with previous studies where differentially expressed genes were assigned putative roles in primary metabolism, gene and protein expression, cell structure and growth, cell division, cell signaling, and cell communication (Posada-Buitrago and Frederick 2005; Tremblay *et al*., 2013).

There is increasing evidence that fungal and oomycete pathogens secrete numerous effector proteins into their hosts (e.g. Kamoun 2007; Hogenhout *et al*., 2009; Tyler 2009; Birch *et al*., 2009; Hok *et al*., 2010; Ravensdale *et al*., 2011). The effector proteins can then manipulate host cell structure and function, thereby facilitating infection (virulence factors or elicitors) (Kamoun 2006). To date, six effector proteins have been identified and characterized in three different rust species. Four effector proteins were isolated from flax rust, *Melampsora lini* (*AvrM*, *AvrL567*, *AvrP123*, and *AvrP4*), one effector protein from bean rust, *Uromyces fabae*, RTP1 (Rust Transferred Protein) and one effector protein from the wheat stem rust fungus *Puccinia graminis* f. sp.*tritici* (PGTAUSPE-10-1) (Kemen *et al*., 2005; Ellis *et al*., 2007; Upadhyaya *et al*., 2014). In terms of ASR, over 20% of all ESTs from germinating urediniospores were found to encode a small secreted, cysteine-rich protein similar to proteins expressed by *Blumeria graminis* f. sp. *Hordei* and *Magnaporthe grisea* during spore germination and appressorium formation (Posada-Buitrago and Frederick 2005), indicating that these proteins may play an important role in the early stages of infection.

The identification of rust and bacterial core effectors and the corresponding resistance proteins from the host plant has been utilized to identify potentially durable resistance genes (R-

genes). The resistance conferred by the R-gene(s) is likely to be durable when the corresponding core effector has an important virulence function in the pathogen and is therefore conserved in the pathogen population. This was observed for *Xanthomonas axonopodis* pv. *manihotis* (*Xam*), the causal agent of cassava bacterial blight (Bart *et al*., 2012). One core effector was identified in ~65 *Xam* strains collected over a 70-year time frame from 12 countries. The core effector was shown to serve as a target activated by R-genes in wild species of *Manihot* and potentially other related plants in the *Euphoribiaceae*. In addition, several oomycete and fungi effectors, including rust pathogens and an ectomycorrhizal fungus, were demonstrated to contain conserved and functional RXLR variants which could mediate their transduction into plant cells (Rafiqi *et al*., 2010; Kale *et al*., 2010; Gu *et al*., 2011; Plett *et al*., 2011; Tian *et al*., 2011).

Classification and Structure of R-genes

To improve the disease resistance of economically important plants, many studies have focused on isolating and cloning R-genes. As mentioned previously, over 70 R-genes have been cloned from different plant species (Liu *et al*., 2007). The recent cloning of disease R-genes and characterization of their predicted products significantly contributed to understanding the molecular basis of plant disease resistance.

There are five major classes of resistance genes. The largest class is the NBS-LRR (nucleotide binding site-leucine-rich repeat) class of resistance genes (Hulbert *et al*., 2001). Depending on the N-terminus structure, the NBS-LRR class of R-genes can be categorized into two subgroups: (1) TIR NBS-LRR class carrying a motif with homology to the cytoplasmic domains of the *Drosophila* Toll protein and the mammalian interleukin-1 receptor (TIR), and (2) non-TIR NBS-LRR class that has a coiled coil (CC) motif or a leucine zipper (LZ). This R-gene

class is predicted to be located in the cytoplasm of the plant cell. Therefore, interaction with the corresponding effector protein has been shown to occur in the cytoplasm (Wang *et al*., 2012). The NBS domain carries several conserved amino acid motifs (P-loop, kinase-2, kinase-3a and GLPL domain) that are responsible for nucleotide binding, downstream signaling and activation of plant defenses (Traut 1994; Tamelin *et al*., 2002). The LRR region usually contains 20-29 residues in length with a conserved 11-residue segment LxxLxLxxNxL or 12-residue segment LxxLxLxxCxxL (Matsushima 2012). It has been shown that the LRR domain interacts with the pathogen effector proteins and is therefore responsible for determining resistance specificity in the host (Kobe and Deisenhofer 1995; Dangl and Jones 2001).

The conservation of several amino acid motifs in the NBS domain has been used to design degenerate primers to isolate R-gene homologs (RGHs) from different plant species with a polymerase chain reaction (PCR)-based approach (Meyers *et al*., 2003; Ameline-Torregrosa *et al*., 2008). Examples of the NBS-LRR class of R-genes are the *RPS* genes (*RPS2*, *RPM2*, and *RPS5*) in *Arabidopsis* that confer resistance to *P*. *syringae* carrying bacterial effectors *AvrRpt2*, *AvrRpm1*/AvrB, and *AvrPphB,* respectively (Dangl and Jones 2001).

The second major class of R-genes has an extracellular LRR a transmembrane domain (TM) and a small cytoplasmic region. The tomato *Cf* resistance genes are the best characterized genes from this class and confer resistance to the biotrophic leaf-mold pathogen *Cladisporum fulvum* carrying *Avr2*, *Avr4*, and *Avr9* effector molecules (Jones *et al*., 1994; Joosten *et al*., 1994; de Wit 1995). Due to the extracellular LRR, interaction with the effector molecule is predicted to occur outside of the plant cell. The cytoplasmic receptor kinase (RLK) is the third major class of resistance genes. The rice *Xa21* gene is the best characterized RLK and confers resistance to effector molecules from *Xanthomonas oryae* pv. *oryzae*. This R-gene has an extracellular LRR, a

transmembrane domain (TM), and a cytoplasmic kinase domain (Ronald *et al*., 1992; White and Yang 2009). Interaction with the effector molecule occurs outside of the plant cell. The recognition signal is then transmitted to the internal kinase domain through the TM domain, leading to activation of downstream signaling cascades and plant defenses. The forth major class of resistance genes is the kinase class and is best described by the *Pto* kinase gene in tomato (Pan *et al*., 2000). *Pto* interacts with another NBS-LRR internal protein (*Prf*) to form a protein complex. Formation of the protein complex results in the activation of plant defenses. The *Rpw8* gene in *Arabidopsis* describes the fifth major class of resistance genes. It has a TM domain and an internal CC domain. *Rpw8* is associated with broad-spectrum resistance to powdery mildew but elicits resistance in a similar manner to NBS-LRR proteins utilizing a conserved signaling pathway dependent on salicylic acid (Xiao *et al.*, 2005)

In addition, there are examples of R proteins with novel structures that do not belong to any of the five major classes. For example, RRS1-R that recognizes *Ralstonia solanacearum* encodes a TIR-NBS-LRR protein carrying a carboxy-terminal nuclear localization signal and WRKY transcriptional activation domain (Deslandes *et al*., 2003). The *Xa27* gene cloned from rice shares no homology with any known R-genes and is only expressed in the vicinity of the tissue infected by *Xanthomonas oryae* pv. *oryzae* expressing the cognate effector protein *AvrXa27* (Gu *et al*., 2005). Two unique rice R-genes are *Xa5* that encoding a TFII transcription factor and *Xa13* that has homologous to the nodulin *MtN3* gene (Chu *et al*., 2006).

R-gene Protein, Effector Protein Recognition Models

Plant R-gene proteins can interact with effector proteins from the pathogen directly or indirectly. However, the majority of the identified R-gene proteins interact with the

corresponding effector proteins indirectly and rely on additional plant accessory proteins that is a pathogen virulence target or a structural mimic of a target (Dodds and Rathjen 2010). The gene-for-gene model describes a direct interaction and was first proposed by H. H. Flor (Flor 1971). There are only three well characterized examples of R-gene direct interactions. The flax *L* and *M* TIR-NBS-LRR genes interact directly with *Melampsora lini* fungal effectors *AvrL567* and *AvrM*, respectively. Additionally, the rice *Pi-ta* CC-NBS-LRR gene directly interacts with the rice blast (*Magnaporthe grisea*) effector *AvrPita*. This recognition and direct interaction leads to defense activation.

Additional models were proposed after several R-genes and their corresponding *Avr* genes were cloned but researchers were unable to demonstrate a direct interaction using several methods including the yeast two hybrid system. This indicated that the majority of the R-gene proteins and effector proteins do not follow the gene-for-gene interaction model proposed by Flor many years ago. After the cloning and characterization of several R-genes and effectors, it has been demonstrated that the gene-for-gene model is in fact the exception and not the rule (Cui *et al*., 2015). Three additional models have been proposed to describe an indirect interaction of R-gene proteins and effector proteins.

The first is described as the "guard model". This model was first proposed by van de Biezen and Jones in 1998 demonstrating that NBS-LRR proteins guard an accessory internal plant protein (guardee) that is targeted and conformationally modified by pathogen effectors (van de Biezen and Jones 1998). The best characterized example of this model is in *A. thaliana*. The *Arabidopsis* NBS-LRR protein *RPM1* serves as the guard for RIN4. RIN4 is an internal plant protein (guardee) that is targeted by two *Pseudomonas syringae* effectors, *AvrB* or *AvrRPM1*. The effectors mediate the phosphorylation of RIN4, leading to a conformational change in the

effector protein-guardee protein complex, which is detected by *A. thaliana* NBS-LRR protein, *RPM1* (Axtell and Staskawicz 2003; Mackey *et al*., 2003). There is no direct interaction between the R-gene protein and the effector protein. Additionally, one individual R-protein can recognize more than one effector with recognition of the modification of a guardee protein. This can explain the phenomenon as to how plant species can defend themselves against multiple pathogens when there are a limited number of functional R-genes in the plant genome.

The "decoy model" is another proposed indirect interaction model with a slight modification of the guard model. The decoy model states that the accessory protein is able to perceive the effector to broaden the recognition capability for its corresponding R-protein. For example, the tomato NBS-LRR protein *Prf* relies on the *Pto* Serine Threonine protein kinase accessory protein that is the target of *P. syringae AvrPto* and *AvrPtoB* (Mucyn *et al*., 2006; Shan *et al*., 2008). Collier and Moffett further modified both the guard and decoy model and proposed the "bait-and-switch model". This model proposes that the pathogen effector protein is brought into the NBS-LRR system via the bait protein/co-factor (guardee protein/effector target), leading to the subsequent recognition between the effector and NBS-LRR protein through a direct manner (Collier and Moffett 2009). However, due to the inadequate evidences or examples to support decoy and the bait-and-switch models, it is important to note these two proposed models still need to be verified and the recognition events are not well understood (Dodds and Rathjen 2010). This indicates that the recognition and interaction of R-gene and effector proteins is not as straight forward as initially believed and can vary depending on several factors including the plants species, the pathogen, and the class of R-genes and effectors.

The Structure and Evolution of R-genes

With the complete genome sequence of many plants available, extensive analyses have been done to identify R-genes in a genome-wide manner in plants such as *Arabidopsis*, rice, soybean, potato, sunflower, corn, lettuce, papaya, *Populus trichocarpa*, *Brassica rapa*, and sorghum (Meyers *et al*., 2003; Yang *et al*., 2006; Xiao *et al*., 2006; Kohler *et al*., 2008; Radwan *et al*., 2008; Mun *et al*., 2009; Porter *et al*., 2009; McHale *et al*., 2009; Li *et al*., 2010; Guo *et al*., 2011; Bakker *et al*., 2011). Soybean is predicted to have a genome size of 1,115 megabases (Mb). Schmutz *et al* (2010) has reported a soybean whole-genome shotgun sequence of *G. max* Williams 82 cultivar, comprised of 950 Mb assembled and anchored sequences. They also demonstrated that this plant has undergone at least two rounds of large-scale duplication approximately 13 and 59 million years ago (Mya), making the genome highly duplicated with about 75% of the genes present in multiple copies. The duplication events, including tandem and segmental duplication may have a critical impact on the expansion and/or contraction of NBS genes. So far, only a few soybean R-genes have been characterized, and all of them belong to the NBS-LRR class (He *et al*., 2003; Ashfield *et al*., 2004; Hayes *et al*., 2004; Wang *et al*., 2004; Gao *et al*., 2005; Meyer *et al*., 2009). To analyze the structure and evolution of R-genes, a couple of studies focusing on a genome-wide identification of NBS-LRR R-genes have been performed in not only in soybean, but also in other legume plants such as *Medicago truncatula*, pigeon pea, common bean and castor bean (Zhang *et al*., 2011; Kang *et al*., 2012; Shao *et al*., 2014). Therefore, analysis of duplication of NBS-LRR genes in soybean will identify the main forces that drive the diversification and rapid evolution of NBS-LRR genes in soybean.

The clustering of R-genes is considered to result from gene duplication events, including tandem (transferring duplicated segment), segmental (duplication of large DNA fragments and

entire chromosomal regions) and ectopic (one or several genes are copied or translocated to distal and probably random locations in the genome) duplication events that are associated with the maintenance of sequence polymorphisms in the R-gene loci (Michelmore *et al*., 1998; Meyers *et al*., 2005; Liu *et al*., 2007). Due to the "arm-race" between plants and pathogens, R-genes and effector genes co-evolve and can generate new resistance genes with novel specificity by different mechanisms, such as genetic recombination, gene conversion, point mutation, or unequal crossing over. These processes play critical roles in the distribution of R-genes, amplification and reduction of the number of R-gene family members, and generation of new specificity against different pathogen effectors (Leister 2004). Both tandem and segmental duplications are involved in the evolution of *Arabidopsis* NBS-LRR genes, and it is suggested that many duplicated sequences might have been lost, or diverged through accumulated mutations to diversify the R-gene families (Meyers *et al*., 2003). Similarly, in maize, a cluster of R-genes (*Rp1* complex) are arranged in a tandem array on the short arm of chromosome 10. There are a total of 14 dominate genes within the *Rp1* complex, providing specific resistance to *Puccinia sorghi*. The new specificity against this pathogen can be generated through inter/intragenic recombination events among gene families. For example, *Rp1-JF* and *Rp1-J* parental haplotypes generated 9 recombinants with the combined resistance of both parents as a result of inter-recombination event, while *Rp1-D* and *Rp1-I* heterozygote generated four recombinant progeny haplotypes with novel non-parental resistance specificity (Hulbert 1997). Selection imposed on the host plant and the pathogen is responsible for the diversification of NBS-LRR gene sequences (Pan *et al*., 2000). Diversifying selection typically favors non-synonymous substitutions over synonymous exchanges in most R-genes, which can change the specificity of the host plant. Therefore, this type of selection leads to high amino acid variation

at the R-gene loci and rapid evolution of R-genes and pathogen effector genes. Conversely, population studies suggested that balancing selection maintains R-genes in a polymorphic state within a population (Meyers *et al*., 2005). However, selection is balanced due to positive impact of the enhanced fitness of the host in the presence of the pathogen and polymorphism in the R-gene loci is well maintained in the population for an extended period of time. For instance, the *Arabidopsis Rpp13* R-gene and the *Peronospora prasitica* effector *Atr13* were stably maintained in natural populations, indicating that both genes may be under balancing selection (Allen *et al*., 2004). Rose *et al* (2004) also analyzed 24 wild accessions of *Arabidopsis* and demonstrated that the polymorphisms that are attributed to diversifying selection, correlated with functional differences in resistance among *Rpp13* alleles, and the ratio of nonsynonymous and synonymous substitutions (Ka/Ks) demonstrated that the amino acid variation in the LRR regions of *Rpp13* genes were favored by balancing selection. Therefore, analysis of the type of selection pressure that is imposed on the R-genes maintained in the plant and the effectors in the pathogen can give some insight as to how these genes are evolving and potentially identify R-genes that are stable in the plant population and can provide durable resistance.

The classic approach to the study of the genetic basis of plant disease resistance has led to a great deal of crop improvement, especially in species like wheat where rust diseases are the major biotic limitation on food production (McIntosh and Brown 1997; Ellis *et al*. 2007). In the genomics era, if one can find a way to analyze a trait by DNA sequence analysis, then it is likely to be quicker, more definitive, and more comprehensive than any other approach. Hence, one of the goals of this proposed project is to characterize ASR field isolates, where we combine classical and molecular methods to identify host resistance genes and pathogen avirulence loci,

then use the gained knowledge to determine which ASR genotypes are the most predominant and most problematic.

Project Objectives

The original research presented in this dissertation aims to analyze the mechanism of soybean resistance to the current ASR population in Georgia by characterizing both host resistance and pathogen variability. This will be a major advance in assessing disease risk. Detailed data on pathogen diversity and host resistance is lacking and is required to develop sustainable management strategies (e.g., to determine whether new races of the pathogen are being introduced into production areas or if new races appear as variants of endemic genotypes). The long-term goals of this project are to define the genetics of resistance to ASR in soybean, create genetic and genomic resources to study resistance, and determine the genetic basis of race specificity in ASR. To date, achieving these goals has been severely hindered by a lack of genetic and genomic resources for soybean as well as for ASR.

The objectives of this research are:

1) To phenotype ASR isolates on a set of soybean differentials, released soybean cultivars, and wild relatives and to analyze the genetic diversity of ASR field isolates in Georgia.
2) To identify candidate ASR effector genes and host R-genes that are expressed during infection with ASR from the current pathogen population.
3) To analyze the structure and evolution of R-genes in soybean and its wild relatives.

Literature cited

1. Agrios G. N. (1997) Plant pathology. 4th edition. Academic Press, London, UK.
2. Akamatsu H., Yamanaka N., Yamaoka Y., Soares R. M., More L. W., Ivancovich A. J. G., Bogado A. N., Kato M., Yorinori J. T., Suenaga K. (2013) Pathogenic diversity of soybean rust in Argentina, Brazil and Paraguay. Journal of General Plant Pathology 79:28-40.
3. Akinsanmi O. A., Ladipo J. L. (2001) First report of soybean (*Phakopsora pachyrhizi*) rust in Nigeria. Plant Disease 85:97.
4. Allen R. L., Bittner-Eddy P. D., Grenville-Briggs L. J., Meitz J. C., Rehmany A. P., Rose L. E., and Beynon J. L. (2004) Host-parasite coevolutionary conflict between *Arabidopsis* and downy mildew. Science 306:1957-1960.
5. Alves Pereira Morales A. M., Borem A., Graham M. A., Abdelnoor R. V. (2012) Advances on molecular studies of the interaction soybean-Asian rust. Crop Breeding and Applied Biotechnology 12:1-7.
6. Ameline-Torregrosa C., Wang B. B., O'bleness M. S., Deshpande S., Zhu H. *et al*. (2008) Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. Plant Physiology 146:5-21.
7. Ashfield T., Ong L. E., Nobuta K., Schneider C. M., Innes R. W. (2004) Convergent evolution of disease R-gene specificity in two flowering plant families. Plant Cell 16:309-318.
8. Axtell M. J. and Staskawicz B. J. (2003) Initiation of *RPS2*-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. Cell 112:369-377.

9. Bai J., Choi S. H., Ponciano G., Leung H., and Leach J. E. (2000) *Xanthomonas oryzae* pv. *oryzae* avirulence genes contribute differently and specifically to pathogen aggressiveness. Molecular Plant-Microbe Interactions 13: 1322-1329.

10. Bakker E., Borm T., Prins P., van der Vossen E., Uenk G., Arens M., de Boer J., van Eck H., Muskens M., Vossen J., vd Linden G., van Ham R., KleinLankhorst R., Visser R., Smant G., Bakker J., Goverse A. (2011) A genome-wide genetic map of NB-LRR disease resistance loci in potato. Theoretical and Applied Genetics 123:493-508.

11. Bart R., Cohn M., Kassen A., McCallum E. J., Shybut M., Petriello A., Krasileva K., Dahlbeck D., Medina C., Alicai T., Kumar L., Moreira L. M., Rodrigues Neto J., Verdier V., Santana M. A., Kositcharoenkul N., Vanderschuren H., Gruissem W., Bernal A., Staskawicz B. J. (2012) High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proceedings of the National Academy of Science USA 109(28):972-1979.

12. Birch P. R., Armstrong M., Bos J., Boevink P., Gilroy E. M., Taylor R. M., Wawra S., Pritchard L., Conti L., Ewan R., Whisson S. C., van West P., Sadanandom A., and Kamoun S. (2009) Towards understanding the virulence functions of RXLR effectors of the oomycete plant pathogen *Phytophthora infestans*. Journal of Experimental Botany 60:1133-1140.

13. Boller T., Felix G. (2009) A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annual Review of Plant Biology 60:379-406.

14. Bonde M. R, Peterson G. L., Dowler W. M. (1988) A comparison of isozymes of *Phakopsora pachyrhizi* from the Eastern Hemisphere and the New World. Phytopathology 78:1491-1494.

15. Bromfield K. R. (1984) Soybean rust, Monograph (American Phytopathological Society), No. 11 American Phytopathological Society. St. Paul, MN.

16. Burdon J. J., and Marshall, D. R. (1981) Evaluation of Australian native species of *Glycine* for resistance to soybean rust. Plant Disease 65:44-45.

17. Burdon J. J., Speer S. S. (1984) A set of differential *Glycine* hosts for the identification of races of *Phakopsora pachyrhizi* Syd. Euphytica 33:891-896.

18. Burdon J. J., Thrall P. H., and Ericson L. "The current and future dynamics of disease in plant communities." Annual Review Phytopathology 44 (2006): 19-39.

19. Caldwell P. and Laing M. (2002) Soybean Rust - A New Disease on the Move. Southern African Society for Plant Pathology. Stellenbosch South Africa. http://www.saspp.org/archived_articles/FeatureMarch.php

20. Calenge F., van der Linden C. G., van de Weg E., Schouten H. J., van Arkel G. *et al*. (2005) R-gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. Theoretical and Applied Genetics 110: 660-668.

21. Chakraborty N., Curley J., Frederick R. D., Hyten D. L., Nelson R. L., Hartman G. L., Diers B. W. (2009) Mapping and confirmation of a new allele at *Rpp1* from soybean PI 594538A conferring RB lesion type resistance to soybean rust. Crop Science 49:783-790.

22. Chanamalikarjuna V., Sonah H., Prasad M., Rao G. J. N., Chand S., Upreti H. C., Singh N. K. and Sharma T. R. (2010) Identification of major quantitative trait loci qSBR1-1 for sheath blight resistance in rice. Molecular Breeding 25:155-166.

23. Choi J., Alkharouf N., Schneider K., Matthews B. and Frederick R. D. (2008) Expression patterns in soybean resistant to *Phakopsora pachyrhizi* reveal the importance of peroxidases and lipoxygenases. Functional & Integrative Genomics 8: 314-359.

24. Chu Z., Fu B., Yang H., Xu C., Li Z., Sanchez A., Park Y. J., Bennetzen J. L., Zhang Q. and Wang S. (2006) Targeting xa13, a recessive gene for bacterial blight resistance in rice. Theoretical and Applied Genetics 112(3): 455-461.

25. Chung G., Singh R. J. (2008) Broadening the genetic base of soybean: a multi-disciplinary approach. Critical Reviews in Plant Sciences 27:295-341.

26. Collier S. M. and Moffett P. (2009) NB-LRRs work a 'bait and switch' on pathogens. Trends in Plant Science 14:521-529.

27. Collmer A. (1998) Determinants of pathogenicity and avirulence in plant pathogenic bacteria. Current Opinion in Plant Biology 1:329-35.

28. Cui H., Tsuda K., Parker J. E. (2015) Effector-triggered immunity: from pathogen perception to robust defense. Annual Review of Plant Biology 66:487-511.

29. Dangl J. L. and Jones J. D. G. (2001) Plant pathogens and integrated defence responses to infection. Nature 411: 826-833.

30. Dangl J. L. *et al*. (2013) Pivoting the plant immune system from dissection to deployment. Science 341:746-751.

31. De Coninck B., Timmermans P., Vos C., Cammue B. P., & Kazan K. (2015) What lies beneath: belowground defense strategies in plants. Trends in plant science 20(2): 91-101.

32. de Wit P. J. (1995) *Cf9* and *Avr9*, two major players in the gene-for-gene game. Trends in Microbiology 3:251-252.

33. Deslandes L., Olivier J., Peeters N., Feng D. X., Khounlotham M., Boucher C., Somssich I., Genin S., and Marco Y. (2003) Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. Proceedings of the National Academy of Sciences of the United States of America 100:8024-8029.

34. Dita M. A., Rispail N., Prats E., Rubiales D., Singh K. B. (2006) Biotechnology approaches to overcome biotic and abiotic stress constraints in legumes. Euphytica 127:1-24.

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

36. Ellis J. G., Dodds P. N., and Lawrence G. J. (2007) Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. Annual Review of Phytopathology 45:289-306.

37. Ellis J., Catanzariti A. M., & Dodds P. (2006) The problem of how fungal and oomycete avirulence proteins enter plant cells. Trends in plant science 11(2): 61-63.

38. Flor H. H. (1955) Host-parasite interaction in flax rust-its genetics and other implications. Phytopathology 45:680-685.

39. Flor H.H. (1971) Current status of the gene-for-gene concept. Annual Review Phytopathology 9: 275-296.

40. Frederick R. D., Snyder C. L., Peterson G. L., Bonde M. R. (2002) Polymerase chain reaction assays for the detection and discrimination of the soybean rust pathogens *Phakopsora pachyrhizi* and *P. meibomiae*. Phytopathology 92:217-227.

41. Fritz-Laylin L. K., Krishnamurthy N., Tor M., Sjolander K. V., and Jones J. D. (2005) Phylogenomic analysis of the receptor-like proteins of rice and *Arabidopsis*. Plant Physiology 138:611-623.

42. Gao H., Narayanan N. N., Ellison L., Bhattacharyya M. K. (2005) Two classes of highly similar coiled coil-nucleotide binding-leucine rich repeat genes isolated from the *Rps1-k* locus encode *Phytophthora* resistance in soybean. Molecular Plant-Microbe Interactions Journal 18:1035-1045.

43. Garcia A., Sanches Calvo E., Souza Kiihl R., Harada A. (2008) Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: discovery of a novel locus and alleles. Theoretical and Applied Genetics 117:545-553.

44. Giesler L., Kemerait R. and Sconyers L. (2007) The sentinel plot system: Monitoring movement of an invasive pathogen, in: A. E. Dorrance, M. A. Draper, and D. E. Hershman (Eds.), Using Foliar Fungicides to Manage Soybean Rust, 2nd edition, The Ohio State University, Columbus, OH. USA. pp. 35-38.

45. Goellner K., Loehrer M., Langenbach C., Conrath U., Koch E. and Schaffrath U. (2010) *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. Molecular Plant Pathology 11:169-177.

46. Grau C. R., Dorrance A. E., Bond J. S., Russin J. S. (2004) Fungal diseases, in: H. R. Boerma and J. E. Specht (Eds.), Soybeans: Improvement, Production and Uses, 3rd edition, ASA, CSSA, and SSSA, Madison, WI, USA. pp. 679-763.

47. Gu B., Kale S. D., Wang Q., Wang D., Pan Q., Cao H., Meng Y., Kang .Z, Tyler B. M., and Shan W. (2011) Rust secreted protein Ps87 is conserved in diverse fungal pathogens and contains a RXLR-like motif sufficient for translocation into plant cells. PLoS ONE 6(11):e27217.

48. Gu K., Yang B., Tian D., Wu L., Wang D., Sreekala C., Yang F., Chu Z., Wang G. L., White F. F., and Yin Z. (2005) R gene expression induced by a type-III effector triggers disease resistance in rice. Nature 435:1122-1125.

49. Guo Y. L., Fitz J., Schneeberger K., Ossowski S., Cao J., Weigel D. (2011) Genome-wide comparison of NB-LRR encoding genes in *Arabidopsis*. Plant Physiology 157:757-769.

50. Hartman G. L., Miles M. R. and Frederick R. D. (2005) Breeding for resistance to soybean rust. Plant Disease 89:664-666.

51. Hartman G. L., Sinclair J. B., Rupe J.C. (Eds.) (1999) Compendium of soybean diseases. 4th edition. APS Press, St. Paul, MN, USA.

52. Hartman G. L., Wang T. C. and Tschanz A. T. (1991) Soybean rust development and the quantitative relationship between rust severity and soybean yield. Plant Disease 75:596-600.

53. Hartman G. L., Wang T. C., and Hymowitz T. (1992) Sources of resistance to soybean rust in perennial *Glycine* species. Plant Disease 76:396-399.

54. Hartwig E. E. (1986) Identification of a fourth major gene conferring resistance to soybean rust. Crop Science 26:1135-1136.

55. Hartwig E. E. and Bromfield K. R. (1983) Relationships among three genes conferring specific resistance to soybean rust. Crop Science 23:237-239.

56. Hayes A. J., Jeong S. C., Gore M. A., Yu Y. G., Buss G. R., Tolin S. A., Maroof M. A. (2004) Recombination within a nucleotide-binding-site/leucine-rich-repeat gene cluster produces new variants conditioning resistance to soybean mosaic virus in soybeans. Genetics 166:493-503.

57. He C. Y., Tian A. G., Zhang J. S., Zhang Z. Y., Gai J. Y., Chen S. Y. (2003) Isolation and characterization of a full-length R-gene homolog from soybean. Theoretical and Applied Genetics 106(5):786-793.

58. Hennings V. P. (1903) A few new Japanese Uredinaceae. Hedwigia 42:S107-108.

59. Hershman D. E. (2009) Changes in the 2010 Sentinel Plot System. In: Proceeding of 4th National Soybean Rust Symposium. American Phytopathological Society.

60. Hershman D. E., Sikora E. J. and Giesler L. J. (2011) Soybean rust PIPE: Past, present, and future. Online. Journal of Integrated Pest Management. doi:http://dx.doi.org/10.1603/IPM11001

61. Hogenhout S. A., Van der Hoorn R. A., Terauchi R., Kamoun S. (2009) Emerging concepts in effector biology of plant-associated organisms. Molecular Plant-Microbe Interactions 22:115-122.

62. Hok S., Attard A. and Keller H. (2010) Getting the Most from the Host: How pathogens force plants to cooperate in disease. Molecular Plant-Microbe Interactions 23:1253-1259.

63. Hu G., and Hulbert S. H. (1994) Evidence for involvement of gene conversion in meiotic instability of the *Rp1* rust resistance genes of maize. Genome 37:742-746.

64. Hulbert S. H. (1997) Structure and evolution of the *rp1* complex conferring rust resistance in maize. Annual Review of Phytopathology 35:293-310.

65. Hulbert S. H., Webb C. A., Smith S. M. and Sun Q. (2001) R-gene complexes: evolution and utilization. Annual Review of Phytopathology 39:285-312.

66. Hymowitz T. (1970) On the domestication of the soybean. Economic Botany 24:408-421.

67. Hymowitz T. (1987) Introduction of the soybean to Illinois. Economic Botany 41:28-32.

68. Hymowitz T. (2008) The history of the soybean, in: L. A. Johnson, P. J. White and R. Galloway (Eds), Soybeans Chemistry, Production, Processing, and Utilization. AOCS Press, Urbana, IL, USA. pp. 1-31.

69. Hymowitz T., Newell C. A. (1980) Taxonomy, speciation, domestication, dissemination, germplasm resources and variation in the genus *Glycine*, in: R. J. Summerfield and A. H. Bunting (Eds.), Advances in Legume Science. Royal Botanic Gardens, Kew, UK. pp. 251-264.

70. Hymowitz T., Shurtleff W. R. (2005) Debunking soybean myths and legends in the historical and population literature. Crop Science 45:473-476.

71. Hyten D. L, Hartman G. L., Nelson R. L., Frederick R. D., Concibido V. C., Narvel J. M., Cregan P. B. (2007) Map location of the *Rpp1* locus that confers resistance to soybean rust in soybean. Crop Science 47:837-840.

72. Hyten D. L., Smith J. R., Frederick R. D., Tucker M. L., Song Q., Cregan P. B. (2009) Bulked segregant analysis using the GoldenGate assay to locate the *Rpp3* locus that confers resistance to soybean rust in soybean. Crop Science 49:265-271.

73. Isard S. A., Russo J. M., and DeWolf E. D. (2006) The establishment of a national Pest Information Platform for Extension and Education. Online. Plant Health Progress doi:10.1094/PHP-2006-0915-01-RV

74. Jarosz A. M., and Burdon J. J. (1990) Predominance of a single major gene for resistance to *Phakopsora pachyrhizi* in a population of *Glycine argyrea*. Heredity 64:347-353

75. Jiang G. H, Xia Z. H., Zhou Y. L., Wan J., Li D. Y., Chen R. S., Zhai W. X., Zhu L. H. (2006) Testifying the rice bacterial blight R-gene *xa5* by genetic complementation and further analyzing *xa5* (*Xa5*) in comparison with its homolog *TFIIAγ1*. Molecular Genetics and Genomics 275:354−366.

76. Jones D. A., Thomas C. M., Hammond-Kosack K. E., Balint-Kurti P. J., and Jones J. D. (1994) Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science 266:789-793.

77. Jones J., Parniske M., Thomas T., Hammond-Kosack K., Romes T., Piedras P., Tai T., Tores M. A., Hatzixanthus K., Brading P. and Wulf B. H. (1997) Evolution and function of tomato *Cf*-disease R-genes In: Plant Disease R-gene Function: An EMBO Workshop, May 18-29, Maratea, Italy

78. Joosten M. H., Cozijnsen TJ., and de Wit P. J. (1994) Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene. Nature 367:384-386.

79. Jordan S. A., Mailhot D. J., Gevens A. J., Marois J. J., Wright D. L., Harmon C. L. and Harmon P. F. (2010) Characterization of kudzu (*Pueraria* spp.) resistance to *Phakopsora pachyrhizi*, the causal agent of soybean rust. Phytopathology 100:941-948.

80. Kale S. D., Gu B., Capelluto D. G. S., Dou D. L., Feldman E., Rumore A., Arredondo F. D., Hanlon R., Fudal I., Rouxel T., Lawrence C. B., Shan W., Tyler B. M. (2010) External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. Cell 142:284-295.

81. Kamoun S. (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. Annual Review of Phytopathology 44:41-60.

82. Kamoun S. (2007) Groovy times: filamentous pathogen effectors revealed. Current Opinion in Plant Biology 10(4):358-65.

83. Kanazin V., Marek L. F. and Shoemaker R. C. (1996) R-gene analogs are conserved and clustered in soybean. Proceedings of the National Academy of Sciences 93:11746-11750.

84. Kang Y. J., Kim K. H., Shim S., Yoon M. Y., Sun S., Kim M. Y., Van K., Lee S. H. (2012) Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. BMC Plant Biology 12:139.

85. Kato M., Yorinori J. T. (2008) A study on a race composition of *Phakopsora pachyrhizi* in Brazil: a difficulty of race identification, in: H. Kudo, K. Suenaga, R. M. Soares, A. Toledo (Eds.), Facing the challenge of soybean rust in South America, JIRCAS Working Report No. 58. JIRCAS, Tsukuba, pp. 94-98.

86. Keen N. T. (1990) Gene-for-gene complimentary in plant-pathogen interactions. Annual Review of Genetics 24:447-463.

87. Kemen E., Kemen A. C., Rafiqi M., Hempel U., Mendgen K., Hahn M., *et al*. (2005) Identification of a protein from rust fungi transferred from haustoria into infected plant cells. Molecular Plant-Microbe Interactions 18:1130-1139. doi: 10.1094/MPMI-18-1130

88. Kemerait B. (2014) Soybean Disease and Nematode Control, in Georgia Soybean Guide 2014. pp 22-36.

89. Kendrick M. D., Harris D. K., Ha B-K., Hyten D. L., Cregan P. B., Frederick R. D., Boerma H. R., Pedley K. F. (2011) Identification of a second Asian soybean rust resistance gene in Hyuuga soybean. Phytopathology 101:535-543

90. Killgore E. and Heu R. (1994) First report of soybean rust in Hawaii. Plant Disease 78:1216.

91. Kirk P. M., Cannon P. F., Minter D. W., Stalpers J. A. (2008) Ainsworth and Bisby's dictionary of the Fungi. CAB International, Wallingford, UK.

92. Kitani X. and Inoue Y. (1960) Studies on the soybean rust and its control measure. Part 1. Studies on the soybean rust. Bulletin of the Shikoku Agricultural Experiment Station (Zentsuji, Japan) 5:319-342.

93. Kobe B. and Deisenhofer J. (1995) Proteins with leucine-rich repeats. Current Opinion in Structural Biology 5:409-416.

94. Koch E., Ebrahimnesbat F. and Hoppe H. H. (1983) Light and electron microscopic studies on the development of soybean rust (*Phakopsora pachyrhizi* Syd) in susceptible soybean leaves. Phytopathology 106:302-320.

95. Kochman J. K. (1977) Soybean rust in Australia, in: R.E. Ford and J. B. Sinclair (Eds.), Rust of soybean, The problem and research needs. INTSOY, Series No. 12 University of Illinois, Urbana, USA. pp. 44-48.

96. Kohler A., Rinaldi C., Duplessis S., Baucher M., Geelen D., Duchaussoy F., Meyers B. C., Boerjan W., Martin F. (2008) Genome-wide identification of NBS R-genes in *Populus trichocarpa*. Plant Molecular Biology 66:619-636.

97. Kuang H., Wei F., Marano M., Wirtz U., Wang X., Liu J., Shum W., Zaborsky J., Tallon L., Rensink W. (2005) The R1 resistance gene cluster contains three groups of independently evolving, type I R1 homologues and shows substantial structural variation among haplotypes of *Solanum demissum*. Plant J. 44 (1):37-51.

98. Leach J. E., Vera-Cruz C. M., Bai J. F. and Leung H. (2001) Pathogen fitness penalty as a predictor of durability of disease resistance genes. Annual Review of Phytopathology 39:187-224.

99. Lee S., Costanzo S., Olsen K., Caicedo A. and Jia Y. (2009) Evolutionary dynamics of the blast resistance gene *Pi-ta* in AA genome group of *Oryza* species during domestication. Genetics 183:1315-1325.

100. Leister D. (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease R-genes. Trends in Genetics 20:116-122.

101. Leonard K. J. (1997) Modeling gene frequency dynamics, in: I. R. Crute, E. B. Holub, J. J. Burdon (Eds.), The Gene-for-Gene Relationship in Plant-Parasite Interactions, CAB International, London, UK. pp. 211-30.

102. Levy C. (2005) Epidemiology and chemical control of soybean rust in Southern Africa. Plant Disease 89:669-674.

103. Li J., Ding J., Zhang W., Zhang Y., Tang P., Chen J.-Q., Tian D. and Yang S. (2010) Unique evolutionary pattern of numbers of gramineous NBS-LRR genes. Molecular Genetics and Genomics 283:427-438.

104. Li S., Smith J. R., Ray J. D., Frederick R. D. (2012) Identification of a new soybean rust resistance gene in PI 567102B. Theoretical and Applied Genetics 125:133-142.

105. Lin F., Zhao M., Baumann D. D., Ping J., Sun L., Liu Y., Zhang B., Tang Z., Hughes E., Doerge R. W., Hughes T. J., Ma J. (2014) Molecular response to the pathogen *Phytophthora sojae* among ten soybean near isogenic lines revealed by comparative transcriptomics. BMC Genomics 15:18.

106. Lin S. Y. (1966) Studies on the physiologic races of soybean rust fungus, *Phakopsora pachyrhizi* Syd. (in Chinese with English summary) Journal of Taiwan Agricultural Research 15:24-28.

107. Liu J., Liu X., Dai L., Wang G. (2007) Recent progress in elucidating the structure, function and evolution of disease R-genes in plants. Journal of Genetics and Genomics 34:765-776.

108. Macho A. P. and Zipfel C. (2014) Plant PRRs and the activation of innate immune signaling. Molecules and Cells 54:263-272.

109. Mackey D., Belkhadir Y., Alonso J. M., Ecker J. R. and Dangl J. (2003) *Arabidopsis* RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. Cell 112:379-389.

110. Matsushima N., Miyashita H. (2012) Leucine-rich repeat (LRR) domains containing intervening motifs in plants. Biomolecules 2:288-311.

111. McDonald B. A. and Linde C. (2002) Pathogen population genetics, evolutionary potential, and durable resistance. Annual Review of Phytopathology 40:349-379.

112. McIntosh R. A, Brown G. N. (1997) Anticipatory breeding for resistance to rust diseases in wheat. Annual Review of Phytopathology 35: 311-326.

113. McGregor R. (1973) The Emigrant Pests. A report to the Administrator of APHIS by Dr. Russell McGregor.

114. McHale L. K., Truco M. J., Kozik A., Wroblewski T., Ochoa O. E., Lahre K. A., Knapp S. J. and Michelmore R. W. (2009) The genomic architecture of disease resistance in lettuce. Theoretical and Applied Genetics 118:565-580.

115. McLean R. J., and Byth D. E. (1980) Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybeans. Australian Journal of Agricultural Research 31:951-956.

116. McLean R. J., Byth D. E. (1976) Resistance of soybean to rust in Australia. Austral Plant Pathology Soceity Newsletter 5:34-36.

117. Melchinger A. E., Lee M., Lamkey K. R. and Woodman W. L. (1990) Genetic diversity for restriction fragment length polymorphisms: Relation to estimated genetic effects in maize inbreds. Crop Science 30:1033-1040.

118. Meyer J. D., Silva D. C., Yang C., Pedley K. F., Zhang C., van de Mortel M., Hill J. H., Shoemaker R. C., Abdelnoor R. V., Whitham S. A., *et al*. (2009) Identification and analyses of candidate genes for *rpp4*-mediated resistance to Asian soybean rust in soybean. Plant Physiology 150:295-307.

119. Meyers B. C., Dickerman A. W., Michelmore R. W., Sivaramakrishnan S., Sobral B. W. and Young N. D. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. The Plant Journal 20:317-332.

120. Meyers B. C., Kaushik S., Nandety R. S. (2005) Evolving disease R-genes. Current Opinion in Plant Biology 8:129-134.

121. Meyers B. C., Kozik A., Griego A., Kuang H., Michelmore R. W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell 15: 809-834.

122. Michelmore R. W., Meyers B. C. (1998) Clusters of R-genes in plants evolve by divergent selection and a birth-and-death process. Genome Research 8:1113-1130.

123. Miles M. R., Bonde M. R., Nester S. E., Berner D. K., Frederick R. D. and Hartman G. L. (2011) Characterizing resistance to *Phakopsora pachyrhizi* in soybean. Plant Disease 95:577-581.

124. Monteros M. J., Missaoui A. M., Phillips D. V., Walker D. R., Boerma H. R. (2007) Mapping and confirmation of the 'Hyuuga' redbrown lesion resistance gene for Asian soybean rust. Crop Science 47:829-836.

125. Morel W. P. (2001) Roya de la soja. Ministerio de Agricultura y Ganaderia, Subsecretaria de Agricultura, Direccion de Investigacion Agricola, Centro Regional de Investigacion Agricola-CRIA, Capitan Miranda, Itapua, Paraguay. Comunicado Tecnico-Report Official, Serie Fitopathologia, 1, Junio de 2001.

126. Mucyn T. S. *et al*. (2006) The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity. Plant Cell 18:2792-2806.

127. Mun J-H., Yu H-J., Park S., Park B-S. (2009) Genome-wide identification of NBS-encoding R-genes in *Brassica rapa*. Molecular Genetics and Genomics 282:617-631.

128. National Potato Council (2003) Potato Statistics: Seed Certification, Approved Acreage, Varieties Planted, and Production.

129. Niblack T. L., Tylka G. L., Riggs R. D. (2004) Nematode pathogens of soybean, in: H. R. Boerma and J. E. Specht (Eds.), Soybeans: Improvement, Production and Uses, 3rd edition, ASA, CSSA, and SSSA, Madison, WI, USA. pp. 821-851.

130. Ogle H. J., Byth D. E. and McLean R. J. (1979) Effect of rust (*Phakopsora pachyrhizi*) on soybean yield and quality in South-eastern Queensland. Australian Journal of Agricultural Research 30:883-893.

131. Ono Y., Buritica P. and Hennen J. F. (1992) Delimitation of *Phakopsora Physopella*, and *Cerotelium* and their species on leguminosae. Mycological Research 96:825-850.

132. Orf J. (2010) Introduction, in: K. Bilyeu, M. B. Ratnaparkhe and C. Kole (Ed.), Genetics, Genomics and Breeding of Soybean, CRC Press and Science Publishers, Enfield, NH, USA. pp. 17-18.

133. Pan Q., Wendel J., and Fluhr R. (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. Journal of molecular evolution 50.3:203-213.

134. Panthee D. R., Yuan J. S., Wright D. L., Marois J. J., Mailhot D. and Stewart J. R. C. N. (2007) Gene expression analysis in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) in an early growth stage. Functional and Integrative Genomics 7:291-301.

135. Parlevliet J. E. (1985) Race and pathotype concepts in parasitic fungi. OEPP/EPPO Bulletin. 15:145-150.

136. Patzoldt M. E., Tyagi R. K., Hymowitz T., Miles M. R., Hartman G. L. and Frederick R. D. (2007) Soybean rust resistance derived from *Glycine tomentella* in amphiploid hybrid lines. Crop Science 47:158-161.

137. Pham T. A., Miles M. R., Frederick R. D., Hill C. B. and Hartman G. L. (2009) Differential responses of resistant soybean entries to isolates of *Phakopsora pachyrhizi*. Plant Disease 93:224-228.

138. Plett J. M., Kemppainen M., Kale S. D., Kohler A., Legue V., Brun A., Tyler B. M., Pardo A. G., Martin F. (2011) A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. Current Biology 21:1197-1203.

139. Porter B. W., Paidi M., Ming R., Alam M., Nishijima W. T., and Zhu Y. J. (2009) Genome-wide analysis of *Carica papaya* reveals a small NBS R-gene family. Molecular Genetics and Genomics 281:609-626.

140. Posada-Buitrago M. L. and Frederick R. D. (2005) Expressed sequence tag analysis of the soybean rust pathogen Phakopsora pachyrhizi. Fungal Genetics and Biology 42:949-962.

141. Pretorius Z. A., Kloppers R. J., Frederick R. D. (2001) First Report of soybean rust in South Africa. Plant Disease 85:1288.

142. Radwan O., Gandhi S., Heesacker A., Whitaker B., Taylor C., Plocik A., Kesseli R., Kozik A., Michelmore R.W., Knapp S. J. (2008) Genetic diversity and genomic distribution of homologs encoding NBS-LRR disease resistance proteins in sunflower. Molecular Genetics and Genomics 280:111-125.

143. Rafiqi M., Gan P. H. P., Ravensdale M., Lawrence G. J., Ellis J. G., Jones D. A., Hardham A. R., Dodds P. N. (2010) Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. Plant Cell 22:2017-2032.

144. Ramstedt M., Hurtado S., and Astrom B. (2002) Pathotypes of *Melampsora* rust on Salix in short-rotation forestry plantations. Plant Pathology 51:185-190.

145. Ratnaparkhe M. B., Singh R. J., Doyle J. J. (2011) Glycine, in C. Kole (Ed.), Wild Crop Relatives: Genomic and Breeding Resources Legume Crops and Forages, Springer, New York, USA. pp. 83-116.

146. Ravensdale M., Nemri A., Thrall P. H., Ellis J. G., Dodds P. N. (2011) Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. Molecular Plant Pathology 12(1):93-102.

147. Richter T. E., Pryor T. J., Bennetzen J. L. and Hulbert S. H. (1995) New rust resistance specificities associated with recombination in the *Rp1* complex in maize. Genetics 141:373-381.

148. Ronald P. C., Albano B., Tabien R., Abenes L., Wu K. S., McCouch S., and Tanksley S. D. (1992) Genetic and physical analysis of the rice bacterial blight disease resistance locus, *Xa21*. Molecular Genetics and Genomics 236:113-120.

149. Rose L. E., Bittner-Eddy P. D., Langley C. H., Holub E. B., Michelmore R. W., Beynon J. L. (2004) The maintenance of extreme amino acid diversity at the disease R-gene, *RPP13*, in *Arabidopsis thaliana*. Genetics 166:1517-1527.

150. Rossi R. L. (2003) First report of *Phakopsora pachyrhizi*, the causal organism of soybean rust in the province of Misiones, Argentina. Plant Disease 87:102.

151. Rupe J., Sconyers L. (2008) Soybean rust. Plant Health Instruction doi:10.1094/PHI-I-2008-0401-01

152. Rytter J. L., Dowler W. M. and Bromfield K. R. (1984) Additional alternative hosts of *Phakopsora* p*achyrhizi*, causal agent of soybean rust. Plant Disease 68:818-819.

153. Saksirirat W., Hoppe H. H. (1991) Degradation of uredospores of the soybean rust fungus (*Phakopsora pachyrhizi* Syd) by cell-free culture filtrates of the mycoparasite *Verticillium psalliotae* Treschow. Journal of Phytopathology 132:33-45

154. Schafer J. F., Roelfs A. P. (1985) Estimated relation between numbers of urediniospores of *Puccinia graminis* f. sp. *tritici* and rates of occurrence of virulence. Phytopathology 75:749-750.

155. Schmutz J. *et al*. (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178-183.

156. Schneider R. W., Hollier C. A. , Whitam H. K., Palm M. E., McKemy J. M., Hernandez J. R., Levy L. and DeVries-Paterson R. (2005) First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. Plant Disease 89:774.

157. Scofield S. R., Tobias C. M., Rathjen J. P., Chang J. H., Lavelle D. T., Michelmore R. W. and Staskawicz B. J. (1996) Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. Science 274:2063-2065.

158. Sels J., Mathys J., de Coninck B. M. A., Cammue B. P. A., de Bolle M. F. C. (2008) Plant pathogenesis-related (PR) proteins: A focus on PR peptides. Plant Physiology and Biochemistry. 46:941-950.

159. Shan L., He P., Li J., Heese A., Peck S. C., Nurnberger T., Martin G. B., and Sheen J. (2008) Bacterial effectors target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. Cell Host and Microbe 4:17-27.

160. Shaner G., Stomberg E., Lacy G., Barker K. and Pirone T. (1992) Nomenclature and concepts of pathogenicity and virulence. Annual Review of Phytopathology 30:47-66.

161. Shao Z. Q., Zhang Y. M., Hang Y. Y., Xue J. Y., Zhou G. C., Wu P., ... & Chen J. Q. (2014) Long-Term Evolution of Nucleotide-Binding Site-Leucine-Rich Repeat Genes: Understanding Gained from and beyond the Legume Family. Plant physiology 166(1):217-234.

162. Silva D. C. G., Yamanaka N., Brogin R. L., Arias C. A. A., Nepomuceno A. L., Di Mauro A. O., Pereira S. S., Nogueira L. M., Passianotto A. L. L., Abdelnoor R. V. (2008) Molecular mapping of two loci that confer resistance to Asian rust in soybean. Theoretical and Applied Genetics 117:57-63.

163. Sinclair J. B. (1989) Threats to production in the tropics: Red leaf blotch and leaf rust. Plant Disease 73:604-604.

164. Sinclair J. B., Hartman G. L. (1999) Soybean rust, in: G. L. Hartman, J. B. Sinclair, J. C. Rupe (Eds.), Compendium of soybean diseases. APS Press, St. Paul, MN, USA pp. 25-26.

165. Singh B. B., Gupta S. C. and Singh B. B. (1975) Sources of field resistance to rust and yellow mosaic diseases of soybean. Indian Journal of Genetics and Plant Breeding 34:400-404.

166. Smith S. M., Pryor A. J. and Hulbert S. H. (2004) Allelic and haplotypic diversity at the *Rp1* rust resistance locus of maize. Genetics 167:1939-1947.

167. Smyth A. H. (1907) Writings of Benjamin Franklin, vol 5. Macmillan, New York, USA.

168. Soares R. M. (2008) Practical soybean rust identification, in: H. Kudo, K. Suenaga, R. M. Soares, A. Toledo (Eds.), Facing the challenge of soybean rust in South America, JIRCAS Working Report No. 58. JIRCAS, Tsukuba, pp. 66-69.

169. Souza T. L, Dessaune S. N., Moreira M. A., Barros E. G. (2014) Soybean rust resistance sources and inheritance in the common bean (*Phaseolus vulgaris* L.) Genetics and Molecular Research 13(3):5626-5636.

170. Staskawicz B. J., Ausubel F. M., Baker B. J., Ellis J. G. and Jones J. D. (1995) Molecular genetics of plant disease resistance. Science 268:661-667.

171. Tameling W. I., Elzinga S. D., Darmin P. S., Vossen J. H., Takken F. L., Haring M. A. and Cornelissen B. J. (2002) The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. Plant Cell 14:2929-2939.

172. Tang X., Frederick R. D., Zhou J., Halterman D. A., Jia Y. and Martin G. B. (1996) Initiation of plant disease resistance by physical interaction of *AvrPto* and *Pto* Kinase. Science 274:2060-2063.

173. Tian M. Y., Win J., Savory E., Burkhardt A., Held M., Brandizzi F., Day B. (2011) 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. Molecular Plant-Microbe Interactions 24:543-553.

174. Traut T. W. (1994) The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. European Journal of Biochemistry 222:9-19.

175. Tremblay A., Hosseini P., Alkharouf N. W., Li S., Matthews B. F. (2010) Transcriptome analysis of a compatible responses by *Glycine max* to *Phakopsora pachyrhizi* infection. Plant science 179:183-193.

176. Tremblay A., Hosseini P., Alkharouf N., Li S. and Matthews B. F. (2010) Transcriptome analysis of a compatible response by *Glycine max* to *Phakopsora pachyrhizi* infection. Plant Science 179:183-193.

177. Tremblay A., Hosseini P., Li S., Alkharouf N. W. and Matthews B. F. (2013) Analysis of *Phakopsora pachyrhizi* transcript abundance in critical pathways at four time-points during infection of a susceptible soybean cultivar using deep sequencing. BMC Genomics 14:614. doi: 10.1186/1471-2164-14-614

178. Tremblay A., Li S., Scheffler B. E. and Matthews B. F. (2009) Laser capture microdissection and expressed sequence tag analysis of uredinia formed by *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. Physiological and Molecular Plant Pathology 73:163-174.

179. Twizeyimana M., Hartman G. L. (2012) Pathogenic variation of *Phakopsora pachyrhizi* isolates on soybean in the United States from 2006 to 2009. Plant Disease 96:75-81.

180. Twizeyimana M., Ojiambo P. S., Sonder K., Ikotun T., Hartman G. L., Bandyopadhyay R. (2009) Pathogenic variation of *Phakopsora pachyrhizi* infecting soybean in Nigeria. Phytopathology 99:353-361.

181. Tyler B. M. (2009) Entering and breaking: virulence effector proteins of oomycete plant pathogens. Cellular Microbiology 11:13-20.

182. U.S. Dep. Agric. (2005) A coordinated framework for soybean rust surveillance, reporting, prediction, management and outreach. Online. Soybean Rust Resources, Emergency and Domestic Programs, Plant Protection and Quarantine (PPQ), USDA-APHIS, Washington, DC, USA.

183. Upadhyaya N. M., Mago R., Staskawicz B. J., Ayliffe M. A., Ellis J. G. and Dodds P. N. (2014) A bacterial type III secretion assay for delivery of fungal effector proteins into wheat. Molecular Plant-Microbe Interactions 27:255-264. doi: 10.1094/MPMI-07-13-0187-FI

184. Vakili N. G. (1979) Field survey of endemic leguminous hosts of *Phakopsora pachyrhizi* in Puerto Rico. Plant Disease Reporter 63:931-935.

185. van de Biezen E. A., Jones J. D. H. (1998) Plant disease-resistance proteins and the gene for gene concept. Trends in Biochemical Sciences 24:47-53.

186. van de Hoorn R. A. and Kamoun S. (2008) From guard to decoy: a new model for perception of plant pathogen effectors. Plant Cell 20:2009-2017.

187. van de Mortel M., Recknor J. C., Graham M. A., Nettleton D., Dittman J. D., Nelson R. T., Godoy C.V., Abdelnoor R. V., Almeida A. M. R., Baum T. J. and Whitham S. A.

(2007) Distinct biphasic mRNA changes in response to Asian soybean rust infection. Molecular Plant-Microbe Interactions 20:887-899.

188. van Loon L. C., van Strien E. A. (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiological and Molecular Plant Pathology 55:85-97.

189. Voegele R. T., Mendgen, K. (2003) Rust haustoria: nutrient uptake and beyond. New Phytologist 159: 93-100.

190. Wan Hongjian., *et al*. (2012) Analysis of TIR-and non-TIR-NBS-LRR disease resistance gene analogous in pepper: characterization, genetic variation, functional divergence and expression patterns. BMC genomics 13.1: 502.

191. Wang B. J., Wang Y. J., Wang Q., Luo G. Z., Zhang Z. G., He C. Y., He S. J., Zhang J. S., Gai J. Y., Chen S. Y. (2004) Characterization of an NBSLRR R-gene homologue from soybean. Journal of Plant Physiology 161 (7):815-822.

192. Welter L. J., Göktürk-Bayder N., Akkurt M., Maul E., Eibach R., *et al*. (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Molecular Breeding 20:359-374.

193. White F. F. and Yang B. (2009) Host and pathogen factors controlling the rice-*xanthomonas oryzae* interation. Plant Physiology. 150:4, 1677-1686.

194. Wicker T., Yahiaoui N., Keller B. (2007) Contrasting rates of evolution in *Pm3* loci from three wheat species and rice. Genetics 177(2):1207-1216.

195. Wilcox J. R. (2004) World distribution and trade of soybean, in: H. R. Boerma and J. E. Specht (Eds.), Soybeans: Improvement, Production and Uses, 3rd edition, ASA, CSSA, and SSSA, Madison, WI, USA. pp. 1-14.

196. Wilson R. F. (2004) Seed Composition, in: H. R. Boerma and J. E. Specht (Eds.), Soybeans: Improvement, Production and Uses, 3rd edition, ASA, CSSA, and SSSA, Madison, WI, USA. pp. 621-671.

197. Wilson R. F. (2008) Soybean: Market Driven Research Needs, in: G. Stacey (Ed.), Genetics and Genomics of Soybean. Springer, New York, USA. pp. 3-15.

198. Xiao W. K., Xu M. L., Zhao J. R., Wang F. G., Li J. S., Dai J. R. (2006) Genome-wide isolation of R-gene analogs in maize (*Zea mays* L.). Theoretical and Applied Genetics 113:63-72.

199. Yamaoka Y., Fujiwara Y., Kakishima M., Katsuya K., Yamada K., Hagiwara H. (2002) Pathogenic races of *Phakopsora pachyrhizi* on soybean and wild host plants collected in Japan. Journal of General Plant Pathology 68:52-56.

200. Yang S., Feng Z., Zhang X., Jiang K., Jin X., Hang Y., Chen J. Q., Tian D. (2006) Genome-wide investigation on the genetic variations of rice disease R-genes. Plant Molecular Biology 62:181-193.

201. Yang S., Zhang X., Yue J., Tian D., Chen J. Q. (2008) Recent duplication dominates the expansion of the NBS-encoding genes in two woody species. Molecular Genetics and Genomics 280:187-198.

202. Yeh C. C. (1983) Physiological races of *Phakopsora pachyrhizi* in Taiwan. Journal of Agricultural Research of China 32:69-74.

203. Yeh C. C., and Yang C. Y. (1975) Yield loss caused by soybean rust, *Phakopsora pachyrhizi*. Plant Protection Bulletin 17:7-8.

204. Yorinori J. T. (2008) Soybean germplasms with resistance and tolerance to Asian rust and screening methods, in: H. Kudo, K. Suenaga, R. M. Soares, A. Toledo (Eds.), Facing

the challenge of soybean rust in South America, JIRCAS Working Report No. 58. JIRCAS, Tsukuba, pp. 66-69.

205. Yorinori J. T., Paiva W. M., Frederick R. D. and Fernandez F. T. P. (2002) Ferrugem da soja (Phakopsora pachyrhizi) no Brasil e no Paraguai, nas safras 2000/01 e 2001/02, p. 94. In: Resumos da II congreso Brasilerio de soja e mercosoja 2002. Embrapa soja, Londrina, 2002 393 p. (Documantos/Embrapa soja: n. 181. Rusemo 094).

206. Yorinori J. T., Paiva W. M., Frederick R. D., Constamilan L. M., Bertagnoli P. F., Hartman G. L., Godoy C. V. and Nunes J. J. (2005) Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. Plant Disease 89:675-677.

207. Young N. D. (2000) The genetic architecture of resistance. Current Opinion in Plant Biology 3:285-290.

208. Zhang X. C., Freire M. C. M., Le M. H., De Oliveira L. O., Pitkin J. W., Segers G., Concibido V. C., Baley G. J., Hartman G. L., Upchurch G., Pedley K.F. and Stacey G. (2012) Genetic Diversity and Origins of *Phakopsora pachyrhizi* isolates in the United States. Asian Journal of Plant Pathology 6:52-65.

209. Zhang X., Feng Y., Cheng H., Tian D., Yang S., Chen J. Q. (2011) Relative evolutionary rates of NBS-encoding genes revealed by soybean segmental duplication. Molecular Genetics and Genomics 285: 79-90.

210. Zhou B., Dolan M., Sakai H., Wang G. L. (2007) The genomic dynamics and evolutionary mechanism of the *Pi2/9* locus in rice. Molecular Plant-Microbe Interactions 20:63-71.

211. Zhou T., Wang Y., Chen J. Q., Araki H., Jing Z., Jiang K., Shen J., Tian D. (2004) Genome-wide identification of NBS genes in rice reveals significant expansion of divergent non-TIR NBS Genes. Molecular Genetics and Genomics 271:402-415.

212. Zipfel C. (2014) Plant pattern-recognition receptors. Trends in Immunology 35:345-351.

CHAPTER 2

CHARACTERIZATION OF HOST RESISTANCE OF SOYBEAN AND PERENNIAL *GLYCINE* SPECIES AND GENETIC DIVERSITY OF ASIAN SOYBEAN RUST IN GEORGIA[1]

[1] P. Tian and S. M. Smith. To be submitted to Phytopathology.

Abstract

Asian soybean rust (ASR) is an economically important disease caused by the fungus *Phakopsora pachyrhizi* Sydow. Resistant germplasm have been identified that carry *Rpp* genes and confer race-specific resistance to different *P. pachyrhizi* isolates. However, race-specific resistance is not durable due to pathogen variation among the current ASR population. There are no resistant commercial soybean cultivars currently grown in the USA. In addition, little is known in regards to the genetic diversity and population structure of the current *P. pachyrhizi* isolates in the USA. In this study, eight field isolates were collected from different Georgia soybean fields in 2012 and 2013. Three isolates were identified based on phenotypic variation when inoculated onto a set of soybean differentials. When inoculated with the characterized Georgia isolates, one Japanese (Hyuuga) cultivar and eight soybean wild relatives were resistant, while all of the Georgia released soybean cultivars were susceptible. As a result, Hyuuga was selected as a candidate for RNA sequence analysis. To further characterize the *P. pachyrhizi* isolates in Georgia and their genetic diversity, the Internal Transcribed Spacer (ITS) sequences and the ADP Ribosylation Factor (ARF) sequences were cloned from the eight field isolates. Variability in the ITS sequences indicated that the Georgia *P. pachyrhizi* isolates were diverse and contained genotypes that were identical to genotypes from other countries (Brazil, Zimbabwe, India, Australia, and Thailand) and neighboring Georgia states (Missouri, Arkansas, Florida, Alabama, Louisiana, and Mississippi). Georgia *P. pachyrhizi* genotypes were also identical to the reference genotypes identified in Australia and Hawaii, suggesting that there might be a possible second path for the migration of ASR from East Asia to the USA in addition to the most prevalent hypothesis of ASR migration. Two predominate genotypes were identified in all of the collected isolates, while one isolate (Expo) carried fifteen different genotypes. The

Expo isolate was considered the most diverse isolate collected from Georgia fields and was therefore selected as a candidate for RNA sequence analysis. This study explored the possible sources of resistance against the current *P. pachyrhizi* population in Georgia, provided comprehensive data for ASR genetic diversity and population structure, and selected a Georgia isolate and a resistant soybean cultivar for RNA sequence analysis, which can assist in developing durable resistance to ASR by integrating both the host resistance and pathogen variability data sets.

Introduction

Asian Soybean Rust (ASR) caused by *Phakopsora pachyrhizi* Sydow, has the potential to cause severe economic losses in soybean production. This disease has been widely distributed in the tropics and subtropics of Asia for many decades and was first reported on *Pachyrhizus erosus* (yam bean, jicama) in Japan in 1902 (Hennings 1903). More recently, ASR was reported in Africa (Akinsanmi and Ladipo 2001; Pretorius *et al.*, 2001), Argentina (Yorinori *et al*., 2002), Australia (Bromfield 1984), Bolivia (Yorinori *et al.*, 2005), Brazil (Rossi 2003), China (Kitani *et al.*, 1960; Yeh *et al.*, 1975), Hawaii (Killgore and Heu 1994), and Paraguay (Morel 2001) with associated yield losses ranging from 60% to 100% in some areas costing growers more than $1 billion dollars. November 6, 2004 marked the first observation of ASR in the continental United States (Schneider *et al*., 2005). ASR has not become a major threat to soybean production in North America because it cannot survive the winters in the main soybean production areas in the absence of an actively growing host, and ASR can only overwinter on kudzu in tropical areas in central and south Florida. However, ASR is a concern for the soybean producers in southern states like Georgia, Alabama, Texas, Mississippi, and Lousiana that are geographically close to areas where the pathogen can overwinter. In addition, the long-distance dispersal of urediniospores by wind currents and climate change also has the potential to trigger an outbreak of this disease in the North.

The occurrence of ASR is monitored throughout the USA by the Soybean Rust-Pest Information Platform for Extension and Education (SBR-PIPE) network, which was established by the U.S. Department of Agriculture (USDA) (Hershman *et al*., 2011). With the assistance of this monitoring and predication system, ASR is mainly managed by the application of strobilurin and triazole fungicides when ASR has been observed in sentinel plots. Although this method has

been successful in the control of widespread epidemics of ASR, it increases production costs and the environmental impact of soybean production. Therefore, an integrated pest management (IPM) approach utilizing resistant soybean cultivars is considered as the most effective and economic method to control ASR.

Efforts have been made to identify resistant soybean accessions and to characterize the resistance against ASR. Miles *et al*. (2006) screened more than 16,500 accessions from the USDA soybean germplasm collection that were inoculated with a mixture of four international ASR isolates. A total of 805 accessions with resistance to the ASR isolate mixture were identified. The resistant accessions were then tested in field nurseries in the southeastern USA (Walker *et al*., 2011). Only 85 of the 805 accessions ranged from highly to moderately resistant to the ASR field isolates from the Southeast. The remaining 720 accessions were considered susceptible. This indicated that there is considerable variation between the U.S. and international ASR populations and that this variation can have a significant impact on host resistance. Screening for resistance has also extended to the other legume plants, such as cowpea, green pea, lima bean, snap bean, common bean, and kudzu (Bonde *et al.*, 2008; Jordan *et al*., 2010; Souza *et al*., 2014). The perennial wild relatives of soybean were screened for ASR resistance (Burdon and Marshall 1981; Hartman *et al*., 1992), and demonstrated that 33% of *G. tomentella* and 32% of *G. tabacina* accessions were highly resistant. However, from 40% to 73% of the *Glycine* accessions were susceptible.

Host resistance against ASR is defined by three natural responses of soybean to *P. pachyrhizi* infection: (1) tan sporulating lesions (TAN), (2) red-brown lesions (RB), or (3) no visible lesions (IMMUNE). Susceptible genotypes are characterized by tan-colored lesions due to the production of numerous urediniospores within the uredinia, whereas RB lesions show

limited fungal growth with limited sporulation caused by plant HR (hypersensitive reaction). The absence of lesions indicates that the soybean plant is immune to ASR. At present, none of the commercially available soybean cultivars are resistant to all of the ASR isolates (Hartman *et al*., 2005). Six loci with major dominant genes that confer the resistance against ASR have been described*:* (1) Resistance to *P. pachyhrizi 1* (*Rpp1*) from PI 200492, (2) *Rpp2* from PI 230970, (3) *Rpp3* from PI 462312, (4) *Rpp4* from PI 459025, (5) *Rpp5* from multiple PIs, and (6) *Rpp6* from PI 567102B (McLean and Byth 1980; Hartwig and Bromfield 1983; Hartwig 1986; Garcia *et al*., 2008; Li *et al*., 2012).

Despite the identification of R-genes, race-specific resistance is suggested not to be durable due to the rapid evolution of the pathogen population which overcomes host resistance. To date, most studies have focused on the characterization of host plant resistance and understanding the genetics of resistance against ASR, while studies focusing on the genetic diversity and population structure of ASR are limited. Although a standard set of soybean differentials and the methods of evaluating reaction types in soybean have yet to be established, most studies that assess the variation of ASR were typically analyzed by inoculating different ASR isolates onto various sets of soybean lines (Yamaoka *et al*., 2002; Twizeyimana *et al*., 2008; Pham *et al*., 2009). The term "pathotype" was used to classify pathogen genotypes whose members express the same pattern of host-specific virulence toward individual clones or species within a specified set (Ramstedt *et al*., 2002). For example, three pathotypes were identified among the 72 U.S. ASR isolates (Twizeyimana and Hartman 2012). Genetic diversity analyses were performed based on DNA sequences of the Internal Transcribed Space (ITS) and the ADP Ribosylation Factor (ARF) gene. Zhang *et al*. (2012) and Freire *et al* (2008) characterized the

polymorphisms of ITS sequences from several U.S. and Brazilian ASR isolates, respectively and predicted the potential origin and migration of ASR.

An analysis of pathogen variation with respect to pathogenicity and the correlation of sequence variation with the host resistance phenotype is needed. Understanding the diversity in ASR field populations will provide insight into the variability that can yield challenges to the available soybean germplasm and to generating new virulence profiles. ITS and ARF markers are routinely used to analyze pathogen diversity and are useful across a broad range of potential pathogen species (van der Merwe *et al.,* 2007; van der Merwe *et al.,* 2008; Liu and Hambleton 2010; Hyde *et al.,* 2014). Therefore, objectives of the present study were to (1) characterize the current Georgia ASR field isolates by inoculating them onto a set of soybean differential lines, (2) test the resistance of candidate soybean lines and perennial wild relatives and, (3) study the genetic diversity of the Georgia ASR population.

Materials and Methods

Plant Materials

Seeds for all of the lines utilized for this work were obtained from the USDA Soybean Germplasm Collection (Dr. Randy Nelson, Urbana, IL). The differential lines were selected based on their ASR resistance genotype (Table 2.1). Each differential line carries a specific *Rpp* R-gene against ASR: (1) PI 200492 (*Rpp1*), (2) PI 230970 (*Rpp2*), (3) PI 462312 (*Rpp3*), (4) PI 459025B (*Rpp4*), (5) PI 200526 (*Rpp5*), (6) PI 200487 (*Rpp5*), and (7) PI 200456 (*rpp5*). The soybean lines can be used to phenotypically differentiate between ASR isolates. Two lines (Williams 82 and Cobb) that are susceptible to all known ASR isolates were used as a positive control with all inoculation experiments (Hartwig and Jamison 1975). Additional seeds for the Japanese germplasm and fourteen released soybean cultivars that are grown in Georgia were

obtained from Dr. Roger Boerma and include: (1) AGS 758RR, (2) Pioneer 97M50, (3) Hyuuga-Janapnese cultivar, (4) USG7732nRR, (5) G003213, (6) H7242RR, (7) Prichard RR, (8) Cook, (9) 5601T, (10) Ozark, (11) NC ROY, (12) AGS787, (13) AGS828, (14) Woodruff, and (15) 97M50. Eight accessions from five soybean wild relatives were utilized to test their resistance to the current ASR population in Georgia: (1) *G. tomentella* (PI 446995, PI 446988, PI 483218), (2) *G. tabacina* (PI 483199, PI 339661), (3) *G. argyrea* (PI 505151), (4) *G. microphylla* (PI 373983A), and (5) *G. pescadrensis* (PI 440996). All of the plants were maintained in a growth chamber with day and night environments of 28/20°C temperature and 14/10 h of photoperiod, respectively.

Asian Soybean Rust Isolates

ASR field isolates were collected from different locations in Georgia in 2012 and 2013. Several isolates were provided by Dr. Robert Kemerait. One isolate designated Griffin was collect from Griffin, GA. in June 2012. Four isolates were collected from Attapulgus, GA. in August 2012 and were designated APK, APS1, APS2-3 and APS4. APK was collected from kudzu and the APS isolates were each collected from three different soybean fields in Attapulgus. One isolate, designated Thomas, was collected in Thomas County in October 2012. In September 2013 two isolates were collected in Moultrie and Tifton and were designated Expo and TVP, respectively. Initially, the eight field isolates were considered different because they were isolated from different fields. Therefore, the ASR field isolates were maintained separately on two known susceptible soybean cultivars (Williams 82 or Cobb) and likely represent a mixture of different ASR races.

Plant Inoculations

Two universally susceptible soybean cultivars (Williams 82 or Cobb) were used to increase rust spores collected from the ASR field isolates. Seeds from both susceptible cultivars were planted in small pots and placed in a growth chamber. The growth chamber was maintained with day and night environments of 28/30°C temperature and a 12/12 hour of photoperiod, respectively, and approximately 500 µmol m^{-2} sec^{-1} photosynthetically active radiations at the top of the canopy. During the day and night the relative humidity was maintained at 70% and 90%, respectively. Twenty-one days after planting, the susceptible soybean lines were inoculated separately with the eight field isolates and maintained separately. A small amount of spores from the eight field isolates were first inoculated separately on to the two known susceptible soybean cultivars to increase the number of spores.

To prepare the inoculum, urediniospores collected from the field were suspended in sterile distilled water containing 0.01% (v/v) Tween 20 (Sigma-Aldrich, St Louis, MO) and filtered with 100 µm cell strainer (Corning Life Sciences, Tewksbury, MA) to remove plant debris and clumps of spores. Spore concentration was counted using a hemocytometer and adjusted as needed to a final solution of 6×10^4 spores/ml. Using a spray bottle, 10ml of the spore suspension was sprayed onto the surface of each leaflet. After inoculation, the plants were placed in a black plastic bag that was used as a humidity chamber with a relative humidity close to 100% to promote infection. At 16 hour post inoculation (hpi), the plants were removed from the bags and placed in a growth chamber. Fourteen days post inoculation (dpi) the leaves of ASR-infected susceptible soybean plants with high sporulation were collected and placed into a 250ml flask. Urediniospores were washed away from the leaf tissue with 0.01% Tween 20, filtered, and adjusted to the appropriate concentration as described previously. The resulting spore suspension

was used to inoculate the eight selected soybean differentials, fifteen candidate soybean lines (one Japanese cultivar and fourteen Georgia adapted cultivars) and eight wild soybean relatives. The soybean seedlings were inoculated 21 days after planting while the soybean wild relatives were inoculated 8 weeks after planting due to slow germination.

Phenotypic Scoring

The soybean seedlings and relatives were phenotypically scored 14 dpi (days post inoculation) when symptoms were clearly seen on the two susceptible cultivars to monitor the disease progress and to characterize the resistance of the seedlings to the ASR field isolates. Experimental units were single seedlings with 6 seeds for each soybean plant line/wild relative. A standard disease rating system was used to score the plants. Seedlings with reddish brown lesions (RB) and some sporulation were considered resistant, while seedlings with tan lesions (TAN) and heavy sporulation were considered susceptible (Fig. 2.1). Seedlings with no visible symptoms were characterized an immune (IMMUNE). Three independent experiments were performed for each field isolate and soybean line or wild relative. The phenotypic data was used to determine how many unique ASR field isolates were present among the field isolates collected from eight different locations. ASR field isolates were considered to be different when the phenotype pattern of the soybean differentials were different when inoculated with each ASR field isolates. Conversely, when the phenotypic pattern was identical, the ASR field isolates were characterized as the same isolate.

Rust Genomic DNA extraction, PCR Amplification, Cloning and Sequencing

ASR genomic DNA was extracted from each field isolate by using ZR Fungal/Bacterial DNA MiniPrep kit as described by the manufacturer (Zymo Research, Irvine, CA). The internal transcribed space (ITS) and the ADP Ribosylation Factor (ARF) gene were PCR amplified from

the genomic DNA of each field isolate with one ITS primer pair (Forward, ITS4-TCCTCCGCTTATTGATATGCTT; Reverse, ITS5- GGAAGTAAAAGTCGTAACAAGG) and one ARF primer pair (Forward, ARF- CACCAACAACTTCAGTAGACAGTTT; Reverse, ARF-GAGCCATTCCAGGCCCTCGTACAA) as described in Zhang *et al*. (2012). The ITS and ARF regions are typically used in PCR-based approaches to differentiate between rust fungal isolates (van der Merwe *et al*., 2007; van der Merwe *et al*., 2008; Zhang *et al*., 2012; Jorge *et al*., 2015). PCR was performed in 50-mL reactions with 10 mM Tris-HCl, pH 8.8; 50 mM KCl; 1.522.5 mM MgCl2; 0.8 mM total deoxyribonucleotide triphosphates; 4U of High Fidelity Taq DNA polymerase; 25 pmol of a forward and reverse primer; and 50 ng of genomic DNA. An annealing temperature gradient was initially used that ranged from 50°C to 56°C to optimize the annealing temperature for different primer pair combinations. The resulting PCR products for the ITS (665 bp) and ARF (1065 bp) region were isolated from a 1.5% agarose gel, purified with an Invitrogen Quick gel extraction kit (Carlsbad, CA), and cloned into the Invitrogen pCR2.1-TOPO cloning vector using the methods described by the manufacturer. Thirty-six independent clones were selected randomly for each isolate (ITS and ARF) and sequenced using the Big-Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Grand Island, NY) following the manufacturer's protocol. This was performed for each isolate (8 isolates; 36 ITS clones; 36 ARF clones= 576 total clones sequenced). Sequences of the ITS and ARF region derived from the eight field isolates were used for diversity analysis studies. All sequencing was performed at the Georgia Genetic Facility.

Sequence Data Analysis

Sequence of the 576 ITS and ARF clones were used to run a BLASTN search against the NCBI non-redundant database to verify putative homologies to known ITS and ARF genes (Hall

1999). Sequences with homologies to known ITS and ARF genes were aligned using ClustalW (Larkin *et al*., 2007). The duplicated genes were removed and all the remaining ITS and ARF sequences were considered as different genotypes present in each field isolates. ITS genotypes were selected to study the population structure of ASR in Georgia. The number of the ITS sequences identical or similar with known ITS genes were recorded in Table 2.4. Alignment of the ITS genotypes and ARF genotypes for each field isolate were subjected to genetic diversity analysis using DnaSP version 5, where the number of indels, single nucleotide polymorphisms (SNPs) as well as nucleotide diversity were analyzed (Librado and Rozas 2009). Two phylogenetic trees were constructed with all of the putative ITS and ARF genotypes using the MEGA (Molecular Evolutionary Genetics Analysis) program version 5 (Tamura *et al*. 2011) and the Neighbor-Joining (NJ) method with distance represented as the number of nucleotide differences. Confidence in the phylogeny was assessed with one-thousand bootstrap replicates.

Results

Georgia ASR Isolate Pathotypes

Phenotypic analysis of the soybean differentials inoculated with the ASR field isolates indicated that the isolates collected from eight different locations in Georgia represent three different ASR field isolates. A total of about 1,200 seedlings were scored. Three distinct phenotypic patterns were observed (Table 2.1). When inoculated with the Griffin, APK, APS1 and APS2-3 field isolates, an identical phenotypic pattern was observed for the seven soybean lines and two susceptible lines indicating that the four field isolates are the same isolate. Five soybean lines (PI 200492- *Rpp1*, PI 230970- *Rpp2*, PI 462312- *Rpp3*, PI 459025B- *Rpp4*, PI 200487- *Rpp5*) responded with a red brown (RB) phenotype (resistant) indicating the presence of a R-gene in the plant that is effective against the corresponding avirulence genes/effectors in

the field isolates collected from the four locations. However, two soybean accessions (PI 200526- *Rpp5* and PI 200456- *rpp5*) and two universally susceptible lines (Williams 82 and Cobb) responded with a TAN (susceptible) phenotype. This indicated that the two soybean lines do not carry an effective R-gene and the field isolates from the four locations carry virulence genes/effectors that are undetected by the host plant, allowing the four field isolates to be pathogenic on the two soybean lines.

An identical phenotypic pattern was observed when inoculated with APS4 and Thomas isolates (Table 2.1). Additionally, the APS4 and Thomas phenotypic pattern was similar to the Griffin, APK, APS1, and APS2-3 pattern, differing only by an immune (IM) response observed for the soybean line carrying the *Rpp1* gene (PI 200492). This indicated that APS4 and Thomas are the same isolate and are different from Griffin, APK, APS1, and APS2-3. The phenotypic pattern observed for Expo and TVP was identical indicating that they are the same isolate. The Expo and TVP phenotypes were most similar to APS4 and Thomas differing only by a TAN response observed for the soybean line carrying the *Rpp4* gene (PI 459025B). This indicated that PI 459025B does not carry an R-gene that is effective against the virulence genes/effectors in the Expo and TVP isolates collected from those locations, resulting in susceptibility. Additionally, the phenotypic pattern observed for Expo and TVP was also different from the Griffin, APK, APS1 and APS2-3 pattern. This was demonstrated by an IM and TAN response for PI 200492 (*Rpp1*) and PI 459025B (*Rpp4*), respectively, for Expo and TVP. This suggested that *Rpp4* is ineffective against virulence genes/effectors in Expo and TVP.

When inoculated with each isolate, five soybean lines carrying *Rpp1*, *Rpp2*, *Rpp3*, *Rpp4* or one of the *Rpp5* resistance genes consistently responded with an RB or IM phenotype (Table 2.1). This indicated that the five soybean lines have an R-gene that is effective against the field

isolates collected from all locations carry the corresponding avirulence genes/effectors, resulting in resistance. However, PI 200526 carrying a second dominate R-gene and PI 200456 carrying the recessive *rpp5* gene, all responded with a TAN phenotype with all of the field isolates indicating the ineffectiveness of these alleles at the *Rpp5* locus against the virulence genes/effectors in the current ASR population in Georgia.

Characterization of Host Resistance of Candidate Soybean Lines and Wild Relatives

Phenotypic analysis of fifteen candidate soybean lines and eight soybean wild relatives was conducted to assess the resistance response when inoculated with the eight characterized ASR field isolates collected from different locations in Georgia (Figure 2.1). A total of about 2000 seedlings were scored. All of the tested candidate soybean lines responded with a TAN resistance reaction to all of the ASR field isolates. This indicated that the released soybean lines do not carry an effective R-gene against any of the ASR field isolates, resulting in susceptibility. Conversely, Hyuuga showed a RB reaction when inoculated with each ASR field isolate indicating that Hyuuga has R-genes that are effective against the field isolates collected from all four locations in Georgia. The resistance response of eight selected soybean wild relatives was also assessed following inoculation with each of the ASR field isolates (Table 2.3). A total of about 1000 seedlings were scored. All of the soybean wild relatives conferred a high level of resistance to each of the ASR field isolates demonstrated by a RB or an IM response. This indicated that the soybean wild relatives carry R-genes that are effective against the corresponding avirulence genes/effectors in the field isolates collected from all four locations in Georgia and are therefore a good source of resistance to ASR. Interestingly, there were several out-breaks of powdery mildew in the greenhouse when the soybean plants were screened for resistance. Therefore, the resistance of the soybean lines and wild relatives to powdery mildew

was assessed (Table 2.2 and Table 2.3). Ten released Georgia soybean lines (AGS 758RR, Pioneer 97M50, G003213, H7242RR, Prichard RR, 5601T, AGS787, AGS828 and 97M50) and two wild relatives (PI 483199 and PI 446995) conferred resistance to powdery mildew indicated by the absences of disease development.

Genetic diversity of Georgia ASR field isolates

A total of 115 ITS genotypes were used to study the genetic variation in each Georgia field isolate (Table 2.4). This number excludes genes that were cloned and were not identified as ITS genes. Based on their sequences, all of the ITS genotypes corresponded to 12 reference genotypes from different countries and from states surrounding Georgia (Brazil-1, MUT Zimbabwe, OK07-1A, FL08-2B, MS08-2D, India 73-1, MS07-1E, TH02-1A, BZ01-1C, AR07-2A, MO06-1D and AU79-1/HW94-1B). Eight new Georgia genotypes were identified that were similar in sequence to the reference genotypes (Table 2.4). For example, two Georgia genotypes (GA1 and GA2) were identical to genotypes from a Brazilian isolate (Brazil-1) and were therefore designated GA1-Brazil and GA2-Brazil. Similarly, two Georgia genotypes were identical to genotypes from an isolate collected in Zimbabwe (MUT Zimbabwe) and were designated GA1-Zimbabwe and GA2-Zimbabwe. One Georgia genotype was identical to a genotype from a Florida isolate (FL08-2B) and was designated GA1-FL08-2B.

Each field isolate is s a mixture of different genotypes and carries from 9 to 15 different genotypes (Table 2.4). The Expo field isolate from Georgia has the most number of genotypes in its population that were identical to the reference genotypes, indicating it is the most diverse field isolate collected. Two genotypes (Brazil-1 and Mo06-1D) were identified in all of the Georgia ASR field isolates at all eight locations and have most number of genotypes (4 to 15)

identical to the genotypes in the Georgia isolates. This indicated that Brazil-1 and Mo06-1D are the two predominate genotypes in current soybean rust population in Georgia.

In addition to the ITS sequences, a total of 110 ARF sequences were selected from 288 clones (Table 2.5). The ARF genotypes in the current ASR population in Georgia corresponded to two major genotypes, GA08-1A and GA08-1D. All of the isolates have 14 to 25 ITS sequences and from 12 to 24 ARF sequences. After removal of duplicated sequences, 20 ITS and 17 ARF genotypes were identified (Table 2.5). In the ITS and ARF region, 12 and 18 polymorphisms were identified in the eight isolates, respectively. A total of 3 SNPs were identified in the ITS sequences and 17 were present in the ARF sequences. A total of 9 indels were identified in the ITS sequences, while only 1 indel was found in the ARF sequences. Indels were the major type of polymorphism identified in the ITS and ARF sequences. The average genetic diversity of the ITS sequences was detected at 0.271%, while the average genetic diversity of the ARF sequences was 0.204%. The average number of nucleotide differences is very low in both ITS and ARF genes and ranged from 0-2.73.Two groups were identified in the ITS sequences. One group has an insertion (TTTA) and a SNP (G nucleotide) close to the insertion, while the other has a different SNP (T or A nucleotide) at the same position. The two ARF groups can be distinguished by the presence of an extra 6 base pair repeat (CCCTCT).

According to the BLAST search, the two groups of ITS sequences were either identical or highly similar (98%-100%) to the two major ITS genotypes in GenBank. One major ITS genotype was Brazil-1 (GenBank accession number EU523736) (Silva *et al*., 2008) from the Brazilian isolates collected in 2001. The ITS sequences are also highly similar to the MUT Zimbabwe (GenBank accession number AF333499) genotype from the isolate collected from Mutare, Zimbabwe in 2000. The other major genotype was AU79-1B (GenBank accession

number JN872987) from the isolate collected from Australia in 1979. This isolate is also highly similar to the HW94-1B genotype (GenBank accession number JN872977) from the isolate collected from Hawaii in 1994, as well as the India 73-1 genotype (GenBank accession number AF333492) from the isolate collected from Pantnagar, India in 1973 (Frederick *et al*. 2002). In terms of ARF sequences, the two groups were either identical or highly similar (98%-100%) to the GA08-1A (GenBank accession number JN899178) and GA08-1D (GenBank accession number JN899203) genotypes from the isolates collected from Georgia in 2008 (Zhang *et al*., 2012).

Phylogenetic analysis of the aligned ITS genotypes shows that two major clades (I and II) were present, suggesting there are at least two major groups of genotypes within ASR field isolates in Georgia (Figure 2.2). Reference genotypes Brazil-1 and MUT Zimbabwe were on the same major clade with the ITS-I genotype, while India 73-1, AU79-1B and HW94-1 were on the major clade with ITS-II. However, phylogenetic analysis indicated that the ARF region is more conserved than the ITS region. The majority of the genotypes were clustered on one clade with only 2 genotypes in the second clade (Figure 2.2). The low bootstrap value for both phylogenetic trees may be due to the high similarity among the sequences analyzed in this study and that the sequences only differ by a few SNPs or indels.

Discussion

To look for more innovative methods to control ASR, a better understanding of the pathogenic variability and population structure is essential to dissect the evolutionary relationship between the host and pathogen. In this study, we used both phenotyping and genotyping techniques to characterize the current ASR field isolates in Georgia, and furthermore identified soybean and wild relative lines with resistance against this disease.

Our collection of ASR isolates was selected from 8 different Georgia fields in 2012 and 2013. Three pathotypes were identified from this collection based on the phenotypic data. The three pathotypes were compared with the pathotypes that Twizeyimana and Hartman (2012) characterized from 72 U.S. ASR isolates based on the virulence. The three pathotypes in this study can be considered as two pathotypes using their criteria. The Georgia pathotypes from this study were identical to pathotype 041 and 011, while pathotype 241, which is more aggressive than all other pathotypes were not identified in Georgia field isolates The difference between pathotype 041 and 011 can be observed with PI 459025B (*Rpp4*) which is resistant to 011 while susceptible to 041. In this study, TVP and Expo were collected in southeast Georgia in 2013 and are considered to be pathotype 011 while PI 459025B is resistant to all other isolates collected in 2012. This indicated that the *Rpp4* gene may be overcome in the Georgia soybean field. To confirm this finding, more field isolates need to be tested. This also suggests that the pathotypes in the Georgia ASR field isolates have been well maintained over time. This may be due to the lack of selection pressure on the pathogen. For example, in this study none of the release Georgia soybean lines were resistant to the ASR field isolated. Therefore, they do not carry an R-gene that is effective against the current ASR pathogen population. In the absence of an effective R-gene in soybean cultivars, there is no selection pressure imposed on the ASR pathogen population to change. This often results in the maintenance of very similar pathotypes in the pathogen population over time. The Expo isolate was considered to be the most diverse Georgia field isolate, therefore was selected for RNA sequence analysis in Chapter 3.

Hyuuga is a Japanese cultivar. This cultivar conferred resistance to all eight Georgia field isolates in this study. The resistance of this collection has been studied because of its resistance to many ASR isolates (Monteros *et al*., 2007; Monteros *et al*., 2010; Kendrick *et al*., 2011).

Hyuuga resistance was originally thought to be controlled by *Rpp?*, a single dominate R-gene that mapped to the same region as *Rpp3*. A second R-gene against ASR was identified recently in Hyuuga, which explains why the resistance of Hyuuga is still effective, while the *Rpp3* resistance was broken down in Brazil. The resistance of Hyuuga was also tested with multiple domestic and international isolates collected over an extended period of time. Only a very few isolates resulted in susceptibility (Kendrick *et al*., 2011; Walker *et al*., 2011; Kim *et al*., 2012), indicating that the Hyuuga resistance has been effective over an extend period of time. As a result, we selected Hyuuga as a candidate cultivar for studying durability of resistance against ASR utilizing RNA sequence analysis (Chapter 3).

The perennial wild species of *Glycine* have been reported to have sources of ASR resistance (Burdon and Marshall 1981; Hartman *et al*., 1992). For example, both accessions of *G. canescens* and *G. argyrea* were reported to carry single dominant R-genes (Burdon 1988; Jarosz and Burdon 1990). In this study, we selected eight soybean wild relatives that demonstrated resistance to ASR isolates collected about 20 years ago. The resistance of the soybean wild relatives against the current Georgia field isolates were confirmed in this study, indicating that their resistance has been well maintained, even though in previous studies the population of ASR was shown to be evolving rapidly based on pathogenic variation and population genetic analyses (Freire *et al*., 2008; Twizeyimana and Hartman 2012; Zhang *et al*., 2012). Although the R-genes have not been transferred to cultivated soybean lines due to the low frequency of successful hybridization and early pod abortion, there have been some progresses in crossing *G. tomentella* with soybean following the modified procedure of embryo rescue, colchicine treatments, and backcrossing (Singh and Hymowitz 1987; Singh 2010). By refining the hybridization procedures between wild relatives and cultivated soybean or cloning target R-genes from wild relative and

transferring them into soybean lines through transgenic techniques, soybean wild relatives may offer additional R-genes that may be stacked with *Rpp* genes and potentially provide more durable resistance against current ASR. It is not clear if the resistance of the wild relatives is controlled by a single dominate resistant gene or by multiple genes. Either way, it is likely that the genes in the wild relatives are still effective against most ASR isolates because the R-genes are not utilized in agriculture, therefore have not imposed selection pressure on the pathogen to change, often resulting in the breakdown of resistance. This suggests the soybean wild relatives may be resource for selecting and cloning effective R-genes against ASR. Although crosses of wild relatives with cultivated soybean may not be plausible, transgenic approaches can be utilized for pyramiding R-genes into cultivated soybean (Jones *et al*., 2014).

Population genetic analyses of the eight Georgia field isolates suggested that each isolate is a mixture of different genotypes. The slight difference in phenotype observed for the isolates may be associated with the different combinations of genotypes. This suggests that each genotype may carry different pathogenic effectors that can be recognized by the corresponding soybean resistance proteins that trigger plant defense, resulting in different resistance reactions. The variation of resistance reactions was demonstrated by the phenotypes that ranged from TAN to IMMUNE. This was observed for the soybean differential lines and the soybean wild relatives.

A high level of genetic diversity has been found in the field isolates collected from Georgia in 2012 and 2013. This is consistent with what have been reported for Brazilian ASR isolates (Freire *et al*., 2008). ASR is considered to increase and amplify the population mainly through the asexual stage by producing urediniospores (the sexual stage has not yet been reported). Therefore SNPs and genetic deletions may be the major force to drive the diversification of the genotypes, which was confirmed by this study with the analysis of the ITS

and ARF gene diversity. The similar high diversity for asexual plant pathogens was also reported for *Puccinia triticina* (Goyeau *et al*., 2007). The mechanism causing this high level of genetic diversity was previously explained as due to parasexuality and heterokaryosis (Kolmer 2001). In addition, the rapid genetic diversification may be also accelerated by somatic hybridization causing the fusion of hyphae known as anastomosis, which results in the migration of nuclei in the germ tube network (Vittal *et al*., 2011). The different genotypes present among the pathogen population provides a large reservoir of genetic traits for the pathogen to select so that any superior traits such as overcoming resistance or fungicide resistance can be accumulated rapidly in the population, which explains why several *Rpp* genes were broken down relatively fast in Brazil and there is a drastic reduction of available soybean germplasm resistant to ASR over the past 10 years (Walker *et al*., 2011). In this study, *Rpp4* resistance was apparently broken down by two field isolates Expo and TVP which were collected in 2013. This suggested that the deployment of race-specific resistance control by single dominant genes is very risky and has the potential to alter the population structure of the pathogen population. In addition, fungicide so far has been heavily applied to control ASR. The combination of stroblurin and Triazol has been very effectively in the USA and Brazil. However, several fungicides which were labeled to control ASR were less effective and resistance to fungicide has been reported among some ASR populations (Kemerait, personal communication). Therefore, having a better understanding of the population structure will also help us to adjust the use of chemicals against pathogen.

ASR was first reported in Japan. It was proposed that ASR migrated from East Asia to Africa and Australia and then moved from Africa to South America, eventually arriving in the continental United States. This theory was also confirmed by this study by the identification of a major group of genotypes that were identical or highly similar to one genotype from Brazil-1 that

was collected from Brazil in 2001. Interestingly, some genotypes were also very similar to Australian and Hawaiian genotypes, indicating that there might be an alternative migration route for ASR from Australia and Hawaii to the continental USA between 1996 and 2000. The exact location of where ASR first arrived in the continental USA is unknown. However, based on this study and previous data, the ASR isolates may be a mixture of ASR isolates from South America and arrived in Central America sometime before 2006. However, to examine this question further, it would require a larger collection of both domestic and international isolates that were collected over an extended period of time.

The study of ASR both regionally and internationally can provide a better understanding of the evolution and migration of this pathogen. Freire *et al* (2008) demonstrated that there are multiple times of entries of ASR from Africa to the USA and that the ones that arrived earlier may be distributed in a larger geographic region. This suggested that new ASR isolates can migrate into the USA during severe weather and increase the genetic diversity of the ASR pathogen population, potentially causing economic loss in the soybean production in North America and make it even more critical to develop durable resistance to ASR.

Literature cited

1. Akinsanmi O. A., Ladipo J. L. (2001) First report of soybean (*Phakopsora pachyrhizi*) rust in Nigeria. Plant Disease 85:97.

2. Arias C. A. A., Toledo J. F. F., Almeida L. A. *et al* (2008) Asian rust in Brazil: varietal resistance. In: Kudo H. *et al* (eds) JIRCAS working rep no. 58: facing the challenge of soybean rust in South America. Japan International Research Center for Agricultural Sciences (JIRCAS), Tsukuba, pp 29–30.

3. Bonde M. R., Nester S. E., Berner D. K., Frederick R. D., Moore W. F., Little S. (2008) Comparative susceptibilities of legume species to infection by *Phakopsora pachyrhizi*. Plant Disease 92: 30-36.

4. Burdon J. J. (1988) Major gene resistance to *Phakopsora pachyrhizi* in *Glycine canescens*, a wild relative of soybean. Theoretical and Applied Genetics 75:923-928.

5. Burdon J. J. and Marshall D. R. (1981) Evaluation of Australian native species of *Glycine* for resistance to soybean rust. Plant Disease 65:44-45.

6. Frederick R. D., Snyder C. L., Peterson G. L., Bonde M. R. (2002) Polymerase chain reaction assays for the detection and discrimination of the soybean rust pathogens *Phakopsora pachyrhizi* and *P. meibomiae*. Phytopathology 92:217-227.

7. Freire M. C. M., Oliveira L. O., Almeida A. M. R., Schuster I., Moreira M. A., Liebenberg M. M. and Mienie C. M. S. (2008) Evolutionary history of *Phakopsora pachyrhizi* (the Asian soybean rust) in Brazil based on nucleotide sequences of the internal transcribed spacer region of the nuclear ribosomal DNA. Genetics and Molecular Biology 31:920-931.

8. Garcia A., Sanches Calvo E., Souza Kiihl R., Harada A. (2008) Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: discovery of a novel locus and alleles. Theoretical and Applied Genetics 117:545-553.

9. Goyeau H., Halkett F., Zapater M. F., Carlier J. and Lannou C. (2007) Clonality and host selection in wheat pathogenic fungus *Puccinia triticina*. Fungal Genetics and Biology 44:474-483.

10. Hall A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95-98.

11. Hartman G. L., Miles M. R. and Frederick R. D. (2005) Breeding for resistance to soybean rust. Plant Disease 89:664-666.

12. Hartman G. L., Wang T. C., and Hymowitz T. (1992) Sources of resistance to soybean rust in perennial *Glycine* species. Plant Disease 76:396-399.

13. Hartwig E. E. (1986) Identification of a fourth major gene conferring resistance to soybean rust. Crop Science 26:1135-1136.

14. Hartwig E. E. and Bromfield K. R. (1983) Relationships among three genes conferring specific resistance to soybean rust. Crop Science 23:237-239.

15. Hartwig E. E. and Jamison K. W. (1975) The uniform soybean tests-Southern states. USDA-ARS, Stoneville, MS.

16. Hennings V. P. (1903) A few new Japanese Uredinaceae. Hedwigia 42:S107-108.

17. Hershman D. E., Sikora E. J. and Giesler L. J. (2011) Soybean rust PIPE: Past, present, and future. Online. Journal of Integrated Pest Management. doi:http://dx.doi.org/10.1603/IPM11001

18. Hyde K. D., Nilsson R. H., Alias S. A., Ariyawansa H. A., Blair J. E., Cai L., & Zhou N. (2014) One stop shop: backbones trees for important phytopathogenic genera: I (2014). Fungal Diversity 67(1):21-125.

19. Jarosz A. M., and Burdon J. J. (1990) Predominance of a single major gene for resistance to *Phakopsora pachyrhizi* in a population *of Glycine argyrea*. Heredity 64:347-353.

20. Jones J. D., Witek K., Verweij W., Jupe F., Cooke D., Dorling S., ... & Foster S. (2014) Elevating crop disease resistance with cloned genes. Philosophical Transactions of the Royal Society B: Biological Sciences 369(1639): 20130087.

21. Jordan S. A., Mailhot D. J., Gevens A. J., Marois J. J., Wright D. L., Harmon C. L. and Harmon P. F. (2010) Characterization of kudzu (*Pueraria* spp.) resistance to *Phakopsora pachyrhizi*, the causal agent of soybean rust. Phytopathology 100:941-948.

22. Jorge V. R., Silva M. R., Guillin E. A., Freire M. C. M., Schuster I., Almeida A. M. R., Oliveira L. O. (2015) The origin and genetic diversity of the causal agent of Asian soybean rust, *Phakopsora pachyrhizi*, in South America. Plant Pathology 64(3): 729-737.

23. Kendrick M. D., Harris D. K., Ha B-K., Hyten D. L., Cregan P. B., Frederick R. D., Boerma H. R., Pedley K. F. (2011) Identification of a second Asian soybean rust resistance gene in Hyuuga soybean. Phytopathology 101:535-543.

24. Killgore E. and Heu R. (1994) First report of soybean rust in Hawaii. Plant Disease 78:1216.

25. Kim K. S., Unfried J. R., Hyten D. L., Frederick R. D., Hartman G. L., Nelson R. L., Song Q., Diers B. W. (2012) Molecular mapping of soybean rust resistance in soybean accession PI 561356 and SNP haplotype analysis of the *Rpp1* region in diverse germplasm. Theoretical and Applied Genetics 125:1339-1352.

26. Kitani X. and Inoue Y. (1960) Studies on the soybean rust and its control measure. Part 1. Studies on the soybean rust. Bulletin of the Shikoku Agricultural Experiment Station (Zentsuji, Japan) 5:319-342.

27. Kolmer J. A. (2001) Molecular polymorphism and virulence phenotypes of the wheat leaf rust fungus *Puccinia triticina* in Canada. Canadian Journal of Botany 79:917-926.

28. Larkin M. A., Blackshields G., Brown N. P., Chenna R., McGettigan P. A., McWilliam H., ... & Higgins D. G. (2007) Clustal W and Clustal X version 2.0. Bioinformatics, 23(21): 2947-2948.

29. Li S., Smith J. R., Ray J. D., Frederick R. D. (2012) Identification of a new soybean rust resistance gene in PI 567102B. Theoretical and Applied Genetics 125:133-142.

30. Librado P. and Rozas J. (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.

31. Liu M. and Hambleton S. (2010) Taxonomic study of stripe rust, *Puccinia striiformis sensu lato*, based on molecular and morphological evidence. Fungal biology, 114(10): 881-899.

32. McLean R. J., and Byth D. E. (1980) Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybeans. Australian Journal of Agricultural Research 31:951-956.

33. Miles M. R., Frederick R. D., and Hartman G. L. (2006) Evaluation of soybean germplasm for resistance to *Phakopsora pachyrhizi*. Plant Health Progress. 10.1094/PHP-2006-0104-01-RS.

34. Monteros M. J., Ha B. K., Phillips D. V., and Boerma H. R. (2010) SNP assay to detect the 'Hyuuga' red-brown lesion resistance gene for Asian soybean rust. Theoretical and Applied Genetics 121:1023-1032.

35. Monteros M. J., Missaoui A. M., Phillips D. V., Walker D. R., and Boerma H. R. (2007) Mapping and confirmation of the 'Hyuuga' redbrown lesion resistance gene for Asian soybean rust. Crop Science 47:829-836.

36. Morales A. M. A. P., Borém A., Graham M. A., & Abdelnoor R. V. (2012) Advances on molecular studies of the interaction soybean-Asian rust. Crop Breeding and Applied Biotechnology 12(1):1-7.

37. Morel W. P. (2001) Roya de la soja. Ministerio de Agricultura y Ganaderia, Subsecretaria de Agricultura, Direccion de Investigacion Agricola, Centro Regional de Investigacion Agricola-CRIA, Capitan Miranda, Itapua, Paraguay. Comunicado Tecnico-Report Official, Serie Fitopathologia, 1, Junio de 2001.

38. Nei M. (1987). Molecular evolutionary genetics. Columbia university press.

39. Pham T. A., Miles M. R., Frederick R. D., Hill C. B., and Hartman G. L. (2009) Differential responses of resistant soybean genotypes to ten isolates of *Phakopsora pachyrhizi*. Plant Disease 93:224-228.

40. Pretorius Z. A., Kloppers R. J., Frederick R. D. (2001) First Report of soybean rust in South Africa. Plant Disease 85:1288.

41. Ramstedt M., Hurtado S., and Åström B. (2002) Pathotypes of Melampsora rust on Salix in short-rotation forestry plantations. Plant Pathology 51:185-190.

42. Rossi R. L. (2003) First report of *Phakopsora pachyrhizi*, the causal organism of soybean rust in the province of Misiones, Argentina. Plant Disease 87:102.

43. Schneider R. W., Hollier C. A. , Whitam H. K., Palm M. E., McKemy J. M., Hernandez J. R., Levy L. and DeVries-Paterson R. (2005) First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. Plant Disease 89:774.

44. Schoen D. J., Burdon J. J., and Brown A. H. D. (1992) Resistance of *Glycine tomentella* to soybean leaf rust *Phakopsora pachyrhizi* in relation to ploidy level and geographic distribution. Theoretical and Applied Genetics 83:827-832.

45. Silva D. C. G., Yamanaka N., Brogin R. L., Arias C. A. A., Nepomuceno A. L., Di Mauro A. O., Pereira S. S., Nogueira L. M., Passianotto A. L. L., Abdelnoor R. V. (2008) Molecular mapping of two loci that confer resistance to Asian Rust in soybean. Theoretical and Applied Genetics 117:57-63.

46. Singh R. J. (2010) Methods for Producing Fertile Crosses between Wild and Domestic Soybean Species. United States Patent No. 7,842,850.

47. Singh R. J. and Hymowitz T. (1987) Intersubgenetic crossability in the genus *Glycine* Willd. Plant Breeding 98:171-173.

48. Souza T. L., Dessaune S. N., Moreira M. A., Barros E. G. (2014) Soybean rust resistance sources and inheritance in the common bean (*Phaseolus vulgaris* L.). Genetics and Molecular Research 13(3):5626-5636.

49. Tamura K., Peterson D., Peterson N., Stecher G., Nei M. *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary

distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739.

50. Twizeyimana M. and Hartman G. L. (2012) Pathogenic variation of *Phakopsora pachyrhizi* isolates on soybean in the United States from 2006 to 2009. Plant Disease 96:75-81.

51. Twizeyimana M., Ojiambo P. S., Sonder K., Ikotun T., Hartman G. L., and Bandyopadhyay R. (2009) Pathogenic variation of *Phakopsora pachyrhizi* infecting soybean in Nigeria. Phytopathology 99:353-361.

52. van der Merwe M. M., Walker J., Ericson L., Burdon J. J. (2008) Coevolution with higher taxonomic host groups within the *Puccinia/Uromyces* rust lineage obscured by host jumps. Mycological research 112(12): 1387-1408.

53. van der Merwe M.M., Ericson L., Walker J., Thrall P. H., Burdon J. J. (2007) Evolutionary relationships among species of *Puccinia* and *Uromyces* (Pucciniaceae, Uredinales) inferred from partial protein coding gene phylogenies. Mycological Research 111: 163-175.

54. Vittal R., Yang H. C., Hartmann G. L. (2012) Anastomosis of germ tubes and migration of nuclei in germ tube networks of the soybean rust pathogen, *Phakopsora pachyrhizi*. European Journal of Plant Pathology 132:163-7.

55. Walker D. R., Boerma H. R., Phillips D. V., Schneider R. W., Buckley J. B., Shipe E. R., Mueller J. D., Weaver D. B., Sikora E. J., Moore S. H., Hartman G. L., Miles M. R., Harris D. K., Wright D. L., Marois J. J., and Nelson R. L. (2011) Evaluation of USDA

soybean germplasm accessions for resistance to soybean rust in the southern United States. Crop Science 51:678-693. doi:10.2135/cropsci2010.06.0340

56. Yamaoka Y., Fujiwars Y., Kakishima M., Katsuya K., and Yamada K. (2002) Pathogenic races of *Phakopsora pachyrhizi* on soybean and wild host plants collected in Japan. Journal of General Plant Pathology 68:52-56.

57. Yeh C. C., and Yang C. Y. (1975) Yield loss caused by soybean rust, *Phakopsora pachyrhizi*. Plant Protection Bulletin 17:7-8.

58. Yorinori J. T., Paiva W. M., Frederick R. D. and Fernandez F. T. P. (2002) Ferrugem da soja (*Phakopsora pachyrhizi*) no Brasil e no Paraguai, nas safras 2000/01 e 2001/02, p. 94. In: Resumos da II congreso Brasilerio de soja e mercosoja 2002. Embrapa soja, Londrina, 2002 393 p. (Documantos/Embrapa soja: n. 181. Rusemo 094).

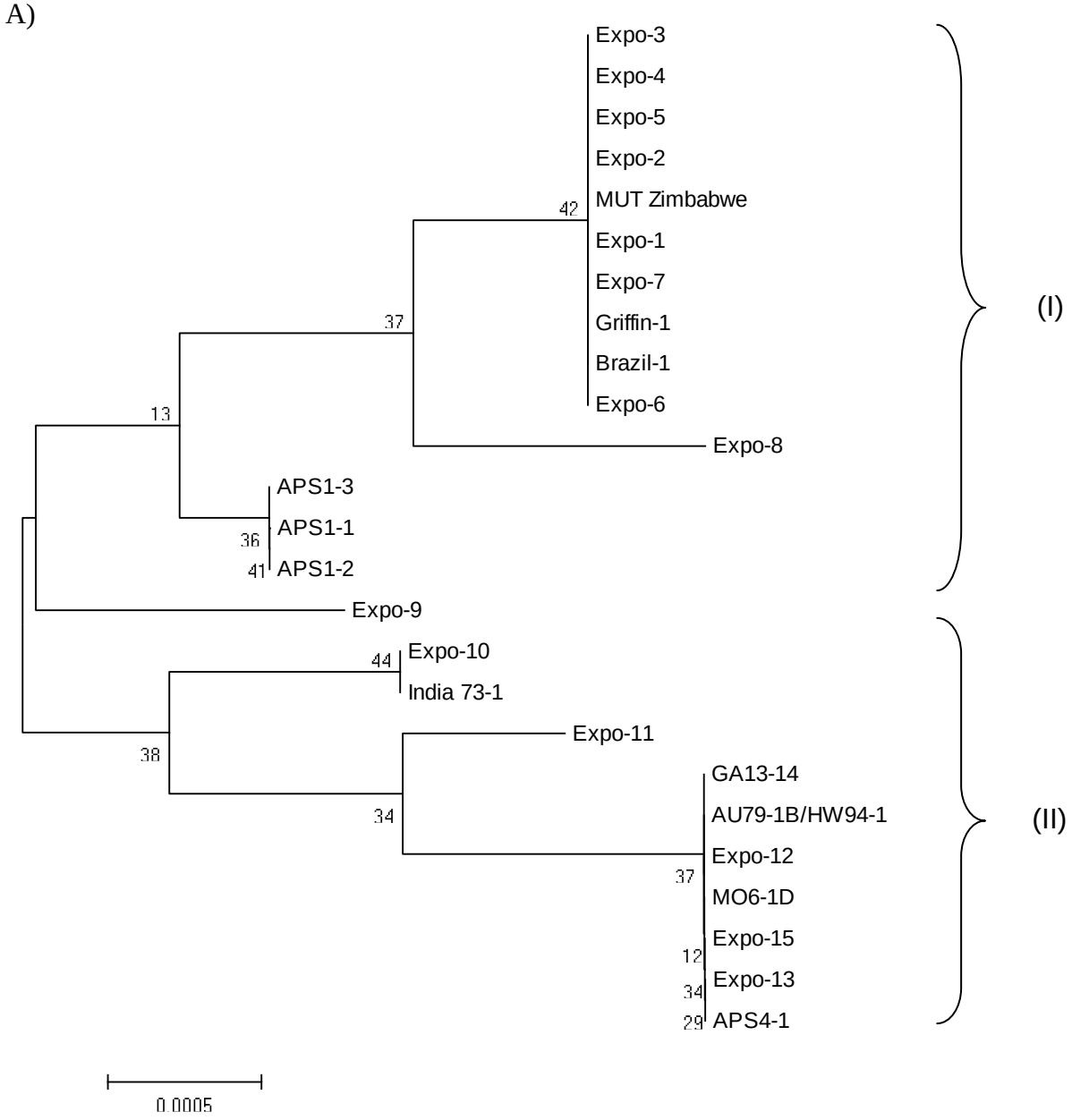
59. Yorinori J. T., Paiva W. M., Frederick R. D., Constamilan L. M., Bertagnoli P. F., Hartman G. L., Godoy C. V. and Nunes J. J. (2005) Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. Plant Disease 89:675-677.

60. Zhang X. C., Freire M. C. M., Le M. H., De Oliveira L. O., Pitkin J. W., Segers G., Concibido V. C., Baley G. J., Hartman G. L., Upchurch G., Pedley K. F. and Stacey G. (2012) Genetic Diversity and Origins of *Phakopsora pachyrhizi* isolates in the United States. Asian Journal of Plant Pathology 6:52-65.

TAN RB IM

Figure 2.1 Disease rating system used to assess the resistance reactions of soybean lines and soybean wild relatives inoculated with ASR. TAN- susceptible reaction; RB- reddish brown, resistant reaction; IM- immune, no symptoms.

A)


Expo-3
Expo-4
Expo-5
Expo-2
MUT Zimbabwe
Expo-1
Expo-7
Griffin-1
Brazil-1
Expo-6
Expo-8
APS1-3
APS1-1
APS1-2
Expo-9
Expo-10
India 73-1
Expo-11
GA13-14
AU79-1B/HW94-1
Expo-12
MO6-1D
Expo-15
Expo-13
APS4-1
42
37
13
36
41
44
38
34
37
12
34
29
0.0005
(I)
(II)

B)

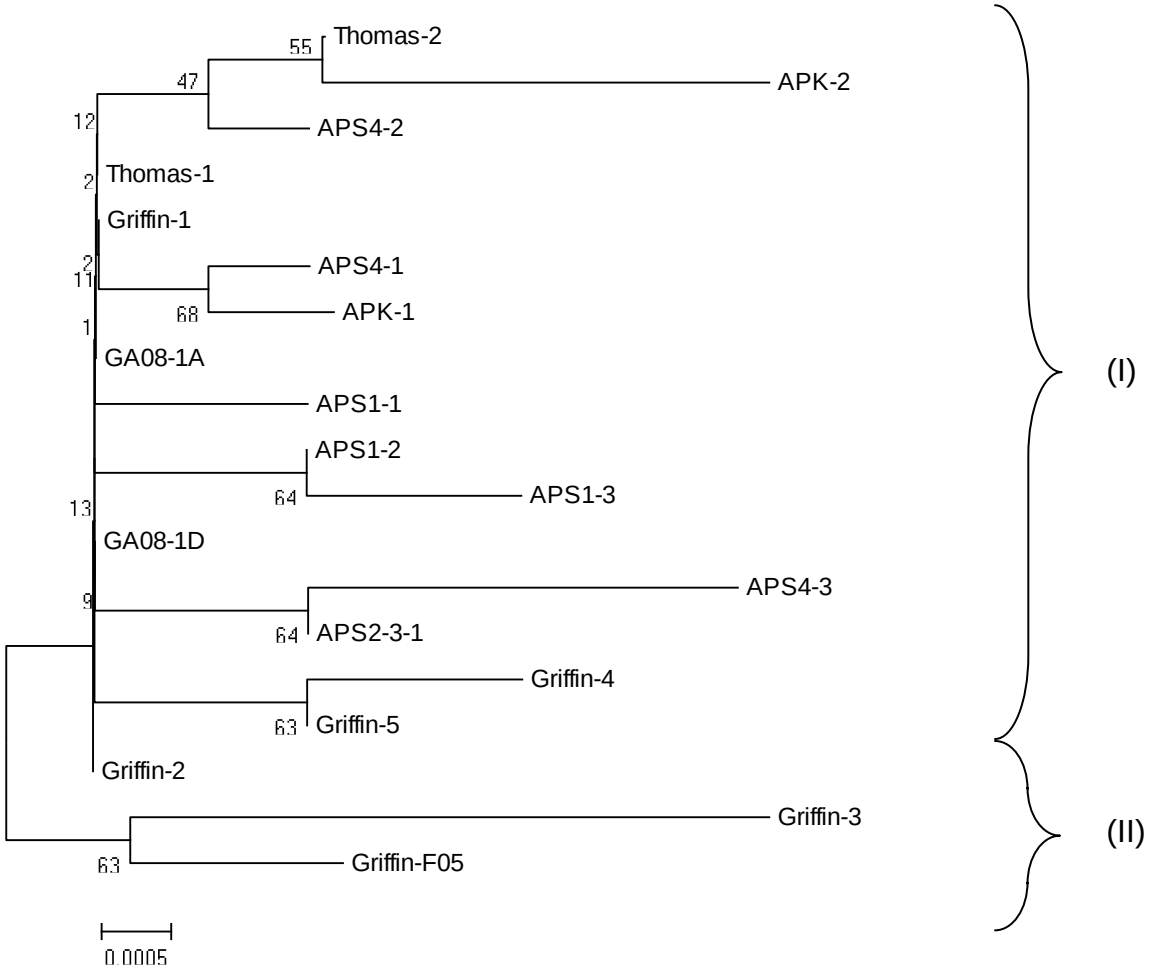


Figure 2.2 Phylogenetic tree of ITS and ARF genotypes present in Georgia ASR field isolates. (A) Phylogenetic tree established by alignment of 20 ITS genotypes from Georgia ASR field isolates. (B) Phylogenetic tree established by alignment of 17 ARF genotypes from Georgia ASR field isolates. The phylogenetic tree was constructed using the Neighbor Joining (NJ) method, with distance represented as the number of nucleotide differences. Confidence in the phylogeny was assessed with one-thousand bootstrap replicates. Clade I and II represent two major genotype groups present in Georgia ASR population.

Table 2.1 Reaction types for nine soybean lines inoculated with eight ASR isolates collected from different locations in Georgia.

Soybean Lines	Griffin	APK	APS1	APS2-3	APS4	Thomas	Expo	TVP
PI 200492 (*Rpp1*)	RB	RB	RB	RB	IM	IM	IM	IM
PI 230970 (*Rpp2*)	RB	RB	RB	RB	RB	RB	RB	RB
PI 462312 (*Rpp3*)	RB	RB	RB	RB	RB	RB	RB	RB
PI 459025B (*Rpp4*)	RB	RB	RB	RB	RB	RB	TAN	TAN
PI 200526 (*Rpp5*)	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
PI 200487 (*Rpp5*)	RB	RB	RB	RB	RB	RB	RB	RB
PI 200456 (*rpp5*)	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Williams 82 (susceptible)	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Cobb (susceptible)	TAN	TAN	TAN	TAN	TAN	-	-	-

TAN- susceptible reaction; RB- reddish brown, resistant reaction; IM- immune, no symptoms.
Source of isolate: Griffin-Griffin, GA.; APK- Attapulgus, GA.; APS1, APS2-3 and, APS4- Three different soybean fields in Attapulgus, GA.; Thomas- Thomas., GA.; Expo- Moultrie, GA.; TVP- Tifton, GA.
Three colors represent three different phenotype patterns and indicate three different isolates.
Dash (-) indicates that the phenotypic data is not available due to unavailability of seeds.

Table 2.2 Reaction types for candidate soybean cultivars inoculated with eight ASR isolates collected from different locations in Georgia.

Cultivars	Griffin	APK	APS1	APS2-3	APS4	Thomas	Expo	TVP
AGS 758RR[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Pioneer	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Hyuuga[b]	RB	RB	RB	RB	RB	RB	RB	RB
USG7732nRR[b]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
G003213[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
H7242RR[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Prichard RR[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Cook[b]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
5601T[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Ozark[b]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
NC ROY[b]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
AGS787[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
AGS828[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
Woodruff[b]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN
97M50[a]	TAN	TAN	TAN	TAN	TAN	TAN	TAN	TAN

[a] Lines resistant to Powdery Mildew; [b] Lines susceptible to Powdery Mildew

TAN- susceptible reaction; RB- reddish brown, resistant reaction.

Source of isolate: Griffin-Griffin, GA.; APK- Attapulgus, GA.; APS1, APS2-3 and, APS4- Three different soybean fields in Attapulgus, GA.; Thomas- Thomas., GA.; Expo- Moultrie, GA.; TVP- Tifton, GA.

Three colors indicate three different isolates.

Table 2.3 Reaction types for eight soybean wild relatives inoculated with eight ASR isolates collected from different locations in Georgia.

PI Number	Species	Griffin	APK	APS1	APS2-3	APS4	Thomas	Expo	TVP
PI 440996[b]	*G. pescadrensis*	RB	RB	RB	RB	RB	RB	RB	RB
PI 483199[a]	*G. tabacina*	RB	RB	RB	RB	RB	RB	RB	RB
PI 339661[b]	*G. tabacina*	RB	RB	RB	RB	RB	RB	RB	RB
PI 446995[a]	*G. tomentella*	RB	RB	N/A	N/A	RB	RB	RB	RB
PI 373983A[b]	*G. microphylla*	RB	N/A	RB	N/A	N/A	N/A	RB	RB
PI 446988[b]	*G. tomentella*	IM	N/A	IM	IM	IM	IM	RB	RB
PI 505151[b]	*G. argyrea*	RB	RB	RB	RB	RB	RB	RB	RB
PI 483218[b]	*G.tomentella*	RB	RB	RB	RB	IM	RB	RB	RB

[a] Lines resistant to Powdery Mildew; [b] Lines susceptible to Powdery Mildew.
Source of isolate: Griffin-Griffin, GA.; APK- Attapulgus, GA.; APS1, APS2-3 and, APS4- Three different soybean fields in Attapulgus, GA.; Thomas- Thomas., GA.; Expo- Moultrie, GA.; TVP- Tifton, GA.
RB- reddish brown, resistant reaction; IM- immune, no symptoms; N/A- indicates data is not available due to failure of seed germination.
Three colors indicate three different isolates.

Table 2.4 Number of genotypes in ASR Georgia field isolates collected from different locations in Georgia based on sequence analysis of the Internal Transcribed Space (ITS) region.

Reference genotypes[a]	Griffin[b]	APS1	APS2-3	APS4	APK	Thomas	Expo	TVP	Total locations[c]
Brazil-1	8	9	13	13	5	6	15	9	8
MUT Zimbabwe	2	1	4	0	2	5	1	3	7
OK07-1A	4	1	0	2	5	2	2	1	7
GA1-Brazil	2	0	0	0	0	1	1	1	4
GA2-Brazil	1	0	1	0	0	1	3	0	4
GA1-Zimbabwe	0	0	0	0	0	0	2	0	1
GA2-Zimbabwe	0	1	0	0	0	0	1	2	3
FL08-2B	0	4	0	1	2	2	0	1	5
GA1-FL08-2B	1	1	1	0	0	0	0	1	4
MS08-2D	0	0	1	0	0	0	1	0	2
India 73-1	0	1	1	1	1	0	1	0	5
MS07-1E	0	0	0	1	0	0	1	0	2
TH02-1A	0	1	0	0	0	0	0	0	1
BZ01-1C	0	0	0	1	0	0	0	0	1
AR07-2A	2	1	1	1	1	1	1	1	8
MO06-1D	10	11	9	7	14	12	4	10	8
GA1-MO	1	0	1	0	1	0	0	0	3
GA2-MO	4	2	0	5	1	3	1	1	7
GA3-MO	0	0	1	0	0	0	1	0	2
AU79-1/HW94-1B	0	0	2	1	1	0	1	2	5
Genotypes[d]	10	11	11	10	10	9	15	11	

[a] Names of the reference genotypes from GeneBank that were identical with genotypes from Georgia ASR field isolates.
[b] Number of clones with the sequences that were identical with reference genotypes from GeneBank.
[c] Number of locations were the reference genotypes were found.
[d] Number of genotypes that each field isolate carried without duplications.

Table 2.5 Genetic diversity of Internal Transcribed Space (ITS) and the ADP-Ribosylation Factor gene isolated from 8 field isolates collected from different locations in Georgia.

Candidate genes	Isolate	Clone[a]	Sites[b]	Indels[c]	SNPs[d]	π[e]	K[f]
ITS	APK	10	668	8	3	0.247%	1.62
	APS1	12	668	5	3	0.219%	1.44
	APS2-3	10	668	4	3	0.247%	1.62
	APS4	11	668	8	3	0.243%	1.60
	Expo	15	668	9	3	0.237%	1.56
	Griffin	10	668	4	3	0.243%	1.60
	Thomas	9	668	8	3	0.228%	1.50
	TVP	11	668	8	3	0.238%	1.56
ARF	APK	4	666	1	4	0.324%	2
	APS1	5	666	1	3	0.212%	1.4
	APS2-3	3	666	1	1	0.101%	0.667
	APS4	5	666	1	5	0.303%	2
	Expo	2	666	1	0	0	0
	Griffin	6	666	1	7	0.414%	2.73
	Thomas	4	668	1	1	0.076%	0.5
	TVP	2	666	1	0	0	0

[a]Clones represents the number of clones sequences after removal of duplicated sequences.
[b]Sites represents the nucleotides in the analyzed region.
[c]Indels represent the number of indels in the analyzed region.
[d]SNPs represents the number of single nucleotide polymorphisms in the analyzed region.
[e]π= (Nucleotide diversity) represents the average number of nucleotide differences per site between two sequences (Nei 1987).
[f]K=Average number of nucleotide differences.

CHAPTER 3

IDENTIFICATION OF DIFFERENTIALLY EXPRESSED SOYBEAN DEFENSE-RELATED GENES AND ASIAN SOYBEAN RUST EFFECTOR GENES IN RESPONSE TO *P. PACHYRHIZI* INFECTION[1]

[1] P. Tian and S. M. Smith. To be submitted to Molecular Plant Microbe Interaction.

Abstract

Obligate basdiomycete *Phakopsora pachyrhizi*, is the causal agent of Asian Soybean Rust and is one of the most destructive foliar soybean diseases. The main objective of this study was to perform transcriptome profiling of *P. pachyrhizi* and resistant and susceptible soybean plants using dual-RNA sequencing in order to characterize putative soybean defense-related genes and soybean rust effector genes that are associated with resistance to ASR. Soybean plants were inoculated with one characterized Georgia field isolate three weeks following planting. A dual-RNA sequencing method was performed and generated two transcriptome datasets from a susceptible (Williams 82) and resistant soybean cultivar (Hyuuga) to identify differentially expressed host genes and putative *P. pachyrhizi* effector genes. A total of 3,768 differentially expressed genes (DEGs) were identified in Williams and 2,759 DEGs were identified in Hyuuga. Comparison of the DEGs in Hyuuga and Williams 82 identified genes that were expressed in the incompatible interaction and compatible interactions, suggesting that several NBS-LRR R-genes, *WRKY* genes, and *MYB* genes may be responsible for the incompatible interaction in Hyuuga. Additionally, 1,786 putative fungal effector genes were also identified from the assembled transcripts. Twenty-six of these genes were further characterized as putative pathogen secreted effector proteins.

Introduction

Asian Soybean Rust (ASR) is caused by *Phakopsora pachyrhizi*, and it is one of the most virulent diseases of soybean. In 2007, the United States Soybean Rust Commentary (http://www.sbrusa.net/) reported the detection of soybean rust spores in soybean fields in Alabama (five counties), Florida (ten counties), Georgia (five counties), Louisiana (five parishes) and Texas (six counties). Currently, much effort has been directed towards identifying soybean resistance against ASR. There are no commercial soybean cultivars that confer resistance to the current ASR population. The predominant disease management strategy used to control ASR is fungicide application, which increases the cost of production and may negatively impact the environment.

One predominate isolate (Expo) was identified among collected Georgia field isolates and one Japanese soybean cultivar (Hyuuga) that was resistant to all of the ASR field isolates was also identified (Chapter 2). However, very little information is available regarding the molecular mechanisms that contribute to *P. pachyrhizi* pathogenesis and defense responses in soybean against ASR. Therefore, understanding the molecular mechanisms involved in host-pathogen interactions is critical for the development of strategies to control ASR in a sustainable manner. Recently, with the development of sequencing techniques and gene expression analysis on a large scale, new bioinformatics tools for data analysis, the soybean response to rust infection can be analyzed with several transcriptional profiling methods such as microarray and RNA sequencing analyses. These techiniques can be used to identify various genes involved in host resistance and susceptibility. For example, Choi *et al*. (2008) performed a large-scale transcript profiling analysis on *Rpp1* soybean plants and revealed an up-regulation in the expression of genes, especially for lipoxygenease and peroxidase from the resistant plants, indicating that these genes confer important functions in *Rpp1*-mediated resistance. Intriguingly,

a bi-phasic response of soybean to the pathogen was found in both *Rpp1*- and *Rpp2*-mediated resistance (Choi *et al*., 2008; van de Mortel *et al*., 2007; Panthee *et al*., 2007). It showed that from 6 to 36 hour post-inoculation (hpi), which is the stage of spore germination and penetration, the first wave of differential gene expression in compatible and incompatible interactions was observed and reached a peak at 12 hpi. Beginning at 72hpi, which corresponded to haustoria mother cell and haustoria formation, the second wave of defense-related gene expression was observed. The increase in gene expression was more rapid in the resistant lines than in the susceptible lines. Therefore, the pattern of gene expression changes during the early infection stage might represent a non-specific recognition of the pathogen and activation of basal defense response in both resistant and susceptible, while the second wave of differential gene expression suggested recognition of the effector proteins by the resistance genes, triggering the rapid resistance response in the resistant line (van de Mortel *et al*., 2007).

Fewer studies have been conducted on transcript abundance in the pathogen since it is hard to isolate pathogen tissue from the host plant material. Previous work utilized mRNA extracted from infected leaves from soybean cultivar Williams 82 at two time-points to construct a cDNA library. Over 5,981 and 6,390 expressed sequence tags (ESTs) corresponded to soybean genes and *P. pachyrhizi* genes, respectively (Posada-Buitrago *et al*., 2006; unpublished data). Based on differential gene expression analysis within *P. pachyrhizi* germinating spores, 189 of 488 identified unique ESTs showed significant similarly (E-value < 10^{-5}) to sequences deposited in the NCBI non-redundant protein database (Posada-Buitrago and Frederick 2005). Additionally, using laser capture microdissection (LCM), Tremblay *et al*. (2009, 2010) isolated susceptible soybean palisade and mesophyll cells that showed signs of infection, extracted RNA transcript, performed transcriptome profiling, and generated a limited number of *P. pachyrhizi*

ESTs. About 80% of identified genes shared no homology to previously described soybean rust genes, which suggested stage-specific gene expression in the development of uredinia.

There is increasing evidence that fungal and oomycete pathogens secrete numerous effector proteins into their hosts (e.g. Kamoun 2007; Hogenhout *et al.*, 2009; Tyler 2009; Birch *et al*., 2009; Hok *et al*., 2010; Ravensdale *et al*., 2011). Those effector proteins can manipulate host cell structure and regulate the host genes function, thereby facilitating infection (virulence factors or elicitors) (Kamoun 2006). To date, six effector proteins have been identified in three different rust species: AvrM, AvrL567, AvrP123, and AvrP4 in the flax rust fungus *Melampsora lini*, the Rust Transferred Protein (RTP1) in the bean rust fungus *Uromyces fabae*, and PGTAUSPE-10-1 in the wheat stem rust fungus *Puccinia graminis* f. sp.*tritici* (Kemen *et al*., 2005; Ellis *et al*., 2007; Upadhyaya *et al*., 2014). In terms of *P. pachyrhizi*, 20% of all ESTs from germinating urediniospores were found to encode a small secreted, cysteine-rich protein similar to proteins expressed by *Blumeria graminis* f. sp. *Hordei* and *Magnaporthe grisea* during spore germination and appressorium formation (Posada-Buitrago and Frederick 2005), indicating that these proteins may play a very important role in early stages of infection.

The identification of rust and bacterial core effectors and the corresponding resistance proteins from the host plant has been utilized to identify potentially durable resistance genes (R-genes). The resistance conferred by the R-gene(s) is likely to be durable when the corresponding core effector has an important virulence function in the pathogen and is therefore conserved in the pathogen population. This was observed for *Xanthomonas axonopodis* pv. *manihotis* (*Xam*), the causal agent of cassava bacterial blight (Bart *et al*., 2012). One core effector was identified in ~65 *Xam* strains collected over a 70-year time frame from 12 countries. The core effector was shown to serve as a target activated by R-genes in wild species of *Manihot* and potentially other

related plants in the *Euphoribiaceae*. In addition, several oomycete were demonstrated to contain conserved and functional RXLR variants which could mediate their transduction into plant cells (Rafiqi *et al*., 2010; Kale *et al*., 2010; Gu *et al*., 2011; Plett *et al*., 2011; Tian *et al*., 2011).

To identify potential host durable resistance associated with characterizing core effector proteins as well as host resistance proteins, both host and pathogen datasets during infection process are essential. Therefore dual RNA-seq is needed to analyze gene expression changes in both the pathogen and the host simultaneously. This technique allows monitoring of gene expression in two different organisms at high level of accuracy during the infection process. RNA-seq is more sensitive, accurate and economical, than microarray-based approaches (Rappuoli *et al*., 2000), providing high-resolution data with very high coverage of the transcriptomes. More importantly, it does not require the design of new chips for experiments analyzing different pathogens or hosts and is therefore a species-independent platform (Westermann *et al*., 2012). Although transcriptomic experiments have predominantly focused on either the host or the pathogen, understanding of the infection process will require the simultaneous analysis of both interaction partners. With the availability of the genome sequence of several plant pathogens, differential gene expression was also analyzed in the sudden oak death pathogen *Phytophthora ramorum* and its most susceptible forest host, tanoak. Analysis of six *Puccinia triticina* isolates and wheat, as well as rice blast fungus and rice, provided simultaneous elucidation of the tactics of host plant defense and pathogen attack, which also assisted with characterizing fungal secretomes during infection (Kawahara *et al*., 2012; Bruce *et al*., 2014; Hayden *et al*., 2014).

In this study, a pair-end dual RNA-seq technique were performed on a susceptible soybean cultivar (Williams 82) and a resistant cultivar (Hyuuga) following inoculation with the

Expo Georgia field isolate *P. pachyrhizi* using an Illumina platform. The main objectives of this study were to analyze the large dataset of DEGs to identify the putative soybean defense-related genes that are involved in host resistance and to characterize the putative *P. pachyrhizi* effector genes that were secreted into plant tissue and potentially triggered the plant defense. This study identified putative targets in the host and pathogen that are responsible for plant defense and pathogen infection, respectively. With the knowledge of the effector proteins present in ASR population, as well as the resistant protein in the host, this research will provide information that can be utilized to assist selecting more effective and durable resistance against ASR.

Materials and Methods

Pathogen Isolation and Plant Inoculation

An ASR field isolate (Expo) was harvested from field-collected soybean leaves in Moultrie, Georgia in November 2013. The identity of this isolate was visually confirmed under the microscope and with polymerase chain reaction (PCR) as previously described in chapter 2. It is a mixture of different genotypes with high genetic diversity. Urediniospores were increased on a susceptible soybean cultivar Williams 82. ASR urediniospores were collected from the Williams 82 cultivar beginning at 10 days post inoculation (dpi) and continued to 14 dpi. To prepare the inoculum, urediniospores collected from the field were suspended in sterile distilled water containing 0.01% (v/v) Tween 20 (Sigma-Aldrich, St Louis, MO) and filtered with 100 µm cell strainer (Corning Life Sciences, Tewksbury, MA) to remove plant debris and clumps of spores. Spore concentration was determined using a hemocytometer and adjusted as needed to a final solution of 6×10^4 spores/ml. Using a spray bottle, 10ml of the spore suspension was sprayed on to the surface of each leaflet. After inoculation, the plants were placed in a black plastic bag that was used as a humidity chamber with a relative humidity close to 100% to

promote infection. At 16 hour post inoculation (hpi), the plants were moved to a growth chamber. On 14 dpi the leaves of ASR-infected susceptible soybean plants with high sporulation were collected and placed in to a 250ml flask. Urediniospores were washed away from the leaf tissue with 0.01% Tween 20, filtered and adjusted to the appropriate concentration as described previously. The resulting spore suspension was used to inoculate two cultivars (Hyuuga and Williams 82) 21 days after planting. Three plants of each cultivar in 10 cm-pot were prepared in three replicates (pots). *P. pachyrhizi*-inoculated and mock-inoculated plants from the two cultivars were kept in two different growth chambers. The first trifoliolate leaves (3-4/time point) were collected at seven time points (6 hpi, 12 hpi, 24 hpi, 48 hpi, 72 hpi, 96hpi, and 120hpi). Each leaf was about 0.6 g and was flash frozen in liquid nitrogen before RNA extraction.

RNA Isolation, cDNA Library Preparation, and Illumina RNA Sequencing

Total RNA was extracted using TRIzol (Invitrogen, CA) according to the manufacturer's instructions and stored in -80 °C freezer. Four experimental groups were utilized: (1) mock susceptible (MS), (2) mock resistant (MR), (3) inoculated susceptible (IS) and (4) inoculated resistant (IR). RNA samples collected 72 hpi were utilized for RNA sequence analysis. Each set of samples represent three biological replicates. RNA sample preparation, mRNA enrichment, cDNA library preparation, and RNA sequencing analysis was performed at the Georgia Genomics Facility (GGF) at the University of Georgia (http://dna.uga.edu/) according to standard protocols. A total of 12 cDNA libraries were generated and sequenced and paired-end 75-bp long RNA-seq reads were generated using KAPA stranded mRNA-seq Kit (Kapa Biosystems, Boston, MA) on a Next-Seq 500 sequencing system (Illumina, San Diego, CA).

RNA Sequencing Data Processing and Assembly

The sequence analysis associated with the RNA sequencing data analysis were performed using zcluster queueing system by GACRC (The Georgia Advanced Computing Resource Center) in The University of Georgia at Athens, GA. A quality control was conducted using FastQC 0.8 software and Phred Score measure (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/). Bases with a length less than 30bp were trimmed from both ends of the raw sequence reads (fastq_quality_trimmer -t 30). About 80% of the trimmed reads with quality scores greater than 20 were retained for assembly (fastq_quality_filter -q 20 -p 80). These tools are part of the FASTX-Toolkit (http://hannonlab.cshl.edu/fastxtoolkit/). Quality-filtered reads were mapped to the reference genome using Tophat, which internally uses Bowtie as its high-throughput read alignment tool. The resulting reads were mapped to version 2.0 of the *G. max* reference genome (Schmutz *et al*., 2010) provided in Phytozome v10.1 (www.phytozome.net/soybean/) using TopHat v2.0.13 (Trapnell *et al*., 2010) in conjunction with Bowtie v1.1.0 (Langmead *et al*., 2009) using all default parameters (Table 1). The reads were concatenated using Cufflinks v2.2.1 (Trapnell *et al*., 2010). The generated transcript assemblies in each sample from the Cufflinks were merged using the Cuffmerge tool (Trapnell *et al*., 2012).

All of the un-mapped reads were converted to the fastaq files and were subsequently re-mapped to the soybean reference genome to remove the rest of the soybean reads. The remaining reads were considered as potential ASR reads. *De novo* transcriptome assembly was accomplished using Trinity version r20140717 (Grabherr *et al*., 2011) with the following settings (--seqType fq --JM 30G --SS_lib_type RF --CPU 4 --output all-trinity-output).

Differential Gene Expression Analysis

The reads from Tophat and merged assemblies from Cuffmerge were fed into Cuffdiff2 (Trapnell *et al*., 2012) for testing differential expression. This allows testing for the presence of differentially-expressed genes between two different treatments and cultivars and the identification of gene sets showing statistically significant changes (fold > 2, p-value < 0.05). The two treatment comparisons were: (1) inoculated resistant (IR) versus mock resistant (MR) and (2) inoculated susceptible (IS) versus mock susceptible (MS). The IR vs IS was not included because this comparison may remove some genes that are associated with plant defense in response to pathogen inoculation. The MR vs MS was also not included because it doesn't contribute to the identification of differentially expressed defense-related genes in response to the infection o*f P. pachyrhizi*.

Gene Ontology Analysis

All the significantly expressed transcripts from resistant cultivar Hyuuga were compared with those of the susceptible Williams 82 cultivar. The overlapped transcripts were subtracted to identify the significantly expressed transcripts that only belong to the resistant or susceptible cultivar. To annotate the transcripts of interest, they were run against the NCBI non-redundant (nr) database using BLASTx with an Expect (E)-value cut-off of 1×10^{-5}. Gene Ontology (GO) terms mapping and annotation analysis were performed using the Blast2GO software with default parameters (Aparicio *et al*., 2006). Genes that were identified in each comparison were categorized based on their function, and were analyzed to determine their potential roles in plant defenses. This was done by comparing the predicted function of each gene with the function of known genes described in an incompatible interaction during a biotrophic interaction.

Specific Gene Expression Comparison

The transcripts that were identified as highly similar with R-genes, transcription factor (TF) *WRKY* genes, and *MYB* genes were selected for structure and phylogenetic analysis. The full-length of the target genes were obtained by searching gene ID and the chromosome location against the soybean genome in Phytozome (http://www.phytozome.net/search.php). Protein domains were annotated in each R-gene, and the R-genes that carry NBS (nucleotide binding site) or LRR (Leucine Rich Repeat) regions were selected. Based on the amino acid sequences of the R-genes, TIR, and Non-TIR genes were characterized by the last residue in the Walker-B (Kinase 2) motif of NBS region (Meyers *et al*., 1999). Full length and truncated NBS-LRR R-genes were characterized by using the HMMER program (Eddy 1998). The resulting NBS-LRR R-genes that were located on chromosome 16 were compared with the RGHs amino acid sequences characterized from the fosmid analysis in chapter 4. Phylogenetic analysis was performed with MEGA (Molecular Evolutionary Genetics Analysis) program version 5 (Tamura *et al*., 2011) using ClustalW Neighbor Joining (NJ) method, with distance represented as the number of amino acid differences. Confidence in the phylogeny was assessed with one-thousand bootstrap replicates. Amino acid sequence similarity tests were performed using pairwise alignment methods by BioEdit (Biological sequence alignment editor) version 7.2.5 (Hall 1999). The name of *WRKY* genes were searched based on the gene ID from Soybase (http://soybase.org/).The names of the *GmWRKY*1-64 were assigned according to Zhou *et al* (2008). *GmWRKY*65-182 was assigned based on the chromosome order. The classification of each *WRKY* genes was done according to Eulgem *et al.* (2000). The expression confirmation was performed based on Soybase ESTs, RNA-seq analysis (in silico analysis) and RNA-seq of *P. pachyrhizi* lesion microdissection (http://soybase.org/). *MYB* genes were classified based on the

number of repeats in the protein structure (Dubos *et al*., 2010). Their amino acid sequences were searched against *Arabidopsis* genome database and other previous reported *MYB* genes to further identify their potential functions (https://www.arabidopsis.org/servlets/Tair).

Identification of *P. pachyrhizi* Effector Proteins

The assembled transcriptome from Trinity were first compared to plant and fungal genome database located in UnitProKB (http://proteomics.ysu.edu/publication/data/PlantSecKB/) to remove the plant genes assembled by Trinity as well as the transcripts that have no known function (Lum and Min 2011). The fungal transcripts were compared with fungal secretome KnoledgeBase (FunSecKB, http://proteomics.ysu.edu/secretomes/fungi.php) using blastx with an E-value cutoff of $<1 \times 10^{-5}$ to identify rust candidate secreted effector proteins (CSEPs) (Lum and Min 2011). The CSEPs were compared with other rust effector proteins that have been reported previously to further identify their putative function and similarity with known rust effectors. The known rust effector genes were obtained from Genebank by searching key words including "*Melampsora*", "*Puccinia*", "*Uromyces*", "secreted", "hesp", "Avr", etc. Only the transcripts with the E-value < 1×10^{-5} and an identity > 70% were selected. The resulting transcripts were converted to amino acids that were then subjected to structural analysis to identify potential conserved regions, signal peptide as well as transmembrane domains. HMMER v3.1b2 was used to search sequence databases for homologs of protein or DNA sequences and further identify conserved regions (Eddy 1998). Phobius program was used for prediction of transmembrane topology and signal peptides from the amino acid sequence of a protein (Käll *et al*., 2004).

Results

Phenotype Confirmation

At 10 dpi, phenotypic analysis confirmed the presence of a hypersensitive reaction on the inoculated resistant line (Hyuuga), while heavy sporulation was found on the inoculated susceptible line (Williams 82) (Figure 3.1A). The mock inoculated plants from the two lines were free of infection, which confirmed that there was no contamination (Figure 3.1B).

RNA Sequencing Analysis of Transcriptome Samples

Twelve RNA samples were sequenced from one time point (72 phi) with three replications, yielding an average of 30 million reads (2 × 75bp paired) per sample using the Next-Seq 500 sequencing system, which resulted in more than 100X coverage. Figure 3.1 schematically illustrates the process flux of the experimental procedure. After trimming and filtering of the raw reads, the mapping ratio of the resulting reads to the *G. max* reference genome ranged from 92.7% to 94.2% (Table 3.1).

The number of un-mapped reads was about 1.7 million. The un-mapped reads were then mapped against the soybean reference genome again to remove potential plant reads, leaving the pathogen reads in the dataset. The ratio of mapped the un-mapped reads to the soybean reference genome was from 1.1% to 1.4%, representing all of the plant reads that were removed afterwards. After trimming and filtering, a total of 4.4 million reads with high quality were used for *de novo* assembly (Table 3.2, figure 3.4).

Identifying Differentially Expressed Genes (DEGs)

DEGs were determined by comparing the expression profile data from inoculated and mock-inoculated treatments at 72 hpi. Two comparisons were made: 1) Inoculated Resistant line

(IR) versus Mock Resistant (MR) and 2) Inoculated Susceptible (IS) versus Mock Susceptible (MS).

A total of 2,759 and 3,768 genes from the IR-MR and IS-MS genotypes were identified, respectively (Table 3.3). A cut-off threshold of >2 fold change in expression and a P-value below 0.05 was set to identify the DEGs that were significantly differentially expressed during the infection of *P. pachyrhizi*.

Identifying DEGs Related to Host Plant Defense

Following the categorization of DEGs based on the predicated function, the genes related to plant defense were selected (Table 3.4). The function of these genes was predicated by GO analysis and included genes involved in defense response, response to fungus, response to bacteria, response to virus, oxidative stress, systemic required resistance, response to jasmonic acid stimulus, response to incompatible interaction, response to innate immune system, response to salicylic acid biosynthetic process, cell death, pathogenesis related genes (PR-genes), R-genes, reactive oxygen species, MAP kinases and other cell rescue activities. Among the predicted defense-related DEGs in the four datasets, 127 were up-regulated and 63 were down-regulated in the IR-MR comparison. The number of up-regulated DEGs was more than twice the number of down-regulated genes. The fold change observed for the up-regulated genes ranged from 2 to 24.57. Four genes in IR (Glyma.07g005700, Glyma.10g070300, Glyma.15g145700, and Glyma.19g213500) were expressed in response to inoculation with ASR Expo isolate. Among the four genes, Glyma.15g145700 was similar to class-10 pathogenesis-related protein 1-like, while Glyma.07g005700 was identified as a serine threonine-protein kinase. Fewer DEGs were down-regulated in the IR and the fold change of the down-regulated genes ranged from 2 to 22. There were four genes (Glyma.13g278000, Glyma.16g212400, Glyma.09g185500, and

Glyma.11g180500) that were not expressed in the IR. Among these four genes, Glyma.13g278000 was classified as a defensin gene, while Glyma.16g212400 was very similar to a pathogen-inducible trypsin-inhibitor-like protein.

In contrast to the large number of differentially expressed genes observed in the defense-related genes category for the IR, the number of differentially expressed defense-related DEGs was almost equal in the IS. Eight-two putative defense-related DEGs were up-regulated in the MS and 84 were down-regulated in the IS. The fold change of up-regulated genes ranged from 2 to 11.23. The expression of one gene (Glyma.13g278000) was turned on and was identified as a defensin. Interestingly, this gene was not expressed in the IR. More defense-related DEGs were down-regulated in IS. The fold change ranged from 2 to 84.44. There were two genes,Glyma.18g088600 and Glyma.02g121500 that were not expressed in the IS and were similar to the disease resistance protein RPM1-like and mads-box protein CMB1.

Specific DEGs such as R-genes and TFs were selected based on their predicted functions associated with the defense response of the host plant (Table 3.4). All of the up-regulated R-genes were identified as defense-related genes, while the majority of the TF genes were not defense-related genes (Table 3.4).

Identification of R-genes Expressed in Compatible and Incompatible Interactions

The gene ID of each R-gene was used to run them against the Phytozome database to obtain the full length DNA sequences. The sequences were subjected to structural analysis to facilitate the identification of target protein domains within each gene. Only the R-genes with an NBS or LRR region or both were selected for further analysis. There were eighteen up-regulated R-genes identified. No down-regulated R-genes were found in the IR, while fifteen were up-regulated in the IS and five were down-regulated in the IS (Table 3.4). Based on the structure of

each R-gene, fourteen of the eighteen IR up-regulated R-genes were from the TIR-NBS-LRR class of resistance genes. Three IS up-regulated and two IS down-regulated R-genes were from the TIR-NBS-LRR R-genes. However, nine of the fourteen R-genes were full-length NBS-LRR R-genes (Glyma.02g026200, Glyma.03g077400, Glyma.12g011700, Glyma.18g093300, Glyma.18g093800, Glyma.16g135200, Glyma.16g033900, Glyma.16g215000, and Glyma.16g137600). The remaining five R-genes carried a partial NBS or TIR region, or only an LRR region.

The location of these NBS-LRR R-genes on the chromosome, their corresponding putative soybean genes and expression levels were determined (Table 3.5). Six R-genes were located on chromosome 16, while all of the other R-genes were distributed on chromosome 2, 3, 8, 9, 12, 13, 18 and 19. None of genes have been cloned or analyzed with the exception of Glyma.16G215000. This gene was highly similar to a soybean NBS-LRR R-gene, LM6.

Glyma.02G026200 and Glyma.02G030700 were located on chromosome 2. Both of the genes were predicted to be putative soybean RGA 3-like genes. The genes were classified as a non-TIR NBS-LRR gene. The identities of nucleotide and amino acid sequences were 74.32% and 74.33%, respectively. In terms of differential expression, Glyma.02G026200 was up-regulated in the IR (>3.5 fold) while Glyma.02G030700 was down-regulated in the IS (>2.29 fold).

Six R-genes on chromosome 16 were up-regulated in the IR with fold a change ranging from 2.1 to 3.6. They were identified as NBS-LRR R-genes carrying a TIR region. Phylogenetic analysis of the amino acid sequences of the NBS region of the four full-length R-genes and one truncated R-genes was performed to analyze their relationship with previously characterized RGHs from a soybean fosmid library (Chapter 2) (Figure 3.3). Glym.16G215000 was very

similar to RGHs 15-3, 15-4, and 15-7. The amino acid similarity ranged from 84.31% to 91.85%, and all of the genes clustered on the clade with RGH 15-3. The other R-gene (Glym.16G215100) was less similar to the other fosmid RGHs. Similarity of NBS amino acid ranged from 57.43% to 64.92%, even though it was grouped with 15-7 on the same clade. Glym.16G137600 was very similar to Glym.16G135200 (91.3%), but they are not grouped with any of RGHs from the fosmid analysis (Table 3.6).

Putative Transcription Factors (TFs) Involved in the Molecular Response to *P. pachyrhizi*

The total numbers of differentially expressed TFs in the IR and the IS were similar and ranged from 35 to 48 (Table 3.4). The IR had the fewest number of up-regulated TFs (35) but had the most number of up-regulated defense-related TFs.

The *WRKY* gene is considered to be involved in the plant defense response of multiple biotic factors (Table 3.7). Seven up-regulated and three down-regulated *WRKY* genes were identified in the IR. The average expression level (5.9 fold) of the up-regulated *WRKY* genes in the IR was higher than the level of *WRKY* genes in the IS (3.3 fold).

Other TFs including *MYB* genes were identified in the DEGs in both the IS and the IR. The IR had the greatest number of up-regulated *MYB* genes while IS had the most number of down-regulated *MYB* genes (Table 3.8). Based on the protein structure of the *MYB* gene, specific regions were used to screen all of the *MYB* genes and identified the potential class (Dubos *et al*., 2010). Eight of eleven *MYB* genes identified in this study belong to the R2R3 class. In terms of the function of *MYB* genes, only Glyma.05G027000 was shown to be associated with defense. Glyma.05G234600 and Glyma.08G042100 were involved in metabolism. The function of most of the *MYB* genes was unidentified.

Rust Transcriptome Analysis and Identification of Candidate Secreted Effector Proteins (CSEPs)

The pipeline for the selection of putative secreted effector transcripts (Figure 3.4) identified a total of 143 *P. pachyrhizi* CSEP transcripts. They were similar to known effector genes from *Melampsora* Spp., *Puccinia* Spp. and *Uromyces* Spp. With high similarity to known effector genes and proper length, twenty-two transcripts were selected (Table 3.9). Nine of the twenty-six CSEP transcripts were identified as genes encoding *Melampsora larici-populina* 98AG31 hypothetical protein while the other CSEPs were identified as genes encoding *Melampsora larici-populina* 98AG31 secreted protein. Since conserved regions have not been identified in fungal effector proteins, structural analysis targeting potential conserved regions, signal peptide, and transmembrane domains was performed. It has been shown the most fungal effector proteins have a signal peptide and some have a transmembrane domains (Djamei and Kahmann 2012; Guyon *et al*., 2014; Petre *et al*., 2014). One CSEP (c22424_g1_i1) had one single signal peptide in the N-terminus and two CSEPs (c13200_g1_i1 and c11267_g1_i1) had one single signal peptide and multiple transmembrane domains. There were two CSEPs (c1581_g1_i1 and c10027_g2_i1) with a transmembrane domain and a cysteine rich domain. They were categorized as cysteine rich proteins.

Discussion

There was a significant difference in the number and diversity of transcripts in defense-related DEGs in the IR and the IS. The increased number of up-regulated defense-related genes in the IR indicated that more genes were activated to participate in the *Rpp?*-mediated resistance reaction and plant defense, in addition to the DEGs associated with basal defense.

Salicylic acid (SA) and jasmonic acid (JA) are very important plant hormones that are involved in multiple signaling pathways during host defense activation (Delaney *et al*., 1994). It

has been shown that SA rather than JA plays a major role in expression of defense-related genes and cell death response in an incompatible interaction, while JA and ethylene responses were induced during a compatible interaction. This leads to the induction of tryptophan biosynthesis and activation of plant genes encoding defensins (Glazebrook 2005; Wasternack 2007). This study identified pathogenesis-related protein 10 (PR-10) (Glyma.15g145700) and serine threonine-protein kinase (Glyma.07g005700) in the IR. Both of these genes are associated with SA-dependent cell death response. Interestingly, the expression of a defensin gene (PR-12) (Glyma.13g278000) was not detected in the IR, while this genes was expressed in the IS. As the induction of plant genes encoding defensins are associated with JA-dependent response in compatible interaction, these findings indicated that JA-dependent signaling pathway was inhibited while SA-dependent signaling pathway was up-regulated in the *Rpp?*-mediated resistance reaction in Hyuuga. Interestingly, the other gene whose expression was not detected (Glyma.16g212400) was a pathogen-inducible trypsin-inhibitor-like protein. This protein is an anti-nutritive protein that competitively inhibits digestive serine proteases of animals and other organisms and it can be induced by wounding and jasmonic acid in several plant species. This supports the idea that SA-induced resistance is inhibiting JA-induced resistance in soybean in response to the infection of ASR (Thaler *et al.*, 1999; Cipollini and Bergelson 2000, 2001; Cipollini *et al.*, 2004; Musser *et al*., 2005).

The expression of one defense-related gene, Glyma.18g088600, was not detected during a compatible interaction. This gene was identified as a typical NBS-LRR R-genes that was very similar to the *Rpm1* gene that encodes a LZ-NBS-LRR protein and confers resistance to the bacterial pathogen *P. syringae* expressing either, *avrRpm1* or *avrB* (Grant *et al*., 1995). It was also reported that the recognition of *RPM1* is through effector-mediated modifications of the

RPM1-interacting protein 4 (RIN4) in the presence of the bacterial type III effector proteins *AvrRpm1* or *AvrB* (Mackey *et al.,* 2002). The suppression of the expression of the *Rpm1* like gene in the IS suggested that this gene was negatively controlled by other genes to confer the resistance function and induce cell death.

NBS-LRR R-genes encode a protein with an LRR region that is responsible for directly or indirectly interacting with the pathogen effector protein. A nucleotide binding site region in the N-terminus is responsible for transduction of a signal to activate plant defense (Dangl and Jones 2001). NBS-LRR R-genes were found to be highly regulated based on several RNA-seq Analysis of different host-path interactions, such as *Citrus reticulata* - *Xylella fastidiosa*, Peach - *Xanthomonas arboricola* pv. *Pruni* and *Arabidopsis* - *Fusarium oxysporum*. However, R-genes differentially expressed during soybean *P. Pachyrhizi* interactions have not been characterized. In this study, a total of 19 NBS-LRR R-genes were significantly differentially expressed in the IR and the IS. All the R-genes were categorized as defense-related genes. More R-genes were up-regulated in the IR than in the IS. None of the R-genes were down-regulated, indicating that one or more of the R-genes in the IR are possibly responsible for *Rpp?*-mediated resistance in Hyuuga and are therefore candidate resistance genes.

Meyer *et al.* conducted RT-PCR analysis of the candidate *Rpp4* gene and demonstrated that the *Rpp4C* gene is the primary candidate for *Rpp4*-mediated resistance. However, the expression level of *Rpp4* were similar in the *P. Pachyrhizi*-inoculated and mock-inoculated plants. This indicated that the NBS-LRR R-gene expression was maintained at a stable level with or without the infection of pathogen (Meyer *et al*., 2009). However, this study showed that the expression of the IR R-genes were up-regulated (>3 fold) at the 72hpi. This indicated that some of the R-genes that are associated with *Rpp?*-mediated resistance in Hyuuga may use different

mechanisms to mediate resistance and may require a higher level of expression to confer a complete resistance function.

In terms of the distribution of the identified R-genes, the clustering of the IR up-regulated R-genes on chromosome 16 and the location of several R-genes located on different chromosomes, demonstrated that R-genes are unequally distributed in soybean genome. At present, over 319 genes were determined to be putative NBS-LRR genes in the soybean genome. Most of these genes clustered on chromosome 3, 6, 13, 15, 16 and 18, and are likely the result of duplication events during the evolution of soybean (Kang *et al*., 2012). The clustering of R-genes will allow different types of sequence exchanges such as recombination, gene conversion, and un-equal crossing-over to maintain beneficial mutations and homogenization of R-genes (Radwan *et al.*, 2008; Zhang *et al*., 2011). Sequence comparison of the identified up-regulated R-genes and resistance gene homologs (RGHs) from the fosmid analysis in chapter 4, the full-length R-gene Glym.16G215000 and truncated R-gene Glym.16G215100 were grouped together with RGHs F15-3, F15-4, F15-7 in the same clade of the phylogenetic tree. They also shared a very high level of amino acid sequence similarity, indicating that they may belong to a same gene family or were derived from the same ancestor gene and evolved together. In addition, the truncation of Glym.16G215100 may have resulted from an unequal recombination event during the R-gene evolution.

Hyuuga-mediated resistance against ASR was previously reported to be associated with two R-genes that were mapped to the *Rpp3* locus on chromosome 6 and the *Rpp5* locus on chromosome 3 (Kendrick *et al*., 2011). There were more than 80 candidate genes present in the *Rpp5* region but none of the genes were annotated as an NBS-LRR gene (Garcia *et al*., 2008). In this study, none of the upregulated R-genes in Hyuuga were found on chromosome 6, indicating

that the *Rpp3*-like gene may not be differentially expressed during the infection and the expression of this gene may be highly regulated to reduce the fitness cost of the plant. In addition, the only upregulated TIR-NBS-LRR R-genes from chromosome 6 in Hyuuga may not be located at the same locus as the previously mapped *Rpp5*-like gene. Although the relationship between the mapped genes and upregulated R-genes in this study is not direct, it is still critical to verify the potential function of these R-genes to further analyze the mechanism of Hyuuga-mediated resistance.

The TFs play a very important role in the plant response in biotic and abiotic stresses. They are directly involved in the control of gene expression and participate in the regulation of biological processes, especially the response to biotic stresses caused by pathogens (Aoyagi *et al*., 2014). Many studies have implicated the importance of the soybean *WRKY* and *MYB* TFs in the response to ASR infection. These TFs corresponded to 50% of all of the TFs induced by ASR infection (van de Mortel *et al*., 2007; Panthee *et al*., 2007; Tremblay *et al*., 2009). It was demonstrated that both *WRKY* and *MYB* regulated the expression of defense genes, modulating immediate downstream target genes or activating/repressing other transcriptional factors (Dubos *et al*., 2010; Pandey *et al*., 2011).

In this study, more *WRKY* genes were up-regulated in the IS and the IR, demonstrating an increased activity of the *WRKY* genes in the compatible and incompatible interactions. Higher level of expression was observed for the IR *WRKY* genes in comparison to the IS, indicating that one or more of the *WRKY* genes that were highly expressed might be responsible for regulating the biological process associated with cell death and *Rpp?*-mediated resistance in Hyuuga. Benche-Malato *et al.* performed a comprehensive analysis to characterize the function of *WRKY* family genes involved in the response to *P. Pachyrhizi* infection (Benche-Malato *et al.*, 2014).

They compared the *WRKY* genes in response to *P. Pachyrhizi* infection that were identified previously using super SAGE, RNA-seq of LCM and microarray experiments and identified 8 induced *WRKY* genes that are expressed at a higher level or earlier in the resistant genotype. However, only one (Gm*WRKY* 125) of the 8 targeted *WRKY* genes was differentially expressed in this study. Gm*WRKY* 125 was up-regulated (>2.5 fold) in the IR when compared to the IS. Additionally, the expression of this gene peaked at 12 hpi. It is, plausible that this gene may be involved in non-specific defense responses. Two *WRKY* genes (Glyma.03G220800 and Glyma.19G217800) were down-regulated in the IR but were up-regulated in the IS, indicating that they may be involved in basal defense or other biological processes.

Another important group of TFs are the MYB genes. They are one of the largest TF families in the plant kingdom (Aoyagi *et al.,* 2014). There are four classes of MYB TFs in plants and the largest class, R2R3-MYB, is characterized by proteins with two R (four Helix-Turn-Helix (HTH) repetition). The high proportion of observed R2R3-MYB genes was also reflected in this study, where eight of eleven identified *MYB* genes belong to the R2R3 class. Only Glyma.05G01080 was shown to be homologous to two Arabidopsis MYB genes, which is associated with defense. This gene was up-regulated in the IR (>2.8 fold) therefore, may be involved in *Rpp?*-mediated resistance.

Plant pathogens can promote infection and inhibit the host plant defense system by secreting effector proteins (Petre *et al*., 2014). Currently, six effector proteins have been identified in three different rust species, *AvrM, AvrL567*, *AvrP123,* and *AvrP4* in the flax rust fungus *Melampsora lini*, the Rust Transferred Protein RTP1 in the bean rust fungus *Uromyces fabae*, and PGTAUSPE-10-1 in the wheat stem rust fungus *Puccinia graminis f.* sp. *tritici* (Kemen *et al*., 2005; Ellis *et al*., 2007; Upadhyaya *et al*., 2014). The six rust effector proteins are

thought to be translocated from the haustoria into the host cells. Therefore, the 72 hpi time point was selected for identifying candidate soybean rust effector proteins in this study (Tremblay *et al*., 2013). To obtain candidate soybean rust effector genes, both rust RNA-sequencing reads from susceptible and resistant lines were merged to perform *de novo* assembly. By setting strict selection condition with E-values or identity ration, the rust assembled transcripts were run against the fungal gene database, fungal secretome database, a known rust effector genes list and the plant cell wall degrading database.

The genes encoding potential secreted protein in *Melampsora Larici-populina* in Genebank provided a rich resource for comparison with the soybean rust CSEPs (Duplessis *et al.*, 2011). Twenty-six CSEPs were identified with this work that were either highly similar to the *Melampsora larici-populina* 98AG31 hypothetical protein or secreted proteins. However, little is known regarding the function of these proteins. Further structural analysis narrowed down the number of CSEPs to five, with a signal peptides and transmembrane domains. Most of fungal effector proteins had signal peptide domains that are normally located in the N-terminus as one single domain. Multiple transmembrane domains were found in several of the fungal effector proteins, although they are not a common characteristic for fungal effector proteins (Djamei and Kahmann 2012; Guyon *et al*., 2014; Petre *et al*., 2014). Two of CSEPs were characterized as cysteine rich proteins. According to previous reports, *Melampsora lini* effector protein *AvrP4*, *AvrP123* and the *Cladosporium fulvum* effector proteins *Avr2*, *Avr4*, and *Ecp2* were identified as cysteine rich proteins. This suggested that the five CSEPs with signal peptides and transmembrane domains were candidate effector proteins that are potentially associated with pathogenesis in the Expo ASR field isolate from Georgia. It is necessary to validate the expression of the CSEPs at different time points to monitor the change in expression ASR in

compatible and incompatible interaction. Following the molecular validation of the expression of the CSEPs, the establishment of functional validation methods will be essential to further verify how the CSEPs are delivered into the plant cell, how they function as an avirulence factor, and how they interact (directly or indirectly) with the host R-gene proteins.

Dual RNA sequence analysis took transcriptomics one step further by analyzing gene expression changes in the pathogen and the host simultaneously (Westermann *et al*., 2012). However, this technique does require high coverage for sequencing both the host and the pathogen reads at a high depth. It is demonstrated that an estimated minimum of about 2,000 million reads from total RNA and 200 million reads from rRNA-depleted samples seem to be require to simultaneously monitor gene expression in both host and pathogen (Westermann *et al*., 2012). In this study, only 4.4 million reads of rust were used for *de novo* assembly, which didn't assist us in establishing a reference transcriptome profile for differential expression analysis due to low quality and short length of the majority of assembled rust transcripts. The reason of low amounts of rust reads was caused by the low quantity of ASR cells present in ASR infected leaves. It was obvious that the majority of the RNA extracted from the soybean leaf infected with ASR was soybean RNA. To solve this issue, laser capture microdissection (LCM) technique is necessary for isolation the infected cells from non-infected cells, which can eliminate contaminating background gene expression of non-infected cells and increase the concentration of rust RNA in the RNA extraction product (Tremblay *et al*., 2010). This will be very useful since it is difficult to detect the presence and development of rust infection especially in the incompatible interactions.

This study provided valuable information and a comprehensive list of defense-related transcripts that showed a distinct difference between Hyuuga and Williams 82. These transcripts

included NBS-LRR genes, *GmWRKY* genes and *Myb* genes that will be used to better understand the mechanism of Hyuuga *Rpp?*-mediated resistance against the current ASR pathogen population in Georgia. Parallel analysis of the host and the pathogen also generated a list of candidate ASR effector proteins, that were similar to previously identified *Melampsora Larici-populina* secreted proteins. This is the first report of simultaneous transcriptome analysis of soybean and *P. Pachyrhizi*. This work provided insight into the complexity of the interaction between pathogen effector genes and host defense-related genes in compatible and incompatible interactions. With the knowledge of both ASR pathogenesis-associated effector proteins and soybean defense-related genes during the host-pathogen interaction, it will allow better selection of effective and potential durable resistance genes for the control of ASR.

Literature cited

1. Aoyagi L. N., Lopes-Caitar V. S., de Carvalho M. C., Darben L. M., Polizel-Podanosqui A., Kuwahara M. K., ... & Marcelino-Guimarães F. C. (2014) Genomic and transcriptomic characterization of the transcription factor family R2R3-MYB in soybean and its involvement in the resistance responses to *Phakopsora pachyrhizi*. Plant Science, 229:32-42.
2. Aparicio G., Gotz S., Conesa A., Segrelles D., Blanquer I., Garcia J. M., ... & Talon M. (2006) Blast2GO goes grid: developing a grid-enabled prototype for functional genomics analysis. Studies in health technology and informatics 120:194.
3. Bai J., Choi S. H., Ponciano G., Leung H., and Leach J. E. (2000) *Xanthomonas oryzae* pv. *oryzae* avirulence genes contribute differently and specifically to pathogen aggressiveness. Molecular Plant-Microbe Interactions 13:1322-1329.
4. Bart R., Cohn M., Kassen A., McCallum E. J., Shybut M., Petriello A., Krasileva K., Dahlbeck D., Medina C., Alicai T., Kumar L., Moreira L. M., Rodrigues Neto J., Verdier

V., Santana M. A., Kositcharoenkul N., Vanderschuren H., Gruissem W., Bernal A., Staskawicz B. J. (2012) High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proceedings of the National Academy of Science USA 109(28):972-1979.

5. Beissbarth T., Speed T. P. (2004) GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics 20:1464-1465.
6. Bencke-Malato M., Cabreira C., Wiebke-Strohm B., Bücker-Neto L., Mancini E., Osorio M. B., ... & Bodanese-Zanettini M. H. (2014) Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to *Phakopsora pachyrhizi* infection. BMC plant biology 14(1), 236.
7. Birch P. R., Armstrong M., Bos J., Boevink P., Gilroy E. M., Taylor R. M., Wawra S., Pritchard L., Conti L., Ewan R., Whisson S. C., van West P., Sadanandom A., and Kamoun S. (2009) Towards understanding the virulence functions of RXLR effectors of the oomycete plant pathogen *Phytophthora infestans*. Journal of Experimental Botany 60:1133-1140.
8. Bromfield K. R., and Hartwig E. E. (1980) Resistance to soybean rust [*Phakopsora pachyrhizi*] and mode of inheritance. Crop Science 20:254-255.
9. Bruce M., Neugebauer K. A., Joly D. L., Migeon P., Cuomo C. A., Wang S., *et al*. (2014) Using transcription of six *Puccinia triticina* races to identify the effective secretome during infection of wheat. Frontiers in plant science 4:520.
10. Choi J., Alkharouf N., Schneider K., Matthews B. and Frederick R. D. (2008) Expression patterns in soybean resistant to *Phakopsora pachyrhizi* reveal the importance of peroxidases and lipoxygenases. Functional & Integrative Genomics 8: 314-359.

11. Cipollini D. F., Bergelson J. (2000) Environmental and developmental regulation of trypsin inhibitor activity in *Brassica napus*. J Chem Ecol 26:1411–1422.
12. Cipollini D. F., Bergelson J. (2001) Plant density and nutrient availability constrain constitutive and wound-induced expression of trypsin inhibitors in *Brassica napus*. J Chem Ecol 27:593–610.
13. Cipollini D. F., Enright S., Traw M. B., Bergelson J. (2004) Salicyclic acid inhibits jasmonic acid-induced resistance of *Arabidopsis thaliana* to *Spodoptera exigua*. Mol Ecol 13:1643–1653.
14. Dangl J. L. and Jones J. D. G. (2001) Plant pathogens and integrated defence responses to infection. Nature 411: 826-833.
15. Delaney T. P., Uknes S., Vernooij B., & Friedrich L. (1994) A central role of salicylic acid in plant disease resistance. Science 266(5188):1247.
16. Djamei A., & Kahmann R. (2012) *Ustilago maydis*: dissecting the molecular interface between pathogen and plant.
17. Dubos C., Stracke R., Grotewold E., Weisshaar B., Martin C., & Lepiniec L. (2010) *MYB* transcription factors in Arabidopsis. Trends in plant science 15(10):573-581.
18. Duplessis S., Cuomo C. A., Lin Y. C., Aerts A., Tisserant E., Veneault-Fourrey C., ... & Martin F. (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proceedings of the National Academy of Sciences 108(22): 9166-9171.
19. Eddy S. R. (1998) Profile hidden Markov models. Bioinformatics 14:755-763.
20. Ellis J. G., Dodds P. N., and Lawrence G. J. (2007) Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. Annual Review of Phytopathology 45:289-306.

21. Eulgem T., Rushton, P. J., Robatzek S., & Somssich I. E. (2000) The *WRKY* superfamily of plant transcription factors. Trends in plant science *5*(5):199-206.
22. Garcia A., Sanches Calvo E., Souza Kiihl R., Harada A. (2008) Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: discovery of a novel locus and alleles. Theoretical and Applied Genetics 117:545-553.
23. Glazebrook J. (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu. Rev. Phytopathol. 43:189-215.
24. Grabherr M. G., Haas B. J., Yassour M., Levin J. Z., Thompson D., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B. W., Nusbaum C., Lindblad-Toh K.,Friedman N., Regev A. (2011) Full-length transcriptome assembly from RNA-seq data without a reference genome. Nature Biotechnology 29(7):644-52.
25. Grant M. R., Godiard L., Straube E., Ashfield T., Lewald J., Sattler A., ... & Dangl J. L. (1995) Structure of the Arabidopsis RPM1 gene enabling dual specificity disease resistance. Science 269(5225), 843-846.
26. Gu B., Kale S. D., Wang Q., Wang D., Pan Q., Cao H., Meng Y., Kang .Z, Tyler B. M., and Shan W. (2011) Rust secreted protein *Ps87* is conserved in diverse fungal pathogens and contains a RXLR-like motif sufficient for translocation into plant cells. PLoS ONE 6(11):e27217.
27. Guyon K., Balagué C., Roby D., & Raffaele S. (2014) Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. BMC genomics 15(1): 336.

28. Hall T. A. "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." Nucleic acids symposium series 41:1999.
29. Hartwig E. E. (1986) Identification of a fourth major gene conferring resistance to soybean rust. Crop Science 26:1135-1136.
30. Hartwig E. E. and Bromfield K. R. (1983) Relationships among three genes conferring specific resistance to soybean rust. Crop Science 23:237-239.
31. Hayden K. J., Garbelotto M., Knaus B. J., Cronn R. C., Rai H., Wright J.W. (2014) Dual RNA-seq of the plant pathogen *Phytophthora ramorum* and its tanoak host. Tree Genetics & Genomes 10.3:489-502.
32. Hernandez J. M., Heine G. F., Irani N. G., Feller A., Kim M. G., Matulnik T., ... & Grotewold E. (2004) Different mechanisms participate in the R-dependent activity of the R2R3 MYB transcription factor C1. Journal of Biological Chemistry, 279(46), 48205-48213.
33. Hogenhout S. A., Van der Hoorn R. A., Terauchi R., Kamoun S. (2009) Emerging concepts in effector biology of plant-associated organisms. Molecular Plant-Microbe Interactions 22:115-122.
34. Hok S., Attard A. and Keller H. (2010) Getting the Most from the Host: How pathogens force plants to cooperate in disease. Molecular Plant-Microbe Interactions 23:1253-1259.
35. Hunter S., Apweiler R., Attwood T. K., Bairoch A., Bateman A., *et al*. (2009) InterPro: the integrative protein signature database. Nucleic Acids Res 37: D211–215
36. Kale S. D., Gu B., Capelluto D. G. S., Dou D. L., Feldman E., Rumore A., Arredondo F. D., Hanlon R., Fudal I., Rouxel T., Lawrence C. B., Shan W., Tyler B. M. (2010)

External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. Cell 142:284-295.

37. Käll L., Krogh A., & Sonnhammer E. L. (2004) A combined transmembrane topology and signal peptide prediction method. Journal of molecular biology, 338(5), 1027-1036.
38. Kamoun S. (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. Annual Review of Phytopathology 44:41-60.
39. Kamoun S. (2007) Groovy times: filamentous pathogen effectors revealed. Current Opinion in Plant Biology 10(4):358-65.
40. Kanehisa M., Goto S. (2000) KEGG: Kyoto Encyclopaedia of Genes and Genomes. Nucleic Acids Research 28:27-30
41. Kang Y. J., Kim K. H., Shim S., Yoon M. Y., Sun S., Kim M. Y., Van K., Lee S. H. (2012) Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. BMC Plant Biology 12:139.
42. Kawahara Y., Oono Y., Kanamori H., Matsumoto T., Itoh T. *et al*. (2012) Simultaneous RNA-seq analysis of a mixed transcriptome of rice and blast fungus interaction. PLOS ONE 7: e49423. doi:10.1371/journal.pone.0049423.
43. Kemen E., Kemen A. C., Rafiqi M., Hempel U., Mendgen K., Hahn M., *et al*. (2005) Identification of a protein from rust fungi transferred from haustoria into infected plant cells. Molecular Plant-Microbe Interactions 18:1130-1139. doi: 10.1094/MPMI-18-1130
44. Kendrick M. D., Harris D. K., Ha B-K., Hyten D. L., Cregan P. B., Frederick R. D., Boerma H. R., Pedley K. F. (2011) Identification of a second Asian soybean rust resistance gene in Hyuuga soybean. Phytopathology 101:535–543

45. Langmead B., Trapnell C., Pop M., Salzberg S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology 10(3): R25

46. Li S., Smith J. R., Ray J. D., Frederick R. D. (2012) Identification of a new soybean rust resistance gene in PI 567102B. Theoretical and Applied Genetics 125:133-142.

47. Lovegrove F. E., Peña-Castillo L., Mohammad N., Liles W. C., Hughes T. R., & Kain K. C. (2006) Simultaneous host and parasite expression profiling identifies tissue-specific transcriptional programs associated with susceptibility or resistance to experimental cerebral malaria. BMC genomics, *7*(1):295.

48. Lum G., Min X. J. (2011) FunSecKB: the Fungal Secretome KnowledgeBase. Database - the Journal of Biological Databases and Curation. Database, 2011:bar001.

49. Mackey D., Holt B. F., Wiig A., & Dangl J. L. (2002) RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. Cell 108(6), 743-754.

50. McLean R. J., and Byth D. E. (1980) Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybeans. Australian Journal of Agricultural Research 31:951-956.

51. Meyer, J. D., Silva, D. C., Yang, C., Pedley, K. F., Zhang, C., van de Mortel, M., ... & Graham, M. A. (2009) Identification and analyses of candidate genes for Rpp4-mediated resistance to Asian soybean rust in soybean. Plant Physiology, 150(1), 295-307.

52. Meyers B. C., Dickerman A. W., Michelmore R. W., Sivaramakrishnan S., Sobral B. W. and Young N. D. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. The Plant Journal 20:317-332.

53. Miles M. R., Frederick R. D., and Hartman G. L. (2006) Evaluation of soybean germplasm for resistance to *Phakopsora pachyrhizi*. Plant Health Progress. 10.1094/PHP-2006-0104-01-RS.

54. Monteros M. J., Missaoui A. M., Phillips D. V., Walker D. R., Boerma H. R. (2007) Mapping and confirmation of the 'Hyuuga' redbrown lesion resistance gene for Asian soybean rust. Crop Science 47:829-836.

55. Motley S. T., Morrow B. J., Liu X., Dodge I. L., Vitiello A., Ward C. K., & Shaw K. J. (2004) Simultaneous analysis of host and pathogen interactions during an *in vivo* infection reveals local induction of host acute phase response proteins, a novel bacterial stress response, and evidence of a host-imposed metal ion limited environment. Cellular microbiology 6(9): 849-865.

56. Oosthuizen J. L., Gomez P., Ruan J., Hackett T. L., Moore M. M., Knight D. A., & Tebbutt S. J. (2011) Dual organism transcriptomics of airway epithelial cells interacting with conidia of *Aspergillus fumigatus*. PLoS One 6(5): e20527-e20527.

57. Pandey A. K., Yang C., Zhang C., Graham M. A., Horstman H. D., Lee Y & Whitham S. A. (2011) Functional analysis of the Asian soybean rust resistance pathway mediated by *Rpp2*. Molecular Plant-Microbe Interactions *24*(2) 194-206.

58. Panthee D. R., Yuan J. S., Wright D. L., Marois J. J., Mailhot D. and Stewart J. R. C. N. (2007) Gene expression analysis in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) in an early growth stage. Functional and Integrative Genomics 7:291-301.

59. Petre B., Joly D. L., & Duplessis S. (2014) Effector proteins of rust fungi. Frontiers in science 5.

60. Plett J. M., Kemppainen M., Kale S. D., Kohler A., Legue V., Brun A., Tyler B. M., Pardo A. G., Martin F. (2011) A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. Current Biology 21:1197-1203.
61. Posada-Buitrago M. L. and Frederick R. D. (2005) Expressed sequence tag analysis of the soybean rust pathogen *Phakopsora pachyrhizi*. Fungal Genetics and Biology 42:949-962.
62. Radwan O., Gandhi S., Heesacker A., Whitaker B., Taylor C., Plocik A., Kesseli R., Kozik A., Michelmore R.W., Knapp S. J. (2008) Genetic diversity and genomic distribution of homologs encoding NBS-LRR disease resistance proteins in sunflower. Molecular Genetics and Genomics 280:111-125.
63. Rafiqi M., Gan P. H. P., Ravensdale M., Lawrence G. J., Ellis J. G., Jones D. A., Hardham A. R., Dodds P. N. (2010) Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. Plant Cell 22:2017-2032.
64. Rappuoli R. (2000) Pushing the limits of cellular microbiology: microarrays to study bacteria-host cell intimate contacts. Proc. Natl Acad. Sci. USA 97, 13467–13469.
65. Ravensdale M., Nemri A., Thrall P. H., Ellis J. G., Dodds P. N. (2011) Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. Molecular Plant Pathology 12(1):93-102.
66. Schmutz J., Cannon S.B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D.L., Song Q., Thelen J.J., Cheng J., *et al.* (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178-183.

67. Schneider R. W., Hollier C. A., Whitam H. K., Palm M. E., McKemy J. M., Hernandez J. R., Levy L. and DeVries-Paterson R. (2005) First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. Plant Disease 89:774.

68. Seo H. S., Song J. T., Cheong J. J., Lee Y. H., Lee Y. W., Hwang I., Lee J. S., Choi Y. D. (2001) Jasmonic acid carboxyl methyl-transferase: a key enzyme for jasmonate-regulated plant responses. Proc. Natl. Acad. Sci. USA. 98. 4788-4793.

69. Tamura K., Peterson D., Peterson N., Stecher G., Nei M. *et al*. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739.

70. Thaler J. S., Fidantsef A. L., Duffey S. S., Bostock R. M. (1999) Tradeoffs in plant defense against pathogens and herbivores: a field demonstration of chemical elicitors of induced resistance. J Chem Ecol 25:1597-1609.

71. Tian M. Y., Win J., Savory E., Burkhardt A., Held M., Brandizzi F., Day B. (2011) 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. Molecular Plant-Microbe Interactions 24:543-553.

72. Trapnell C., Roberts A., Goff L., Pertea G., Kim D., Kelley D. R. & Pachter L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols *7*(3): 562-578.

73. Trapnell C., Williams B. A., Pertea G., Mortazavi A., Kwan G., van Baren M. J., Salzberg S. L., Wold B. J., Pachter L. (2010) Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28:511–515.

74. Tremblay A., Hosseini P., Alkharouf N. W., Li S., Matthews B. F. (2010) Transcriptome analysis of a compatible responses by *Glycine max* to *Phakopsora pachyrhizi* infection. Plant science 179:183-193.

75. Tremblay A., Hosseini P., Li S., Alkharouf N. W. and Matthews B. F. (2013) Analysis of *Phakopsora pachyrhizi* transcript abundance in critical pathways at four time-points during infection of a susceptible soybean cultivar using deep sequencing. BMC Genomics 14:614. doi: 10.1186/1471-2164-14-614

76. Tremblay A., Li S., Scheffler B. E. and Matthews B. F. (2009) Laser capture microdissection and expressed sequence tag analysis of uredinia formed by *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. Physiological and Molecular Plant Pathology 73:163-174.

77. Tyler B. M. (2009) Entering and breaking: virulence effector proteins of oomycete plant pathogens. Cellular Microbiology 11:13-20.

78. Van de Mortel M., Recknor J. C., Graham M. A., Nettleton D., Dittman J. D., Nelson R. T., Godoy C.V., Abdelnoor R. V., Almeida A. M. R., Baum T. J. and Whitham S. A. (2007) Distinct biphasic mRNA changes in response to Asian soybean rust infection. Molecular Plant-Microbe Interactions 20:887-899.

79. Vanderschuren H., Gruissem W., Bernal A., Staskawicz B. J. (2012) High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proceedings of the National Academy of Science USA 109(28):972-1979.

80. Walker D. R., Boerma H. R., Phillips D. V., Schneider R. W., Buckley J. B., Shipe E. R., Mueller J. D., Weaver D. B., Sikora E. J., Moore S. H., Hartman G. L., Miles M. R.,

Harris D. K., Wright D. L., Marois J. J., and Nelson R. L. (2011) Evaluation of USDA soybean germplasm accessions for resistance to soybean rust in the southern United States. Crop Science 51:678-693. doi:10.2135/cropsci2010.06.0340

81. Wasternack C. (2007) Jasmonates: an update on biosynthesis, signal transduction and action in plant stress response, growth and development. Ann. Bot. (Lond.), 100, 681–697.
82. Westermann A. J., Gorski S. A., Vogel J. (2012) Dual RNA-seq of pathogen and host. Nat Rev Microbiology 10: 618-630.
83. Zdobnov E. M., & Apweiler R. (2001) InterProScan-an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17(9):847-848.
84. Zhang X., Feng Y., Cheng H., Tian D., Yang S., Chen J. Q. (2011) Relative evolutionary rates of NBS-encoding genes revealed by soybean segmental duplication. Molecular Genetics and Genomics 285:79-90.
85. Zhou Q., Tian A., Zou H., Xie Z., Lei G., Huang J., Wang C., Wang H., Zhang J., Chen S. (2008): Soybean *WRKY*-type transcription factor genes, *GmWRKY*13, *GmWRKY*21, and *GmWRKY*54, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants. Plant biotechnology journal 6(5): 486-503.

A)

B)

Figure 3.1 Inoculation of Hyuuga (resistant line) and Williams 82 (susceptible line) with *P. pachyrhizi* 10 dpi (days post-inoculation). A) Hyuuga leaf inoculated with water (left) and *P. pachyrhizi* Expo isolate (right); B) Williams 82 leaf inoculated with water (left) and *P. pachyrhizi* Expo isolate (right).

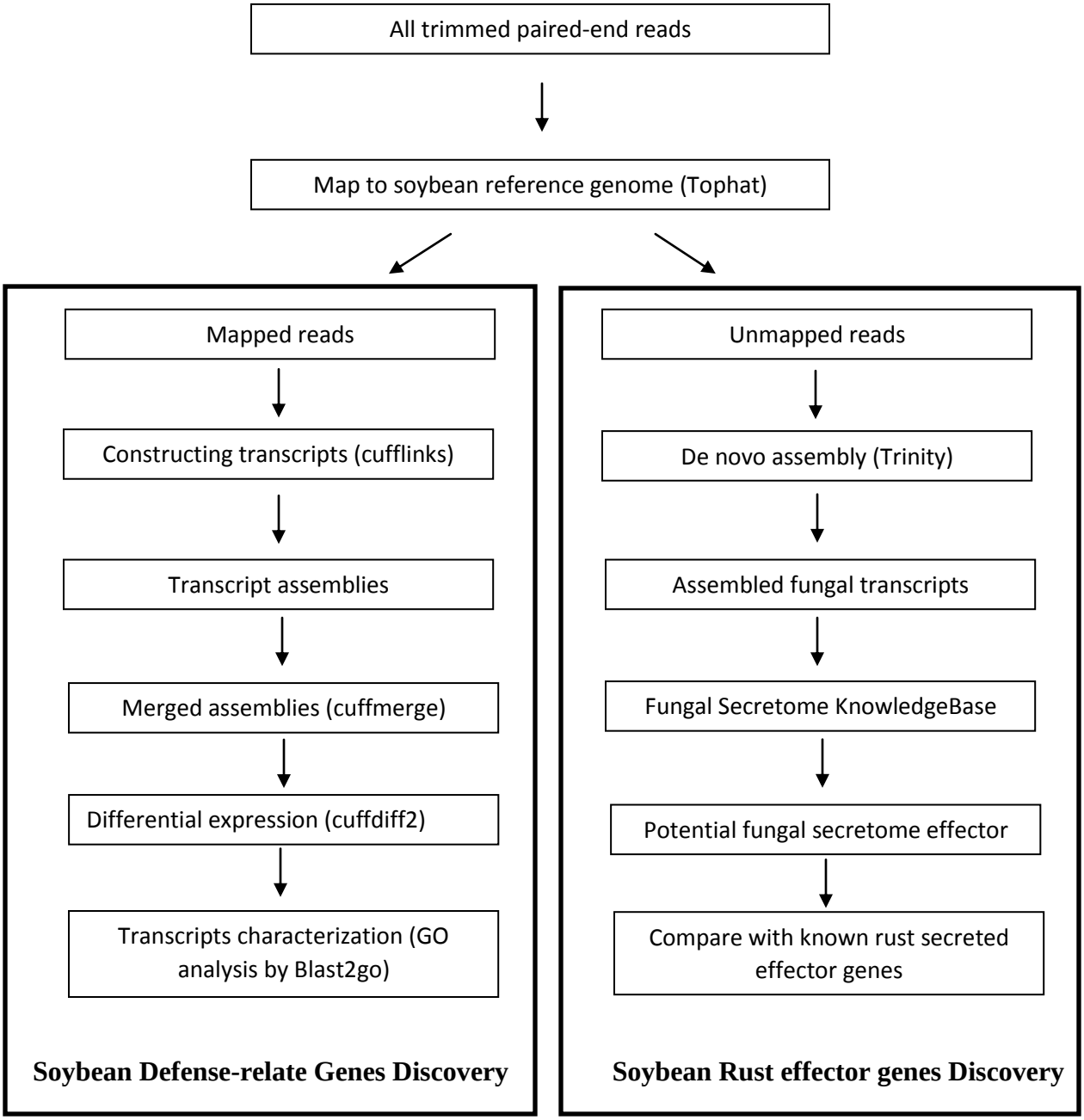


Figure 3.2 Flow chart of experimental procedure to identify potential soybean defense-related genes and Asian Soybean Rust effector genes. All the jobs were performed using zcluster queueing system by GACRC (The Georgia Advanced Computing Resource Center) at The University of Georgia at Athens, GA. The programs used for each step were shown in the parenthesis.

Table 3.1 Statistical results from RNA sequencing analysis (number and percentage of raw, filtered, and aligned reads).

Sample[a]	Number of raw reads[b]	Number of the trimmed reads[c]	Number of mapped reads[d]	Mapped reads (%)[e]
MS72-1	37328461	33845082	31509771	93.10%
MS72-2	30387897	25485166	23956056	94.00%
MS72-3	26946860	24228165	22507965	92.90%
MR72-1	25017567	22251989	20627594	92.70%
MR72-2	22070571	20000839	18720785	93.60%
MR72-3	35872515	31840285	29993548	94.20%
IS72-1	32914241	29787746	27940906	93.80%
IS72-2	27645382	24636831	22912253	93.00%
IS72-3	32670996	28611655	26837732	93.80%
IR72-1	28957036	26738926	25161329	94.10%
IR72-2	24969254	22250411	20715133	93.10%
IR72-3	35425577	32017486	30032402	93.80%

[a] MS: Reads from mock inoculated susceptible line, MR: Reads from mock inoculated resistant line, IS: Reads from *P. pachyrhizi*-inoculated susceptible line, IR: Reads from *P. pachyrhizi*-inoculated resistant line.
[b] The raw reads obtained from RNA-sequencing for three replicates in four treatments.
[c] The reads that were trimmed and filtered from raw sequencing reads by FASTX-Toolkit.
[d] The reads that were mapped to the *G. max* reference genome v.2.
[e] The ratio between the number of mapped reads versus the number of trimmed reads.

Table 3.2 Statistical results from RNA sequencing analysis for potential rust reads (number and percentage of raw, filtered, and aligned reads).

Sample[a]	Number of reads unmapped[b]	mapped reads (%)[c]	unmapped reads (%)[d]	after trimming and filtering
IS72-1	930516	1.20%	98.80%	923054
IS72-2	688677	1.30%	98.70%	682201
IS72-3	738546	1.30%	98.70%	732215
IR72-1	674202	1.10%	98.90%	668993
IR72-2	657072	1.10%	98.90%	651249
IR72-3	773552	1.40%	98.60%	766207

[a] IS: Reads from *P. pachyrhizi*-inoculated susceptible line, IR: Reads from *P. pachyrhizi*-inoculated resistant line.

[b] The reads that were not mapped to the *G. max* reference genome v.2.

[c] The reads that were not mapped to the *G. max* reference genome v.2. were re-mapped to *G. max* reference genome to remove all the plant-related reads. The ratio represents the proportion of potential soybean reads among the unmapped reads.

[d] The ratio represents the proportion of potential soybean rust reads among the unmapped reads.

Table 3.3 Statistical results of comparison of differentially expressed genes (DEGs) between inoculated resistant line (IR) vs mock resistant line (MR) and inoculated susceptible line (IS) vs mock susceptible line (MS).

Comparison	DEGs	Up-regulated	Down-regulated	Significantly Up-regulated	Significantly Down-regulated
IR vs MR	2759	1842	917	917	595
IS vs MS	3768	1613	2155	619	1064

The significantly up- or down-regulated DEGs listed have statistically significant change (fold > 2, p-value < 0.05) in response to the inoculation of *P. pachyrhizi*.

Table 3.4 Identification of all up-regulated and down-regulated DEGs in inoculated resistant line (IR) and inoculated susceptible line (IS).

	IR		IS	
	Up-regulated	Down-regulated	Up-regulated	Down-regulated
DEGs	917	595	619	1064
Defense-related [a]	127	63	82	84
Disease resistance	18	0	15	5
NBS-LRR R-genes[b]	14	0	3	2
Transcription factor	35 (15)[c]	40 (5)	46 (9)	48 (8)

[a] The defense-related transcripts were classified by Gene Ontology (GO) analysis performed by Blast2GO software using default parameters (Aparicio *et al*., 2006).

[b] NBS-LRR R-genes were characterized based on the structures of nucleotide binding site and leucine-rich repeats.

[c] The number in the parenthesis represents the number of transcription factor transcripts that were classified as defense-related transcripts based on GO analysis.

Table 3.5 NBS-LRR disease resistance genes identified in compatible (IS) and incompatible interaction (IR). Up- represents up-regulated genes and Down- represents down-regulated genes.

Gene ID [a]	Chr[b]	Chromosome location	Classification[c]	Predicted gene[d]	Log_2fold change
IR-up-Glyma.02G026200	2	Chr02:2338744..2341398	Non-TIR	RGA3-like	1.82457
IS-down-Glyma.02G030700	2	Chr02:2819404..2823496	Non-TIR	RGA3-like	-1.20089
IR-up-Glyma.02g188400	2	Chr02:35008500-35027979	Non-TIR	CC-NBS-LRR gene	1.01702
IR-up-Glyma.03G077400	3	Chr03:19180262..19187533	TIR	TMV N-like	1.14915
IR-up-Glyma.08G119200	8	Chr08:9192002..9196512	Non-TIR	RFL1-like	2.57829
IS-down-Glyma.09G075500	9	Chr09:8234968..8241071	TIR	SUPPRESSOR of npr1-1	-1.05991
IR-up-Glyma.12G011700	12	Chr12:853112..856650	Non-TIR	RPP13-like	1.3494
IS-up-Glyma.13G076200	13	Chr13:18016880..18022614	TIR	TMV N-like	1.05158
IS-up-Glyma.13G260600	13	Chr13:36467412..36479638	Non-TIR	At4g27220-like	1.17888
IR-up-Glyma.16G135200	16	Chr16:29255829..29266445	TIR	TMV N-like	1.51069
IR-up-Glyma.16g137000	16	Chr16:29418204-29487400	Non-TIR	TMV N-like	1.81624
IR-up-Glyma.16G137600	16	Chr16:29468438..29472890	TIR	TMV N-like	1.81624
IR-up-Glyma.16G033900	16	Chr16:3174939..3181746	TIR	TMV N-like	1.5409
IR-up-Glyma.16G215000	16	Chr16:37261497..37265361	TIR	LM6 precursor RNA	1.84041
IR-up-Glyma.16G215100	16	Chr16:37267759..37269360	TIR	KR11 pseudogene	1.14808
IR-up-Glyma.18G093800	18	Chr18:9412428-9581853	Non-TIR	RPM1-like	1.70219*
IR-up-Glyma.18g093300	18	Chr18:9412428-9581853	Non-TIR	RPM1-like	1.70219*
IR-up-Glyma.18g094100	18	Chr18:9412428-9581853	Non-TIR	RPM1-like	1.70219*
IS-up-Glyma.19G054700	19	Chr19:9126321..9131041	TIR	TMV N-like	1.20039

[a] The gene identification number is based on the name in Wm82.a2.v1.

[b] Chromosome number.

[c] The classification of R-genes was based on the GO analysis performed by Blast2GO software using default parameters (Aparicio *et al*., 2006).

[d] The prediction of the gene function were performed by searching the sequences against GeneBank by BLAST.

* One transcripts was mapped to three soybean genes located in the same region on the chromosome

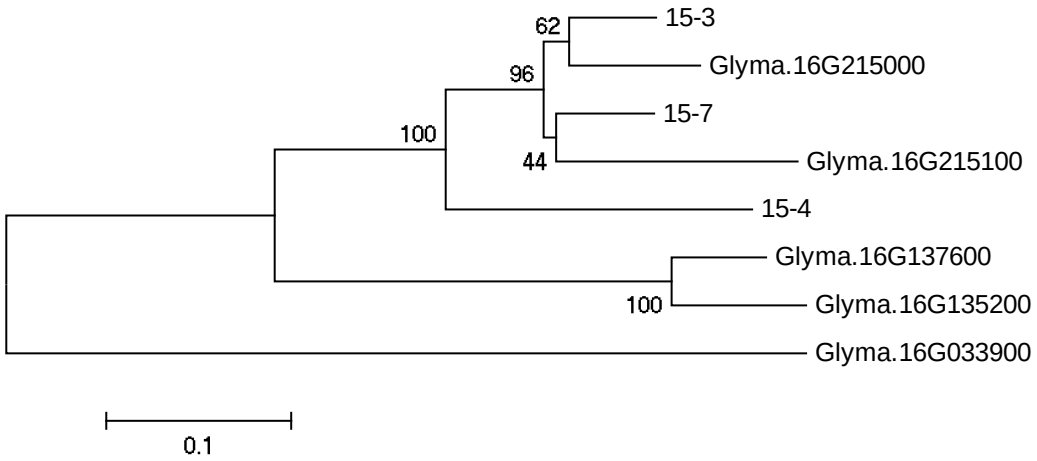


Figure 3.3 Phylogenetic tree of NBS regions of IR up-regulated NBS-LRR R-genes and resistance gene homologs (RGHs) characterized from fosmid analysis from Chapter 4. The phylogenetic tree was constructed using the Neighbor Joining (NJ) method, with distance represented as the number of nucleotide differences. Confidence in the phylogeny was assessed with one-thousand bootstrap replicates.

Table 3.6 Similarity between the amino acid sequences of NBS regions of the NBS-LRR R-genes using paisewise alignment by BioEdit.

	15-3	15-4	15-7	Glym. 16G215000	Glym. 16G215100	Glym. 16G137600	Glym. 16G135200	Glym. 16G033900
15-3		82.11 %	92.67 %	91%	64.92%	75.1%	73.69%	62.58%
15-4			83%	84.31%	57.43%	71.6%	71.94%	63%
15-7				91.85%	64.77%	76.71%	75.1%	62.52%
Glym. 16G215000					63.32%	76.91%	74.5%	64.21%
Glym. 16G215100						53.21%	52.72%	46.1%
Glym. 16G137600							91.3%	58.84%
Glym. 16G135200								58.23%
Glym. 16G033900								

Table 3.7 Identification of soybean transcription factors *WRKY* gene from up-regulated and down-regulated DEGs in IR and IS.

Treatment	Expression	Chr.	Gene ID (Wm82.a2.v1)	Name [a]	Groups [b]	Expression confirmed [c]	Soybase EST ID	Log_2_fold_change
IR	Up-regulated	4	Glyma.04G054200	*GmWRKY*79	IIc	+	FK004547.1	2.79519
		4	Glyma.04G218700	*GmWRKY*21	IIc	+	DQ322691.1	2.25495
		6	Glyma.06G142000	*GmWRKY*97	III	+	AK244967.1	2.57709
		9	Glyma.09G274000	*GmWRKY*125	III	+	BT098696.1	1.37766
		15	Glyma.15G139000	*GmWRKY*151	IIc	+	BW658350.1	2.76867
		17	Glyma.17G222300	*GmWRKY*30	IIa	+	EU019570.1	2.52933
		17	Glyma.17G224800	*GmWRKY*165	IIc	+	-	3.6644
	Down-regulated	3	Glyma.03G220800	*GmWRKY*51	IIc	+	BT098285.1	-1.05778
		8	Glyma.08G082400	*GmWRKY*114	IIc	+	-	-1.11011
		19	Glyma.19G217800	*GmWRKY*177	IIc	+	BI893911.1	-1.68901
IS	Up-regulated	2	Glyma.02G010900	*GmWRKY*67	IIc	+	BT096212.1	1.27641
		3	Glyma.03G220800	*GmWRKY*51	IIc	+	BT098285.1	1.40415
		4	Glyma.04G061400	*GmWRKY*81	IIa	+	EV275620.1	1.98211
		5	Glyma.05G127600	*GmWRKY*89	IIc	+	BW658787.1	1.18227
		5	Glyma.05G160800	*GmWRKY*90	IIe	+	GR846019.1	1.54869
		6	Glyma.06G307700	*GmWRKY*105	IIb	+	CO979705.1	1.71825
		8	Glyma.08G118200	*GmWRKY*48	IIe	+	GR844807.1	1.38437
		9	Glyma.09G080000	*GmWRKY*117	IIb	+	CF808085.1	2.54922
		9	Glyma.09G254800	*GmWRKY*124	IIe	+	BW658775.1	1.28832
		10	Glyma.10G011300	*GmWRKY*54	IIc	+	DQ322698.1	2.00395
		10	Glyma.10G138300	*GmWRKY*1	IIb	+	BI426286.1	2.40498
		18	Glyma.18G056600	*GmWRKY*62	I	+	AK245515.1	1.58592
		19	Glyma.19G217800	*GmWRKY*177	IIc	+	BI893911.1	2.32358
	Down-regulated	13	Glyma.13G370100	*GmWRKY*139	IIa	+	AK244287.1	-1.59524

[a] The names *GmWRKY*1-64 are given according to Zhou *et al* ; *GmWRKY*65-182 are given according to the chromosome order

[b]The classification according to Eulgem *et al*.

[c]The expression confirmation according to Soybase ESTs and RNA-seq analysis (in silico analysis) and RNA-seq of *P. pachyrhizi* lesion microdissection.

Table 3.8 Identification of soybean transcription factor *MYB* genes from up-regulated and down-regulated DEGs in IR and IS.

Treatment	Expression	Chr.	Gene ID (Wm82.a2.v1)	Gene name [b]	Groups[c]	AtMYB[d]	Function	Log_2_fold_ changee
IR	Up-regulated	5	Glyma.05G027000	MYB360	R2R3	AtMYB031 and AT5G56840	defense	1.58546
		8	Glyma.08G029400	MYB127	1R			2.34342
		10	Glyma.10G180800	MYB29	R2R3			2.09733
		20	Glyma.20G209700	MYB29B2	R2R3			1.69118
	Down-regulated	8	Glyma.08G059900	MYB3	R2R3			-1.31164
		13	Glyma.13G247200	MYB77	R2R3			-1.16375
IS	Up-regulated	5	Glyma.05G234600	MYB84	R2R3	Glyma08g04670.1	Metabolism[a]	2.94411
		8	Glyma.08G042100	MYB184	R2R3	Glyma08g04670.1	Metabolism[a]	1.78705
	Down-regulated	1	Glyma.01G003000	MYB128	1R	AT5G56840		-1.31446
		3	Glyma.03G225200	MYB86	R2R3			-1.18252
		5	Glyma.05G222600	MYB125	1R	AT5G56840		-1.77633

[a] Pandey *et al*, 2011

[b] The name of the myb genes were searched against NCBI non-redundant (nr) database using BLASTx

[c] The myb genes were classified according to Dubos *et al*., 2010

[d] The AtMYB were identified according to *Arabidopsis* genome database (https://www.arabidopsis.org/servlets/Tair)

IR indicates inoculated resistant. IS indicated incoculated susceptible.

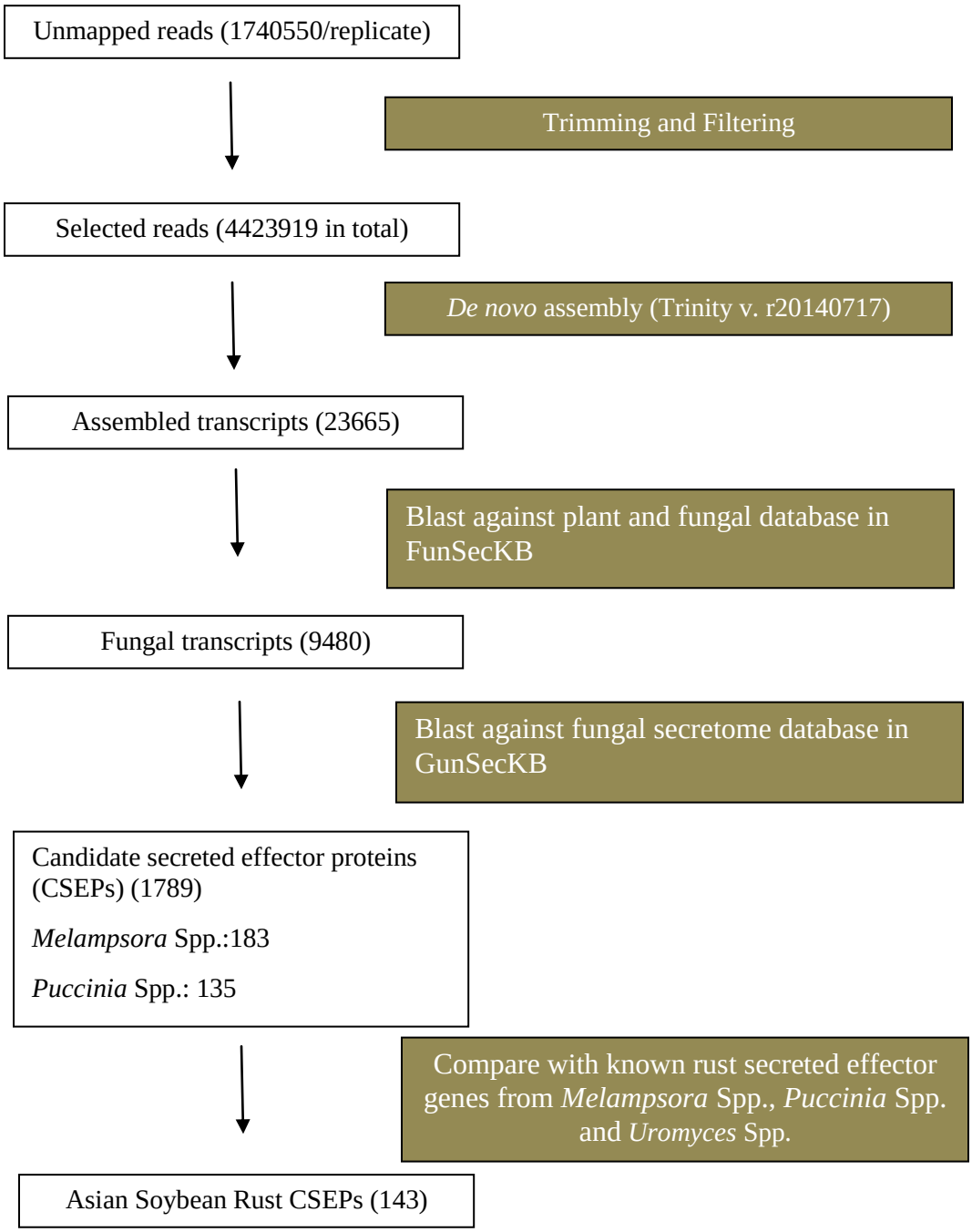


Figure 3.4 Schematic of bioinformatics pipeline used to identify *P. pachyrhizi* putative secreted effector genes. The mapping and *de novo* assembly were performed using zcluster queueing system by GACRC (The Georgia Advanced Computing Resource Center) in The University of Georgia at Athens, GA.

Table 3.9 Identification of *P. pachyrhizi* genes encode for candidate secreted effector proteins (CSEPs).

No.	Gene ID[a]	Blast result[b]	Corresponding rust known effector genes	Identity	E-value
1	c13500_g1_i2	pumilio domain-containing protein (sp: *rhizoctonia solani ag-1 ib*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_112338)	96.67	2.00E-06
2	c15620_g1_i1	Endonuclease-mitochondrial nuclease (sp: *neurospora tetrasperma*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_118194)	96.15	3.00E-05
3	c7460_g1_i1	sphingomyelin phosphodiesterase (sp: *coprinopsis cinerea*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_115419)	93.1	9.00E-05
4	c13765_g2_i1	heat shock protein ssa1 (sp: *candida albicans*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_90972)	90.32	3.00E-05
5	c12088_g1_i1	Putative uncharacterized protein (UM03217.1) (sp: *Ustilago maydis*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_89259)	88.24	4.00E-05
6	c68_g1_i1	peroxisomal matrix protein (sp: *colletotrichum orbiculare*)	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_116059)	87.88	8.00E-05
7	c13450_g1_i2	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_72205)	85.71	8.00E-07
8	c1092_g1_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_53368)	82.22	8.00E-06
9	c1581_g1_i1	secreted|secreted protein (sp: melampsora larici-populina)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_85020)	82.14	5.00E-09

10	c2736_g1_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_124586)	80.33	6.00E-08
11	c2680_g1_i1	gtpase rab11 ypt3 (sp: *aspergillus oryzae*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_112338)	79.59	5.00E-05
12	c10027_g2_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_87549)	79.01	2.00E-13
13	c11886_g1_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_88747)	78.85	9.00E-06
14	c12358_g1_i1	related to arl3-adp-ribosylation factor-like member of the arf-sar family in the ras	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_117298)	78.33	7.00E-06
15	c13658_g1_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_124586)	77.11	3.00E-11
16	c8662_g1_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_118194)	73.64	6.00E-13
17	c15207_g1_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_123928)	72.97	6.00E-32
18	c1382_g1_i1	secreted|secreted protein (sp: *melampsora larici-populina*)	*Melampsora larici-populina* 98AG31 secreted protein (MELLADRAFT_73322)	72.03	2.00E-11
19	c22424_g1_i1	ribonuclease t2-like protein (sp: *coccidioides posadasii*)	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_75620)	71.3	4.00E-09

20	c13200_g1_i1	expansin family protein (sp: *rhodosporidium toruloides* np11)	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_76387)	71.25	1.00E-12
21	c19334_g1_i1	gilt domain-containing protein (sp: *rhizoctonia solani* ag-1 ia)	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_124023)	71.22	1.00E-13
22	c13313_g1_i1	cell wall beta-glucan synthesis family protein (sp: *rhodosporidium toruloides*	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_72374)	70.71	3.00E-05
23	c13318_g1_i1	glycoprotein (sp: *coprinopsis cinerea*)	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_73714)	70.52	3.00E-88
24	c8943_g1_i1	macrofage activating glycoprotein (sp: *coprinopsis cinerea*)	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_124554)	70.5	1.00E-57
25	c5420_g1_i1	ferritin ribonucleotide reductase-like protein (sp: *rhodosporidium toruloides*	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_124085)	70.24	1.00E-50
26	c11267_g1_i1	sh3 domain-containing protein (sp: *rhizoctonia solani* ag-1 ia)	*Melampsora larici-populina* 98AG31 hypothetical protein (MELLADRAFT_72584)	70.05	3.00E-47

[a] The gene identification number was assigned by Trinity v. r20140717 during *de novo* assembly.
[b] The blast results were derived from searching the assembled soybean rust transcripts against the fungal secretome KnowledgeBase (FunSecKB, http://proteomics.ysu.edu/secretomes/fungi.php) using blastx with an E-value cutoff of $<1 \times 10^{-5}$.

CHAPTER 4

STRUCTURE AND EVOLUTION OF NBS-LRR RESISTANCE GENES IN SOYBEAN AND PERENNIAL *GLYCINE* SPECIES[1]

[1] P. Tian and S. M. Smith. To be submitted to Molecular Genetics and Genomics

Abstract

There are numerous disease resistance genes (R-genes) that have been shown to interact (directly or indirectly) with the pathogen effector protein, initiating host defenses. This triggers programed cell death which limits the colonization and development of the pathogen inside the plant tissue, resulting in resistance. The nucleotide-binding site (NBS)-Leucine-rich repeat (LRR) gene family accounts for the largest class of R-genes. To identify the soybean R-genes distribution pattern and characterize the structure of R-genes and their adjacent gene elements, a fosmid library that was developed from genomic DNA isolated from the Bingnan soybean cultivar (PI459025B) has been sequenced. Specific primers targeting conserved regions in the NBS-LRR R-genes of soybean were used to screen the fosmid library. Sixteen positive fosmid clones carrying potential resistance gene homologs (RGHs) were identified. Thirteen RGHs including NBS-LRR and kinase disease resistance genes were identified on the positive fosmids. Extensive structural analyses performed on each RGH in soybean genome allowed us to better understand the potential evolutionary mechanisms that result in such different pattern of R-genes distribution and structures. A pair of primers designed based on the alignment of soybean RGHs from fosmid analysis were used to amplify RGHs from eight soybean wild relatives genomic DNA in a PCR manner. A total 170 NBS and 119 LRR amino acid sequences were identified, which provided us the first report of molecular characterization of NBS-LRR-RGHs in the soybean wild relatives.

Introduction

To protect themselves against the attack of microbial pathogens, nematodes, and insects, plants have developed a complicated and efficient defense system. Significant efforts have been made in the past few years to better understand plant defenses by analyzing the evolution of disease resistance genes (R-genes) in various plant species. These analyses have begun to provide some insights into the source of R-gene specificity and the creation of new resistance specificities. Plant disease resistance often conforms to a gene-for-gene process (Dangl and Jones 2001) in which resistance reaction results from a molecular recognition (directly or indirectly) between the plant proteins encoded by resistance genes (R-genes) with their corresponding effector proteins encoded by pathogen avirulence (*Avr*) genes (Flor 1955). This R-gene and *Avr* gene protein recognition is typically followed by rapid activation of many defense responses and is most often characterized by a hypersensitive response (HR). HR is the death of local cells infected by the pathogen. Most characterized R-genes belong to one of five major classes of genes which include nucleotide binding site (NBS) leucine-rich repeat (LRR) proteins, kinases, receptor kinases, protein with transmembrane and internal CC, and extracellular LRR proteins with small intracellular domains (Hulbert *et al*., 2001; Shao *et al*., 2014).

Most of the work concerning the evolution of R-genes has focused primary on the NBS-LRR class because this is the largest class of resistance genes. Additionally, no other function, outside of disease resistance, has yet been associated with this class of genes. Conversely, the other types of disease resistance gene proteins are involved in many types of processes. The LRR domains in these proteins are typically predicted to be involved in protein-protein interactions and pathogen recognition specificity (Kobe and Deisenhofer 1995; Leister and Katagiri 2000; Dangl and Jones 2001). Other regions, including the NBS, have also shown to be important in a

few instances but this is not typical (Luck *et al*., 2000). The NBS domain is often associated with signaling. This domain contains several conserved motifs that are responsible for DNA binding and sending a downstream signal to activate plant defenses (Tameling *et al*., 2002). About three-fourths of the plant disease resistance genes that have been cloned to date are from the NBS-LRR class of resistance genes.

NBS-LRR genes are considered to be one of the most abundant gene families found in plants, with most plant genomes containing several hundred family members (Ribas *et al*., 2011). Previous diversity analysis demonstrated extremely high levels of inter and intraspecific variation of NBS-LRR genes, which presumably evolved rapidly in response to changes in pathogen populations (Meyers *et al*. 2003; Yang *et al*. 2008). To study the specific evolutionary routes of NBS-LRR encoding genes, comparative sequence analyses of R-gene clusters have been performed across haplotypes or related genomes between certain plant species and their closely related wild species, such as potato and wild potato (Kuang *et al*., 2005), wheat and *T. urartu* (Wicker *et al*., 2007), and cultivated rice and *O. rufipogon* (Seonghee *et al*., 2009).

To improve the disease resistance of economically important plants, many studies have focused on the isolation, cloning, and structural analysis of R-genes. To date, over 70 R-genes have been cloned from different plant species (Liu *et al*., 2007). The conservation of several motifs in the NBS domain has assisted in the design of degenerate primers to isolate Resistance Gene Homologs (RGHs) from different plant species utilizing a PCR based approach (Meyers *et al*., 2003; Ameline-Torregrosa *et al*., 2008). Comparative analysis of resistance gene homologues (RGHs) from heterologous species has provided evidence that R-genes are subject to positive selection, particularly in the LRR region. The LRR solvent exposed residues are predicted to interact either directly or indirectly with the corresponding *Avr* proteins in the pathogen and are

therefore under selective pressure (Kobe and Deisenhofer 1995). Sequence comparisons among several groups of *Arabidopsis* NBS-LRR gene family members demonstrated that selection has acted to diversify the LRR domain (Mondradon-Palamino *et al*., 2002). Moreover, analysis of tomato and *Arabidopsis* RGHs revealed remarkably rapid evolution of RGHs during radiation of plant families (Pan *et al*., 2000).

Comparative studies of R-genes from tomato, lettuce, rice, flax, and *Arabidopsis* have revealed that solvent-exposed positions of the LRRs are hypervariable (Mondradon-Palamino *et al*., 2002) and selective forces imposed by the pathogen incite allelic diversity (Hulbert *et al*., 2001). Therefore, the selective advantage of carrying an R-gene and the pressure imposed on the R-gene to diversify depends on the frequency of the corresponding *Avr* gene in the pathogen population. Investigation of resistance gene/RGH distribution and variation patterns have proven to be powerful tools used to estimate R-gene abundance and the type of selection acting in a plant species. It is therefore, practical to study the evolutionary changes of resistance genes/ RGHs in soybean and soybean wild relatives using comparative genomic analysis.

This study provides insight into the abundance, diversity, and distribution of resistance genes in soybean and its wild relatives, thus generating a resource for future use of this class of genes for improved soybean cultivar performance. The objectives of this study were to: 1) characterize the structure and evolutionary pattern of disease resistance genes in soybean, and 2) study the abundance and structure of NBS-LRR disease resistance genes in soybean wild relatives.

Materials and Methods

Plant Material and Genomic DNA Isolation

All of the plant introductions (PIs) were obtained from the USDA Soybean Germplasm Collection (Randy Nelson, Urbana, IL). Soybean cultivar Bingnan (PI 459025B) has not been sequenced and it confers resistance to the majority field isolates of Asian soybean rust, therefore it was selected for this work to construct a fosmid library and to analyze resistance gene homologues (RGHs). The soybean wild relatives selected for this work include: three *G. tomentella* accessions (PI 446995, PI 446988, and PI 483218), two *G. tabacina* accessions (PI 483199 and PI 339661), one *G. argyrea* accession (PI 505151), one *G. microphylla* accession (PI 373983A) and one *G. pescadrensis* accession (PI 440996). The wild relatives were selected based previous reports indicated that they were resistant to ASR (Hartman *et al*., 1992). Seeds for each soybean genotype were planted in a pot and maintained in a growth chamber (Conviron, Winnipeg, Canada) with day and night environments of 28/20°C temperature and 14/10 h of photoperiod, respectively. Leaves were collected from soybean plants three weeks following planting, while leaves were collected from the wild relatives eight weeks following planting. Genomic DNA was isolated from fresh leaf tissue collected from each cultivar using the CTAB method described by Murray and Thompson (1980).

Primer Design for Fosmid Library Screening

Seven combinations of degenerate primers were designed based on the alignment of NBS-LRR R-genes from soybean, rice, maize, and sorghum (Table 4.1). Four conserved motifs (P-loop, Kinase, MHD, and GLPL) within the NBS domain of resistance proteins were used for primer design. Due to the conservation of these motifs in many plant species, the four conserved NBS motifs have been used to design PCR-based cloning and mapping approaches to

characterize R-genes in many plant species (Meyers *et al*. 2003). The P-loop (GGVGKTT) and kinase motif within the NBS domain was targeted with three forward primers, while the GLPL (GLPLAL) and MHD motif were targeted with seven reverse primers (Table 4.1 and Figure 4.1).

The NBS domain of RGHs was PCR amplified from the genomic DNA of Bingnan soybean cultivar (PI 459025B) with seven combinations of the degenerate primers. PCR products ranged from 806bp to 1,023bp. PCR was performed in 50-mL reactions with 10 mM Tris-HCl, pH 8.8; 50 mM KCl; 1.522.5 mM MgCl2; 0.8 mM total deoxyribonucleotide triphosphates; 4U of High Fidelity Taq DNA polymerase; 25 pmol of a forward and reverse degenerate primer; and 50 ng of genomic DNA. An annealing temperature gradient was initially used that ranged from 50°C to 60°C to optimize the annealing temperature for different primer pair combinations. PCR reactions were performed on a Bio-Rad DNA Engine Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA). The resulting PCR products were isolated from a 1.0% agarose gel, purified with a Qiagen QIAquick Gel Extraction Kit (Qiagen, Redwood City, CA), and cloned into the Invitrogen pCR2.1-TOPO (Carlsbad, CA) cloning vector following the manufacturer's protocol. Two-hundred independent plasmids were selected randomly and sequenced per gel purification product using the Big-Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Grand Island, NY), following the manufacturer's protocol. All sequencing was performed at the Georgia Genomic Facility at the University of Georgia.

All of the sequences from the conserved motifs within the NBS domain were used as a query to verify putative homologies to known RGHs. Sequences with putative homologies to known RGHs were aligned using ClustalW (Larkin *et al*., 2007) and used to design eight pairs of conserved primers (Table 4.1) to screen a soybean fosmid library.

Fosmid Genomic Library Construction and Screening

Soybean genomic DNA isolated from the soybean cultivar Bingnan was used to construct a fosmid library. The genomic DNA was randomly sheared in a size range of ~35kb to ~45kb. The DNA fragments were end-repaired and ligated into the fosmid vector pCC1FOS according to the manufacturer's protocols (Epicentre Technologies, WI). Fosmid clones were packaged using MaxPlax Lambda Packaging Extract. Packaged fosmid clones were stored at 4°C over chloroform in 1 ml of Phage Dilution buffer. The library coverage is ~6 times the soybean genome size (~1115 Mbp) and is arranged into 34 superpools. Each superpool contains 48 pools.

The soybean fosmid library was screened with a PCR-based approach in four steps: 1) DNAs from the 34 superpools were amplified with conserved primers designed from the P-loop, Kinase, MHD, and GLPL motifs within the NBS domain; 2) Once a superpool was identified as positive in the first step, the 48 pools contained in the positive superpool were then amplified with the same primers to identify individual positive pools; 3) 96 randomly selected clones from each positive pool were amplified with the same primers; 4) PCR amplification products from any positive clones among the 96 individual clones were sequenced. BLASTN was then run against the NCBI non-redundant database to verify that the R-genes were amplified and cloned from soybean. The PCR fragments with homology to NBS-LRR R-genes were selected as soybean fosmids that carried putative R-genes. Positive fosmid clones with homology to NBS-LRR R-genes were then shotgun sequenced.

Fosmid Structure Annotation

The fosmids were annotated with the assistance of various gene finding programs. Each program has strengths and weakness. Therefore, combining the results generated from each program can improve the quality of an annotation prediction. The fosmid sequences were first

searched in Phytozome (http://www.phytozome.net/search.php) by utilizing the BLAST program with the target organism set as "*Glycine max* v1.1". This program can provide a prediction of the putative location of the fosmid insertion on the soybean chromosome. Fosmid structure was annotated with the assistance of FgeneSH (Salamov and Solovyev 2000). CDD (Conserved Domains Database) (Marchler-Bauer *et al*., 2011) was introduced to identify the conserved regions in the predicted protein sequences. A soybean transposable elements database (http://soybase.org/soytedb/) was used to search transposable elements among the predicted genes by running a BLAST program against *G. max* and *G. soja* EST (Expressed Sequences Tag), GSS (Genome Survey Sequence), and Core Nucleotide databases and the SoyBase Whole Genome Viewer. This analysis provides a way to see all of the results of a BLAST search simultaneously on the twenty soybean chromosomes.

The repeat sequences within the fosmids were identified by Tandem repeats finder (Benson 1999). A HMMER (biosequence analysis using profile hidden Markov models) search performed with a NB-ARC Pfam HMM PF00931 was used to identify sequences encoding an NBS domain and LRR domains were identified by LRR finder (Eddy 1998; Offord *et al*., 2010). TIR domain was identified by comparing with related sequences from Genebank and visual inspection. The conserved motifs in the N-terminal, NBS, and LRR domain were analyzed by the Multiple Expectation Maximization for Motif Elicitation (MEME) system 4.9.1 (http://meme.nbcr.net/meme/cgi-bin/meme.cgi). Each gene element on all the fosmids, including the NBS-LRR genes, genes with unknown functions as well as transposons were numerically assigned a gene ID number based on the order that they are arranged on each fosmid. MEGA (Molecular Evolutionary Genetics Analysis) program version 5 (Tamura *et al*., 2011) was used to align the NBS amino acid sequences in the predicted RGHs using ClustalW and construct

phylogenetic tree using the Neighbor Joining (NJ) method, with distance represented as the number of nucleotide differences. Confidence in the phylogeny was assessed with one-thousand bootstrap replicates.

Cloning and Sequence Analysis of RGHs from Eight Soybean Wild Relatives

A PCR-based approach was used to identify RGHs from eight soybean wild relatives (PI 446995; PI 446988; PI 483218; PI 483199; PI 339661; PI 505151; PI 373983A; PI 440996). NBS-LRR genes identified on positive soybean fosmid clones were aligned and used to design primers to amplify full-length RGHs from soybean wild relatives. The forward primer (P-Loop 2; 5'- TAGGWAARWCRACHMTTGC-3') targeted the P-Loop motif within the NBS domain, while the reverse primer (SRP7-4 5'-GCTTCTACATGAGGAAGTC-3') targeted the LRR domain. Genomic DNA isolated from the eight soybean wild relatives was PCR amplified with a conserved primer pair targeting the NBS and LRR domain. PCR amplification and cloning of the resulting 2 kb fragment was performed as described previously for the fosmid clones. Ninety-six clones were sequenced for each wild relative per gel purification product using the Big-Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Grand Island, NY), following the manufacturers protocol. Two universal primers (M13 forward and M13 reverse) were used to sequence the clones, performed by Georgia Genetic Facility at The University of Georgia.

BLASTN was run against the NCBI non-redundant database to verify that NBS-LRR genes were cloned. Sequences from the wild relatives with putative homologies to known NBS-LRR genes were aligned using ClustalW (Larkin *et al*. 2007). Duplicated sequences identified in each wild relative were removed from the analysis after the alignment. The NBS and LRR sequences were analyzed as separate domains. They were first translated into amino acid sequences and subjected to protein domain and motif analysis. Sequences with an uninterrupted

open reading frame were further analyzed. MEGA (Tamura *et al*., 2011) was used to align the NBS and the LRR amino acid sequences separately and construct a phylogenetic tree using the Neighbor Joining (NJ) method, with distance represented as the number of nucleotide differences. Confidence in the phylogeny was assessed with one-thousand bootstrap replicates.

Results

Structure and Clustering of Cloned Soybean RGHs

A total number of sixteen soybean fosmids were selected from many hundreds of PCR positive clones amplified with RGH-targeting primers. The RGH primers targeted the P-loop, Kinase, MHD, and GLPL motifs within the NBS domain (Table 4.1 and Figure 4.1). The nucleotide sequences in each fosmid were annotated by FgeneSH. Fosmid 9 was shown to be a vector fragment and fosmid 13 failed to be sequenced, therefore fourteen fosmids were selected for further analysis (Figure 4.2 and Table 4.2). A total of 61 genes and 29 transposable elements were predicted to be in the fourteen sequenced fosmids (Table 4.2). The fosmid insert sizes were found to be 40 kb or less, and inserts exhibited a combined GC content of 44.9%. The size of the selected fosmids ranged from 2362 bp (Fosmid 12) to 35240 bp (Fosmid 6). Each fosmid carries from one to eleven genes ranging in length from 78 to13788 bp. Four partial (truncated) or intact protein kinase genes, nine partial or intact genes with homolog to NBS-LRR resistance genes, and twenty-nine transposable element genes were identified on all the positive fosmids (Table 4.2). Genes with very low similarity to known genes in the database or were characterized as soybean genomic clones were predicted to be genes with an unknown function (Figure 4.2). Genes other than resistance genes were also identified in the fosmids. Fifteen predicted gene elements were identified with homolog to functional proteins such as mitochondrial RNA-splicing protein, ribosome biogenesis protein, transcription factor, centromere protein, or enzyme

with certain function. In addition, several predicted protein kinase genes with potential signaling functions were identified and some kinase genes were clustered together in fosmid 2 and 8 while a single protein kinase was predicted on fosmid 6 and 11 (Figure 4.2).

The HMMER search identified eight putative NBS-LRR genes in fosmid 7, 14, 15, and 16 (Figure 4.2). Five of the identified genes are partial and three are intact R-genes. All of the R-genes are orientated in the same direction with the exception of RGH F16-5 (The No. 5 gene in fosmid 16). By comparison of the RGH amino acid sequences, several conserved domains were identified, including the TIR, NBS, and LRR domains (Figure 4.3). Interestingly, one truncated R-gene identified in fosmid 12 is a truncated version of the TIR-NBS-peptidase gene that was predicted by the CDD program to carry a conserved plant basic secretory protein (BSP) region (Peptidase of plants and bacteria). This gene is predicted to be involved in plant defense mechanism against many different pathogens (Kuwabara *et al*., 1999). In addition, there are no other R-genes identified with homologs to Non-TIR-NBS-LRR R-genes.

In order to understand the fine structure of the NBS-LRR R-proteins, the MEME program was used to compare three soybean reference R-genes (N, SR1, and KR1) with the RGH amino acid sequences from the soybean positive fosmids. There was high amino acid conservation observed in TIR and NBS domains and a relatively low level of conservation in LRR domains (Figure 4.4; Table 4.3; Table 4.4). The TIR domain was conserved in all of the compared RGH proteins identified in the fosmids. TIR1, TIR2, TIR3, and TIR4 were identified in the RGH proteins (Table 4.3). They share homology with the reference resistance genes and the order of TIR motifs (1-4) are well conserved in the soybean RGHs (Meyers *et al*., 2003). Adjacent to the TIR domain, a TN (TIR-NBS) linker and eight motifs (Walker-A/P-loop, RNBS-A-TIR, Walker B/Kinase-2, RNBS-B/ Kinase-3a, RNBS-C, RNBS-D, GLPL and MHD) were identified

and are identical to the motifs identified in the NBS region of known NBS-encoding genes in most plant species (Hayes *et al*., 2000; Graham *et al*., 2000) (Table 4.4). There are ~65 amino acids between the NBS and LRR domains. This region has two motifs (PGKRSRαxxxEDxxαLxxxT and αxxαxxLRxα) that are conserved in the soybean RGHs and they were therefore designated as the NL (NBS-LRR) linkers.

Eleven LRR motifs (L1-L11) were identified in the soybean RGHs (Figure 4.4). L1 is the most prevalent LRR domain with a copy number ranging from 4 to 7 in the soybean RGHs. In the C-terminal region of the soybean RGHs, two potential transmembrane segments (L8 and L9) were identified. TRANS1 (L8) was identified in four RGHs (F15-1, 15-3, 15-4, and 15-7) and TRANS2 (L9) was found in one RGH (F7-4). RGH F14-6 contains a TRANS1 and TRANS2 domain similar to the SR1 reference R-gene.

Identification of Kinase Genes on Soybean fosmids

Several kinases were predicted to be in the soybean fosmids. The kinase genes are either maintained in a cluster (fosmid 2 and 8), or as a single gene (fosmid 6 and 11) (Figure 4.2). In addition, one RGH (F2-6) was predicted to be a *G. max* transcription factor TCP2-like protein. This protein was reported to be associated with regulation of LIPOXYGENASE2 (LOX2) expression and jasmonate (JA) production, which are involved in the *Arabidopsis* defense response to *M. quadrilineatus* (Sugio *et al*., 2011).

Two RGHs (F6-3 and F11-1) on fosmid 6 and 11, respectively, were identified as a putative serine/threonine-protein kinase which can phosphorylate serine and threonine residues and interact with other proteins to effect a wide array of processes ranging from disease resistance and developmental regulation to self- versus non-self-recognition (Afzal *et al*., 2008). One RGH (F7-1) on fosmid 7 was predicted to be a *G. max* heparanase-like protein 1-like.

Haparanase is indirectly involved in H_2O_2 degradation, and generates phenolic compounds that may be used for cell wall fortification (Eckey *et al*., 2004). One RGH (F10-8) was predicted to be a *G. max* kinesin-like calmodulin-binding protein homolog (KCBP) that is responsible for Ca^{2+}/calmodulin regulation of its ATPase activity, interaction with MTs, and motor activity (Abdel-Ghany *et al*., 2005). Interestingly, one full-length putative protein kinase (F16-4) was predicted to be a *G. max* probable ferric reductase transmembrane component (NADPH oxidase). The activation of NADPH can result in reactive oxygen species (ROS) production and program cell death (Lherminier *et al*., 2009). A phytozome search also showed that one RGH (F8-6) carries a 300aa protein kinase domain (Pfam:00069) and is characterized as a full length protein kinase. RGH F8-7 was not predicted to be a kinase. It does however have homology to a *G. max* putative 1-aminocyclopropane-1-carboxylate deaminase-like which is shown to play an important role in plant defense system (Gonzales-Vigil *et al*, 2011). Similarly, RGH F1-3 carries a Ras-GTPase Activating Domain that plays an important role in mediating plant defense responses (Cheung *et al*., 2008).

Transposable Elements on the Soybean Fosmids

The transposable elements identified on the soybean fosmids fall into two groups, LTR (long terminal repeat) retrotransposons or DNA transposons (Figure 4.2). Retrotransposons can be further divided into two subgroups, Ty1-copia and Ty3-gypsy, and are very common in *G. max*. There were seven LTR Copia retrotransposons predicted to be on six fosmids (1, 2, 4, 7, 10, and 14) while seventeen LTR gypsy retrotransposons were predicted in eight fosmids (1, 2, 3, 4, 5, 11, 14, and 15). DNA transposons were less prevalent than LTR retrotransposons, as they were only predicted to be on four fosmids (4, 6, 10, and 11). No long Interspersed Elements (LINE-1 or L1) or Short Interspersed Elements (SINEs) were found in any of the soybean

fosmids. The simple repeat and tandem repeat sequences were identified using tandem repeats finder. Twelve simple repeat sequences with copy numbers ranging from 12.7 to 28.5 were identified, while a hundred tandem repeats were revealed with copy numbers ranging from 1.9 to 10 and unit length varying from 5 to 221 bp.

Gene Duplication and Phylogeny of RGHs in Fosmid 15

Fosmid 15 carried the most number of RGHs and was predicted to be a segment of Chromosome 16 from soybean that has over 40 NBS-LRR genes (Kang *et al*., 2012). Five RGHs were predicted to be in fosmid 15 and are clustered together, therefore may be the result of gene duplication (Figure 4.2; 4.3). Based on phylogenetic analysis of the TIR (Figure 4.5A), NBS (Figure 4.5B), and LRR region, (Figure 4.5C) of the RGH amino acid sequences, the majority of the RGHs identified in fosmid 15 were grouped together on the same major clade (Figure 4.5). Both the NBS and LRR regions from the F14-6 RGH clustered with the *KR1 G. max* resistance gene. The LRR region of the F16-5 RGH clustered with the *Medicago truncatula* R-gene *GS4-5*, but the TIR and NBS region of this gene did not cluster with any of other identified RGHs. Three RGHs consistently (F15-3, 4 and 7) are grouped together in the same clade. Two soybean RGH from fosmid 15 (F15-1 and F15-8) were not full-length, therefore the TIR and NBS region of F15-1 and the LRR region of F15-8 could not be compared with other RGHs.

The nucleotide sequence identity for the RGHs ranged from 70.2 to 75%, while amino acid similarity ranged from 73.6-83.2%. The LRR copy number and distribution pattern were compared among the RGHs based on their predicted structure (Figure 4.4). After the NL linker region, two RGHs (F15-1 and F15-3) have two L1 LRR motifs, one RGH (F15-4) has one, and one RGH (F15-7) has four L1 motifs. Interestingly, the pattern of LRR motifs between the L1 motifs and transmembrane motifs (L6 to L9) are very similar.

Structure of Fosmid 16 with Different Elements

Fosmid 16 is the most interesting fosmid analyzed because it carries seven gene elements that are correspond to genes with a known function with the exception of RGH F16-2 (Figure 4.2). RGH F16-5 was predicted to be a full-length TIR-NBS-LRR R-gene that is corresponding to the *G. max* TMV R-protein N-like (LOC100815486) transcript variant, X1. Interestingly, RGH F16-6 and F16-7 are adjacent to RGH F16-5, oriented in the same direction and are also predicted to be an X1 gene. F16-4 RGH is oriented in the opposite direction of the F16-5 RGH and is predicted to be a *G. max* ferric reduction oxidase 2-like (LOC100790114) transcript variant 1 with homology to *M. truncatula* NADPH oxidase (MTR_7g038480The F16-1 RGH corresponds to a partial glycosyltransferase gene in *A. thaliana*. Previous reports suggested that some glycosyltransferase genes are necessary for resistance to *Pseudomonas syringae* pv. *tomato* in Arabidopsis (Langlois-Meurinne *et al*., 2005). The F16-3 RGH was predicted to be similar to *G. max* UDP-glucose flavonoid 3-O-glucosyltransferase gene, which is essential for *A. thaliana* nonhost resistance to Asian Soybean Rust (Langenbach *et al*., 2013). F16-2 is an uncharacterized protein but it has some similarity to the HXXXD-type acyl-transferase family protein of *Theobroma cacao*. There is a Serine Carboxypeptidase-Like Acyltransferase that has been shown to be critical for synthesis of antimicrobial compounds and disease resistance in oats (Mugford *et al*., 2009).

Structure and Evolution of RGHs Cloned from Soybean Wild Relatives

A total of 329 NBS and 375 LRR nucleotide sequences and 170 NBS and 119 LRR amino acid sequences were identified in the eight soybean wild relatives (Table 4.5). The NBS amino acid sequences were further annotated and predicted that most of the RGHs contain conserved motifs (P-Loop, RNBS-A, Walker B-Kinase-2, RNBS-B, RNBS-C, GLPL, and partial

RNBS-D). The LRR region of the RGHs was shown to start at a region close to NL-linker and spanned bout 200 amino acids. Due to the variability of the last residue of the Walker-B motif in the NBS region of the RGHs, the NBS sequences were separated into two major classes. One class encoded 116 TIR-NBS sequences with an aspartic amino acid at last residue, while 54 Non-TIR NBS sequences have either a tryptophan, cysteine, or serine amino acid at that same position (Table 4.5).

To assess the sequence diversity and relationships between the NBS and LRR region in the wild relatives, the NBS and LRR amino acid sequences characterized from each soybean wild relatives were used to generate a phylogenetic tree (Figure 4.6 A and B). Sixteen phylogenetic trees were generated and the structures of all them are very similar with each other. We selected PI 505151 because its structure fully represents all the phylogenetic trees of other soybean wild relatives. Twenty-five NBS and eighteen LRR amino acid sequences identified in PI 505151 were aligned to generate two phylogenetic trees. The patterns of two phylogenetic trees were similar, as both of trees had two major clades. Interestingly, the four RGH clones (FB02, FH05, FB04, and FC11) in the second clade of the NBS phylogenetic tree (Figure 4.6A) were identified as Non-TIR-NBS sequences, while all of the other RGHs in the first clade belong to TIR-NBS class. In addition, the corresponding LRR amino acid RGH sequences (RB02, RH05, RB04, and RC11) are also clustered together in the same clade.

Discussion

Much of the work concerning the evolution of R-genes in recent years has focused on the NBS-LRR class because this is the largest class of resistance genes. Furthermore, no other function (besides disease resistance) has yet been associated with this class of genes while the other types of proteins are involved in many types of processes. The LRR domains in these

proteins are typically the motifs predicted to be involved in ligand binding and determination of specificity, although other regions have also been demonstrated to be important (Luck *et al*., 2000).

At present, the resources available for high-resolution genetic and physical mapping in soybean have dramatically improved over the last few years. There has also been an increase in the number of R-genes identified and characterized from soybean. Despite these efforts, little is known regarding the detailed structure of the R-genes in soybean, how the resistance genes evolve and the structure of the regions adjacent to the R-genes. In addition, many of the perennial *Glycine* wild species (wild relatives of soybean) have been screened for soybean disease resistance but many questions still remain regarding the evolution of R-genes in wild species as well. To our knowledge, there is no report of cloning RGHs from different *Glycine* perennial wild species. Continued characterization of R-genes in soybean and its wild relatives should allow us to establish clearer trends in how resistance gene loci evolve and thus allow us to utilize them more efficiently in crop improvement programs and may ultimately allow us to design them for in vitro construction.

In this study, we screened a fosmid library containing the genomic DNA of the soybean cultivar Bingnan to identify and characterize soybean resistance genes/resistance gene homologues (RGHs). The average insert size for this library was found to be ~40 kb, ranging from ~35kb to ~45kb, so the library represents approximately 6X coverage of the soybean genome. A library of this size guarantees a 99.5% probability to find one specific fosmid clone when screening for specific sequences within the genome. The screening identified 14 positive clones (fosmids) that carry a total of 13 RGHs including NBS-LRR and kinase disease resistance genes. Extensive structure analyses of each fosmid enabled us to assess the distribution and

evolution of the RGHs as well as assist with cloning the corresponding RGHs from soybean perennial wild species.

Fourteen soybean fosmids carried full-length and truncated RGHs and five of the fourteen fosmids carried more than one RGH, often in tandem arrays as seen in other species. Five NBS-LRR resistance genes were identified in fosmid 15. This fosmid corresponds to chromosome 16 in the soybean genome. Interestingly, we also performed RNA sequencing on a soybean cultivar that was resistant to ASR isolates collected from Georgia (Chapter 3). Fourteen NBS-LRR resistance genes were up-regulated in the resistant cultivar and five of these genes were also located on chromosome 16 in soybean genome. This indicated that similar to many disease resistance genes, few RGHs are maintained as a single gene, while most soybean RGHs are maintained in linked clusters. Additionally, the tandem RGHs were also similar in sequence suggesting that the linked clusters were potentially derived from duplication and/or recombination events. The truncated RGHs were typically missing part of the LRR region. This suggests that the RGHs likely participated in an unequal recombination with another NBS-LRR gene, resulting in the loss of the C-terminal region. Previous analysis of RGHs has demonstrated that plant NBS-LRR disease resistance genes frequently occur in tightly linked clusters and undergo unequal recombination (Pryor 1987; Michelmore and Meyers 1998). This clustering feature is considered to facilitate the expansion of R-gene numbers and race-specificities through duplication, recombination, and positive selection (Michelmore and Meyers 1998).

RGHs that were truncated or contain a stop codon were also identified and were considered to be pseudogenes. Pseudogenes are still considered to be an important component of the evolutionary process of disease resistance genes because that can participate in recombination events with functional R-genes and potentially create new resistance gene that are

effective against the current pathogen population. Truncated versions of RGHs have been identified in soybean and these truncated products have been presumed to have an important role in disease resistance (Graham *et al*., 2000).

All of the NBS-LRR genes identified in the soybean fosmids were from the TIR-NBS-LRR class with the exception of one RGH (F12-1). In most plant species, the CC-NBS-LRR class of resistance genes is predominate. However, this has not been the case for the leguminous species where, the TIR-NBS-LRR genes are typically predominate (Meyers *et al*., 2003). This research suggests that the TIR-NBS-LRR class of resistance genes is the predominate class in soybean and the non-TIR (CC-NBS-LRR) class is a minor class. We anticipate that the use of primers designed to specifically target the CC-NBS-LRR class would identify soybean fosmids with genes from the non-TIR class.

The detailed structure of the RGHs was also analyzed. The TIR and the NBS domain were conserved in all of the NBS-LRR RGHs. Interestingly, the motifs within the NBS region were also highly conserved. The NBS region is predicted to be responsible for sending the signal after the presence of the pathogen has been detected (Dangl and Jones 2001). This suggests that the signaling component of the R-gene-Avr gene interaction is conserved. This is likely to avoid inappropriate signaling events and to maintain the efficiency of the response when plant defense is activated. Conversely, variation was detected within the LRR region of the RGHs. RGHs differed greatly in the number of LRR motifs ranging from 1 to 11 (L1-L11). The L1 motif is the most prevalent LRR domain ranging from four to seven repeats in the soybean RGHs. Many factors can affect the variation in the number of LRR motifs identified in the RGHs including, recombination and pathogen populations that are specific to each R-gene. The LRR regions has been shown to interact with the Avr protein, thus controls specificity (Martin *et al*., 2003).

Therefore, the variation observed in the LRR region of the soybean RGHs may be due to adaptive plasticity for disease resistance and responses to other environmental variables (Clay and Kover 1996).

Many transposable elements were identified in close proximity to RGHs in the soybean fosmids. Plant transposable elements can play very important roles in the evolution of resistance genes, primarily because their insertion near a gene can create novelty in transcriptional regulation, especially epigenetic regulation (Fedoroff 2000; Lisch and Bennetzen 2011). Transposons also can internally capture genes, can serve as sites of homology for unequal crossing-over, or can initiate a chromosome breakage-repair cycle (Purugganan and Wessler 1995; Michelmore and Meyers 1998; Bennetzen 2005). All of these processes can amplify R-genes, and can also move them to new genomic locations. For example, the F7-4 RGH is a truncated gene and it is in close proximity to a transposon. Additionally, some regions in the fosmids were unable to be assembled due to the repetitive nature of that region. Genomes of higher plants vary significantly in their size and complexity. Soybean has been shown to undergo two duplications. Repetitive DNA sequences have been shown to be the major determinant of genome sizes in higher plants. It has been hypothesized that transposable elements play a major role in the expansion and diversification of transmembrane receptor kinase type disease resistance *Xa21* gene family (Song *et al*, 1997).

Crop breeders have exploited wild relatives for resistance to diseases for over a century and they continue to search extended gene pools for genes that convey resistance to major crop pests and diseases (Prescott-Allen 1986). Currently, it has been reported that over 100 beneficial traits have been derived from approximately 60 wild species and incorporated in 13 cultivated crops and out of these 100 beneficial traits over 80% were associated with pest and disease

resistance (Hajjar and Hodgkin 2007). In this study, we selected eight soybean wild relatives that conferred a high level of disease resistance against the current Asian soybean rust population in Georgia. They likely carry interesting R-genes that can be potentially introduced into future breeding programs to control not only ASR but also powdery mildew.

Over 170 NBS and 119 LRR amino acid sequences identified in the eight soybean wild relatives were subjected to structural analyses and phylogenetic analysis. TIR-NBS-LRR R-genes were more abundant than the Non-TIR-NBS-LRR class of R-genes, demonstrating that the TIR-NBS-LRR genes are typically the predominate type of R-genes in leguminous species. In addition, the NBS domain is conserved and the TIR-NBS-LRR can be distinguished with Non-TIR type of R-genes based on the Walker B (Kinased 2) motif of NBS domain. However, although the LRR regions typically divergent, phylogenetic analysis revealed that the LRR region of these two types of R-genes conserved. This indicated that both the NBS domain and LRR region in the two types of R-gene are well maintained in the process of R-gene evolution. This also suggested that less selection pressure is imposed on the R-genes in the wild relatives to diversify and change resulting in less diversity in the LRR region of these R-genes.

In this study, extensive structural analyses of RGHs of eight soybean wild relatives provide fundamental knowledge of the structure, classification, and evolution of disease resistance genes. With the phenotypic data associated with resistance against the current ASR population in Georgia presented in chapter 2, the wild relatives that confer a high level of resistance are considered as candidate lines to study the durability of resistance to ASR. Therefore expression analysis and further function analysis will be essential to dissect the host-pathogen interaction.

Literature cited

1. Abdel-Ghany S. E., Day I. S., Simmons M. P., Kugrens P. and Reddy A. S. (2005) Origin and evolution of kinesin-like calmodulin-binding protein. Plant Physiology 138:1711-1722.
2. Afzal A. J., Wood A. J., Lightfoot D. A. (2008) Plant receptor-like serine threonine kinases: roles in signaling and plant defense. Molecular Plant-Microbe Interactions Journal 21:507-517.
3. Ameline-Torregrosa C., Wang B. B., O'bleness M. S., Deshpande S., Zhu H. et al. (2008) Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant Medicago truncatula. Plant Physiology 146:5-21.
4. Bennetzen J. L. (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. Current opinion in genetics & development, 15(6): 621-627.
5. Benson G. (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research 27:573-580.
6. Cheung M. Y., Zeng N. Y., Tong S. W., Li W. Y., Xue Y., Zhao K. J., Wang C., Zhang Q., Fu Y., Sun Z., Sun S. S., Lam H. M. (2008) Constitutive expression of a rice GTPase-activating protein induces defense responses. New Phytologist 179:530-545.
7. Clay K., & Kover P. (1996) Evolution and stasis in plant-pathogen associations. Ecology 997-1003.
8. Dangl J. L. and Jones J. D. G. (2001) Plant pathogens and integrated defense responses to infection. Nature 411: 826-833.

9. Eckey C., Korell M., Leib K., Biedenkopf D., Jansen C., Langen G., Kogel K-H. (2004) Identification of powdery mildew-induced barley genes by cDNA-AFLP: functional assessment of an early expressed MAP kinase. Plant Molecular Biology 55:1-15.
10. Eddy S. R. (1998) Profile hidden Markov models. Bioinformatics 14:755-763.
11. Flor H. H. (1955) Host-parasite interaction in flax rust-its genetics and other implications. Phytopathology 45:680-685.
12. Gonzales-Vigil E., Bianchetti C. M., Phillips G. N. Jr, Howe G. A. (2011) Adaptive evolution of threonine deaminase in plant defense against insect herbivores. Proceedings of the National Academy of Science USA 108:5897-5902.
13. Graham M. A., Marek L. F., Lohnes D., Cregan P., Shoemaker R. C. (2000) Expression and genome organization of resistance gene analogs in soybean. Genome 43:86-93.
14. Hajjar R. and Hodgkin T. (2007) The use of wild relatives in crop improvement: A survey of development over the last 20 years. Euphytica 156:1-13.
15. Hartman G. L., Wang T. C. and Hymowitz T. (1992) Sources of resistance to soybean rust in perennial Glycine species. Plant Disease 76:396-399.
16. Hayes A. J. (2000) Expression of two soybean resistance gene candidates shows divergence of paralogous single-copy genes. Theoretical Applied Genetics 101:789-795.
17. Hulbert S. H., Webb C. A., Smith S. M. and Sun Q. (2001) R-gene complexes: evolution and utilization. Annual Review of Phytopathology 39:285-312.
18. Kobe B., and Deisenhofer J. (1995) Proteins with leucine-rich repeats. Current Opinion in Structural Biology 5:409-416.
19. Kuang H., Wei F., Marano M. R., Wirtz U., Wang X., Liu J., ... & Baker B. (2005) The R1 resistance gene cluster contains three groups of independently evolving, type I R1

homologues and shows substantial structural variation among haplotypes of Solanum demissum. The Plant Journal 44(1):37-51.

20. Kuwabara C., Arakawa K. and Yoshida S. (1999) Abscisic acid induced secretory proteins in suspension-cultured cells of winter wheat. Plant Cell Physiology 40:184±191.
21. Langenbach C., Campe R., Schaffrath U., Goellner K., Conrath U. (2013) UDP-glucosyltransferase UGT84A2/BRT1 is required for Arabidopsis nonhost resistance to the Asian soybean rust pathogen Phakopsora pachyrhizi. New Phytologist 198: 536-545.
22. Langlois-Meurinne M., Gachon C., and Saindrenan P. (2005) Pathogen-responsive expression of glycosyltransferase genes UGT73B3 and UGT73B5 is necessary for resistance to Pseudomonas syringae pv tomato in Arabidopsis. Plant Physiology 139:1890-1901.
23. Larkin M. A., Blackshields G., Brown N. P., Chenna R., Mcgettigan P. A. et al., (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-2948.
24. Leister D. (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease R-genes. Trends in Genetics 20:116-122.
25. Lherminier J., Elmayan T., Fromentin J., Elaraqui K.T., Vesa S., Morel J., Verrier J.L., Cailleteau B., Blein J.P., and Simon-Plas F. (2009) NADPH oxidase-mediated reactive oxygen species production: Subcellular localization and reassessment of its role in plant defense. Molecular Plant-Microbe Interactions Journal 22:868-881.
26. Lisch D., & Bennetzen J. L. (2011) Transposable element origins of epigenetic gene regulation. Current opinion in plant biology, 14(2): 156-161.

27. Liu J., Liu X., Dai L., Wang G. (2007) Recent progress in elucidating the structure, function and evolution of disease R-genes in plants. Journal of Genetics and Genomics 34:765-776.

28. Lu C., & Fedoroff N. (2000) A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. The Plant Cell, 12(12): 2351-2365.

29. Luck J. E., Lawrence G. J., Dodds P. N., Shepherd K. W., & Ellis J. G. (2000) Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. The Plant Cell 12(8):1367-1377.

30. Martin G. B., Bogdanove A. J. and Sessa G. (2003) Understanding the functions of plant disease resistance proteins. Annual Review of Plant Biology 54:23–61.

31. Meyers B. C., Dickerman A. W., Michelmore R. W., Sivaramakrishnan S., Sobral B. W. and Young N. D. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. The Plant Journal 20:317-332.

32. Meyers B. C., Kozik A., Griego A., Kuang H., Michelmore R. W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell 15: 809-834.

33. Michelmore R. W., Meyers B. C. (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res 8:1113-1130

34. Mondragón-Palomino M., Meyers B. C., Michelmore R. W., & Gaut B. S. (2002) Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. Genome research, 12(9):1305-1315.

35. Mugford S. T., Qi X., Bakht S., Hill L., Wegel E., Hughes R. K., Papadopoulou K., Melton R., Philo M., Sainsbury F., Lomonossoff G. P., Roy A. D., Goss R. J. M., Osbourn A. (2009) A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. Plant Cell 21:2473-2484.
36. Murray M. G., & Thompson W. F. (1980) Rapid isolation of high molecular weight plant DNA. Nucleic acids research 8(19): 4321-4326.
37. Offord V., Coffey T.J., Werling D. (2010) LRR finder: a web application for the identification of leucine-rich repeats and an integrative Toll-like receptor database. Developmental and Comparative Immunology 34:1035.
38. Pan Q., Wendel J., & Fluhr R. (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. Journal of molecular evolution, 50(3): 203-213.
39. Prescott-Allen R. (1988) Genes from the Wild: Using Wild Genetic Resources for Food and Raw materials. International Institute for Environment and Development, London
40. Pryor T. (1987) The origin and structure of fungal disease resistance genes in plants. Trends in Genetics 3:157-161.
41. Purugganan M. D., & Wessler S. R. (1995) Transposon signatures: species-specific molecular markers that utilize a class of multiple-copy nuclear DNA. Molecular ecology 4(2): 265-270.
42. Ribas A. F., Cenci A., Combes M. C., Etienne H., & Lashermes P. (2011) Organization and molecular evolution of a disease-resistance gene cluster in coffee trees. BMC genomics, 12(1): 240.

43. Shao Z. Q., Zhang Y. M., Hang Y. Y., Xue J. Y., Zhou G. C., Wu P., ... & Chen J. Q. (2014) Long-Term Evolution of Nucleotide-Binding Site-Leucine-Rich Repeat Genes: Understanding Gained from and beyond the Legume Family. Plant physiology 166(1):217-234.
44. Song W. Y., Pi L. Y., Wang G. L., Gardner J., Holsten T., & Ronald P. C. (1997) Evolution of the rice Xa21 disease resistance gene family. The Plant Cell 9(8): 1279-1287.
45. Sugio A., Kingdom H. N., MacLean A. M., Grieve V. M., Hogenhout S. A. (2011) Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. Proceedings of the National Academy of Science USA 108:E1254–E1263.
46. Tameling W. I., Elzinga S. D., Darmin P. S., Vossen J. H., Takken F. L., Haring M. A. and Cornelissen B. J. (2002) The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. Plant Cell 14:2929-2939.
47. Tamura K., Peterson D., Peterson N., Stecher G., Nei M. *et al*., (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739.
48. Wicker T., Guyot R., Yahiaoui N., Keller B. (2003) CACTA transposons in Triticeae: a diverse family of high-copy repetitive elements. Plant Physiology 132:52-63.
49. Yang S., Feng Z., Zhang X., Jiang K., Jin X., Hang Y., Chen J. Q., Tian D. (2006) Genome-wide investigation on the genetic variations of rice disease R-genes. Plant Molecular Biology 62:181-193.

Table 4.1 Degenerate and common primers used to screen a soybean fosmid library for NBS RGHs.

Name	Sequences (5'-3')	Orientation	Specificity
(Degenerate primers used to amplify putative NBS RGHs from genomic DNA)			
PloopF	ggHWtgggHggRWtRggVaag	Forward	PloopF
Kinase	ctBStYgtYYTsgATgAygT	Forward	Kinase/GLPLAL
GLPLAL	CARRgYCARWggAAgTCC	Reverse	PloopF/GLPLAL
MHD	ATCWYKWAgDWKRTCRTgCAT	Reverse	PloopF/MHD
Ploop1	ggIggI RTIggIAAIACIAC	Forward	Ploop1
GLP1	IAgIgYIAgIggIAgICC	Reverse	Ploop1/GLP1
GLP2	IATIgCIAgIggIAAICC	Reverse	Ploop1/GLP2
GLP3	IATIgCIAAIggIAgICC	Reverse	Ploop1/GLP3
GLP4	IAAIgYIAgIggIAgICC	Reverse	Ploop1/GLP4
GLP5	IAgIgCIAAIggIAgICC	Reverse	Ploop1/GLP5
(Common primers used to amplify target NBS RGHs from fosmid library)			
GLPF-29	GGAAGACGACACTTGC	Forward	P-Loop
GLPR-318-A03	TAGAGTCATCGTAACTACTCG	Reverse	GLPL
GLPF-29	GGAAGACGACACTTGC	Forward	P-Loop
GLPR-318-D10	CAGAGTCATCATTACCACTCG	Reverse	GLPL
GLPF-29	GGAAGACGACACTTGC	Forward	P-Loop
GLPR-485	GAATCGAGCAGCAACATTCGT	Reverse	GLPL
GLPF-29	GGAAGACGACACTT	Forward	P-Loop
GLPR-276	ACAGGTACTTGCTGGAG	Reverse	GLPL
GLPF-213-D06	TGGGAAAATGATAATAGC	Forward	P-Loop
GLPR-372	GTGCACCTAACTTGTAAG	Reverse	GLPL
GLPF-72	TCTCCTCAATATGATGCTCGT	Forward	P-Loop
GLPR-492	GTACTGAAGTATGCCAATGG	Reverse	GLPL
GLPF-213-D05	AAACTTCACAGTCACATC	Forward	P-Loop
GLPR-330	TGGTGTCTCAGATGCCAT	Reverse	GLPL
GLPF-213-F05	AGACGACATGGTCATACG	Forward	P-Loop
GLPR-361	AGCACTTCAGAATCAGCCAC	Reverse	GLPL

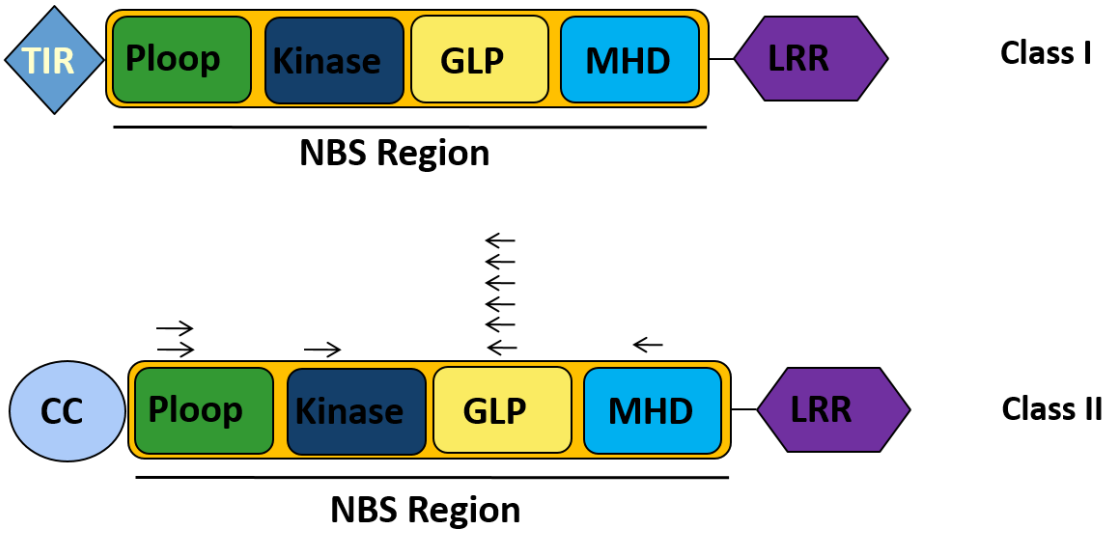


Figure 4.1 Location on the NBS-LRR class resistance genes that degenerate primers were designed from.The arrow represents the location and direction of each degenerate primer. Class I represents the NBS-LRR R-genes with toll interleukin receptor (TIR) in the N-terminus while Class II represents the NBS-LRR R-genes with Coiled-coil domain (CC) in the N-terminus All the degenerate primers were designed based on the alignment of NBS-LRR R-genes from soybean, rice, maize and sorghum. Arrows indicate the direction of the primers designed in the conserved domains. Six reverse primers were designed based GLP motifs, one reverse primer was designed based on MHD motif. Two forward primers were designed from the P-Loop motif and one forward primer was designed from Kinase motif within NBS domain.

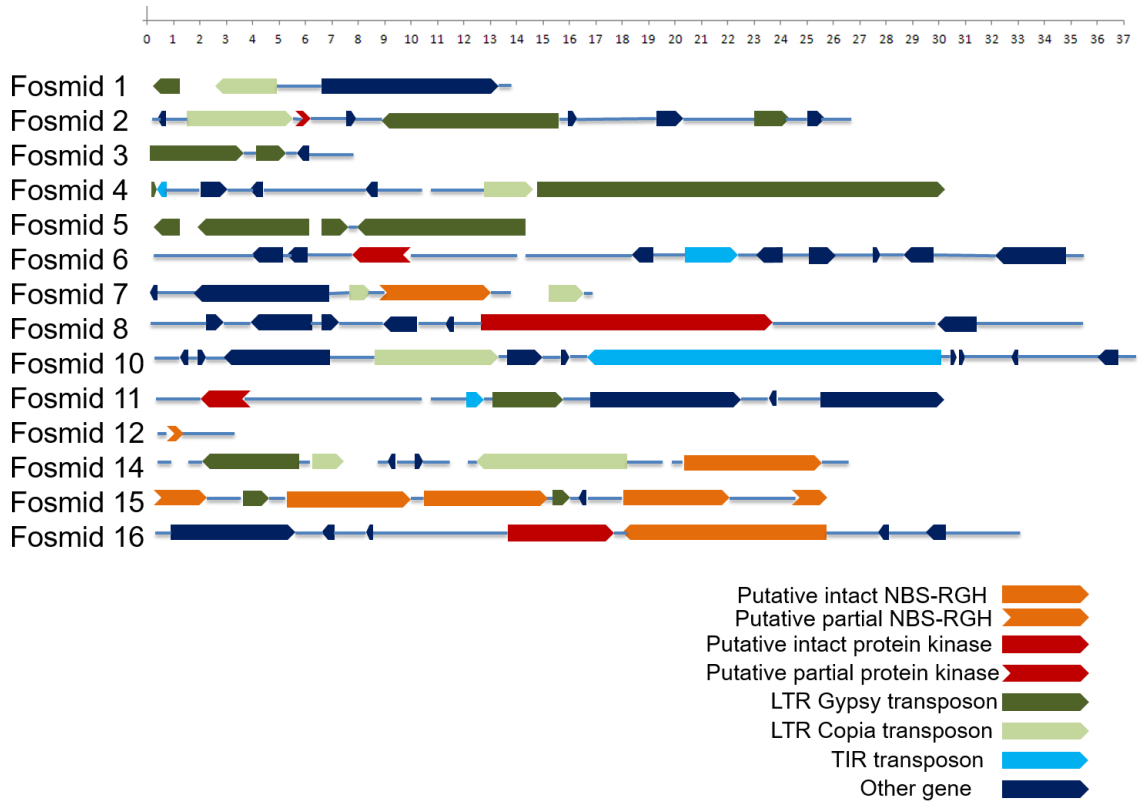


Figure 4.2 Gene structures of 12 fosmids containing NBS RGHs in Bingnan soybean cultivar. Genes were predicted by gene finding programs FgeneSH. The direction of the arrow represents the direction of transcription of each gene predicted by FgeneSH. Different colors represent different categories of genes (orange: putative intact or partial NBS-RGHs; red: putative intact or partial protein kinase; dark green: LTR Gypsy transcposons; light green: LTR Copia transposons; light blue: TIR transposons, dark blue: all other genes). Number scale corresponds to the size of the fosmids in kilo-bases.

Table 4.2 Summary statistics for 16 soybean fosmids containing RGHs

Fosmid Name	No. of Genes [a]	No. of intact Protein Kinase [b]	No. of Partial Protein Kinase	No. of intact NBS-LRR RGHs	No. of partial NBS-LRR RGHs [b]	TE Number[c]
1	3	0	0	0	0	2
2	9	0	1	0	0	3
3	3	0	0	0	0	2
4	7	0	0	0	0	4
5	4	0	0	0	0	4
6	10	0	1	0	0	1
7	5	0	0	0	1	2
8	7	1	0	0	0	0
9*	-	-	-	-	-	-
10	11	0	0	0	0	2
11	6	0	1	0	0	2
12	1	0	0	0	1	0
13*	-	-	-	-	-	-
14	9	0	0	1	0	3
15	8	0	0	3	2	2
16	7	1	0	1	0	0

[a] Number of genes on each fosmid were predicted by Fgenesh
[b] No. of partial NBS-RGHs were predicted by NCBI conserved region search and HMMER search.
[c] Transposable elements (TEs) included in the fosmid. TEs were identified by Williams 82 Transposable Element Database.
*Fosmid 9 sequence was shown to be a vector sequence. Fosmid 13 was failed to be sequenced.

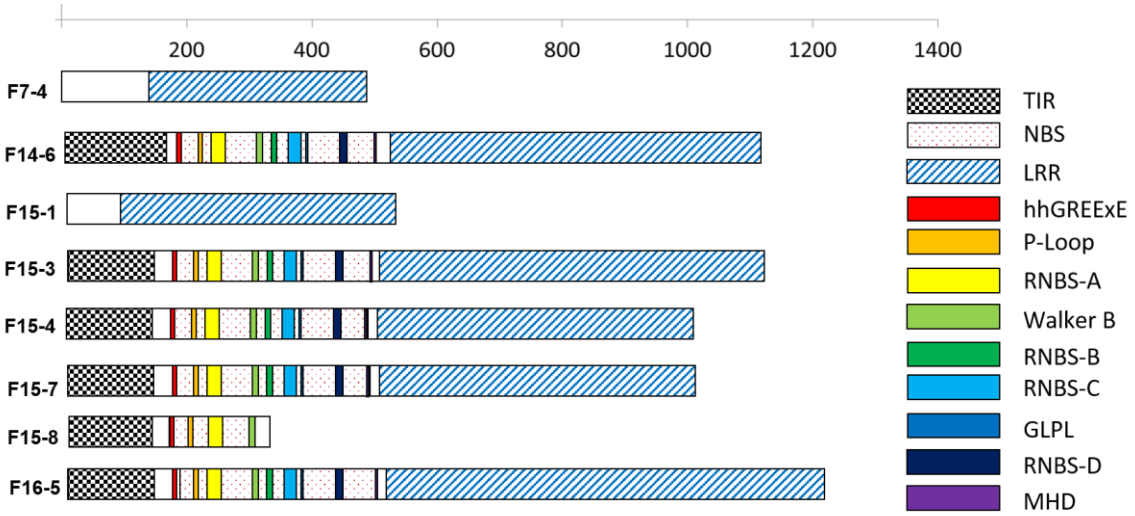


Figure 4.3 Structure of R-proteins encoded by soybean RGHs from fosmid 7, fosmid 14, fosmid 15 and fosmid 16. Number scale corresponds to the size of the fosmids in amino acids. Different colors represent different motifs of proteins (Black and white: TIR; White with red spots: NBS; Blue stripe line: LRR; Red: hhGREExE; Orange: P-loop, Yellow: RNBS-A; Light green: Walker B; Dark green: RNBS-B; Light blue: RNBS-C; Blue: GLPL; Dark blue: RNBS-D; Purple: MHD).

Table 4.3 Structural comparison of the TIR regions of RHG proteins from fosmids and resistance proteins encoded by reference R-genes. Four conserved regions including TIR1, TIR2, TIR3 and TIR4 were identified by MEME program and were then aligned together.

	YDVFLSFRGxDTR (TIR1)	LxxAIxxS (RK) αxααααSx (TIR2)	(NG) YAxSxWCLxEL (AV) x1 VαPαFYDVDPSxQTGx (YF) xKAFxK (TIR3)	WRxALxxαAxαxG (TIR4)
F7-4	N/A	N/A	N/A	N/A
F14-6	NDVFLSFRGEDTR	LEKAIEESRIFIIVLSE	NYASSSFCLNELDYILKFIKGKGILILPVFYKV DPSDVRNHTGSFGKALTN	WKMALNKVANLSG
F15-1	N/A	N/A	N/A	N/A
F15-3	*DVFLNFRGEDTR	LLKAIQESRIAITVLSE	NYASSSFCLDELVTILHCKRKGLLVIPVFYNV DPSDVRHQKGSYEEEMAK	WRIALKQVADLCG
F15-4	YDVFLSFRGEDTR	LLKAIQDSRIAITVLSE	DFASSSFCLDELATILFCAQYNGMMVIPVFY KVYPCDVRHQKGTYGEALAK	WERALRQVANLSG
F15-7	YDVFLNFRGEDTR	LSKAIQESRIAITVLSQ	NYASSSFCLDELVTILHCKREGLLVIPVFHNV DPSAVRHLKGSYGEAMAK	WRMALHQVADLSG
F15-8	YDVFLSFTGQDTR	LSNAIQESRIAITVLSQ	NYASSSFCLDELVTILHCKSQGLLVIPVFYKV DPSHVRHQKGSYGEAMAK	WRMALHQVADLSG
F16-5	NHVFLSFRGDDTR	LIEAIEESMFALIILSS	NYASSTWCLDELQKILECKKEVFPIFLGVDPS DVRHQRGSFAKAFRD	WRHALREVASYSG
N	NDVFLSFRGKDTR	LEKAIEESRIFIIVLSE	NYAWSSFCLNELDYILKFIKGKGLLVLPVFY KVDPSDVRNHTGSFGESLAY	WKMALNQVANLSG
SR1	YDVFLSFTGQDTR	LSDAIQGSRIAITVLSQ	NYAFSTFCLDELVTILHCKSEGLLVIPVFYKV DPSHVRHQKGSYGEAMAK	WRMALQQVADLSG
KR1	NDVFLSFRGEDTR	LEKAIEESRIFIIVLSE	NYASSSFCLNELDYILKFIKGKGILILPVFYKV DPSDVRNHTGSFGKALTN	WKMALNKVANLSG

Table 4.4 Structural comparison of the NBS regions of RGH proteins from soybean fosmids and resistance proteins encoded by reference R-genes. Nine conserved regions including TN (TIR-NBS) linker, P-Loop, RNBS-A, Walker-B, RNBS-B, RNBS-C, GLPL, RNBS-D and MHD were identified by MEME program and were then aligned together.

	hhGRExE (TN linker)	P-Loop/Walker A	RNBS-A-TIR	Walker B/Kinase -2	RNBS-B/ Kinase-3a	RNBS-C	GLPL	RNBS-D	MHD
F7-4	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
F14-6	PVGLESRI	GLGGVGKT	RETSKKHGLQ HLQRNLLSET	LLILDDVDK	GSRVIITTR	TYEVNEL NEEYALE LLNWKAF	GLPL	FLDIACCF	LHD
F15-1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
F15-3	PVGLGSQV	GMGGLGKT	REESNKHGLK HLQSILLSKL	LLILDDVDK	GSRVIITTR	TYEVKVL NQSAALQ LLTWNAF	GLPL	FLDIACCF	MHD
F15-4	PVGLESKV	GMGGIGKS	RESSNNHGLQ HLQSILLSEI	LLILDDVDK	GSIIIITTR	RYEVEVL NQNAALQ LLTWNAF	GLPL	FLDIACCF	MHD
F15-7	PVGLGSQV	GMGGLGKT	REESNKHGLK HLQSILLSKL	LLILDDVDK	GSRVIITTR	TYEVKVL NHNAALQ LLTWNAF	GLPL	FLDIACCF	MHD
F15-8	PVGLESEV	GMGGLGKT	REESNKHGLK HLQSILLSKL	LLILDDVDK	N/A	N/A	N/A	N/A	N/A
F16-5	LVGIDSRM	GRGGIGKT	REVSKTNGLV HIQKELSNL	LLVLDDVSE	GSRVIITTR	TCKARAL AQNEALQ LICLKAF	GLPL	FLDIACFF	MHD
N	PVGLESRI	GLGGIGKT	RETSKTHGLQ YLQRNLLSET	LLILDDVDK	GSRVIITTR	TYEVNEL NEEYALQ LLSWKAF	GLPL	FLDISCCL	LHD
SR1	PVGLGSQV	GMGGLGKT	REESNLKHLQ	LLILDDVDK	GSRVIITTR	TYEVKVL	GLPL	FLDIACCF	MHN

			SSLLSKLLGEK			NHNAALH LLTWNAF				
KR1	PVGLESRI	GLGGVGKT	RETSKKHGLQ HLQRNLL SEM	LLILDDVDK	GSRVIITTR	TYEVNEL NEEYALE LLNWKAF	GLPL	FLDIACCF	LHD	

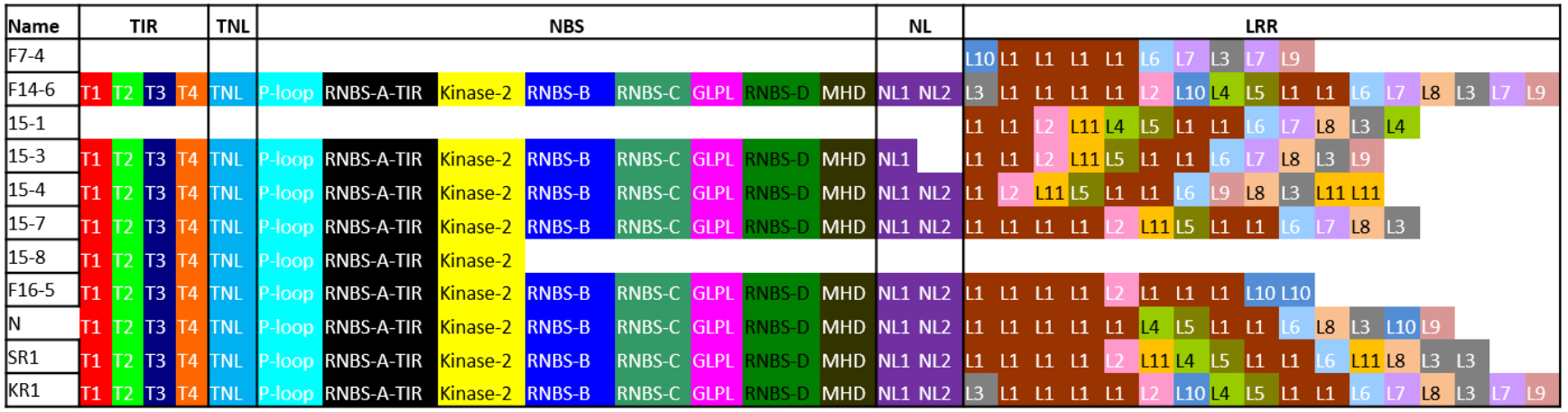
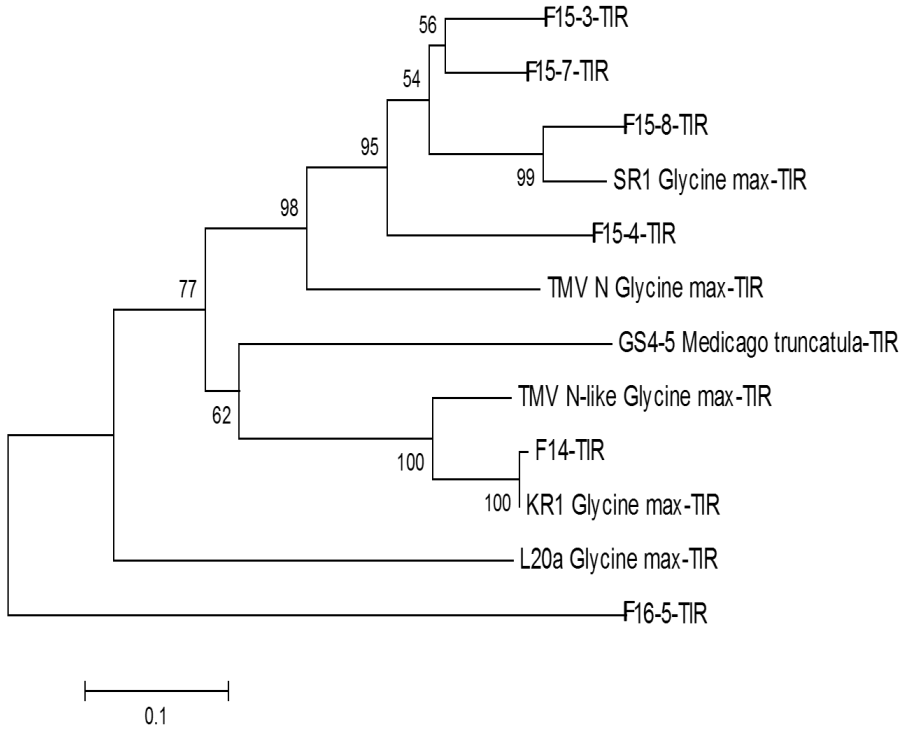
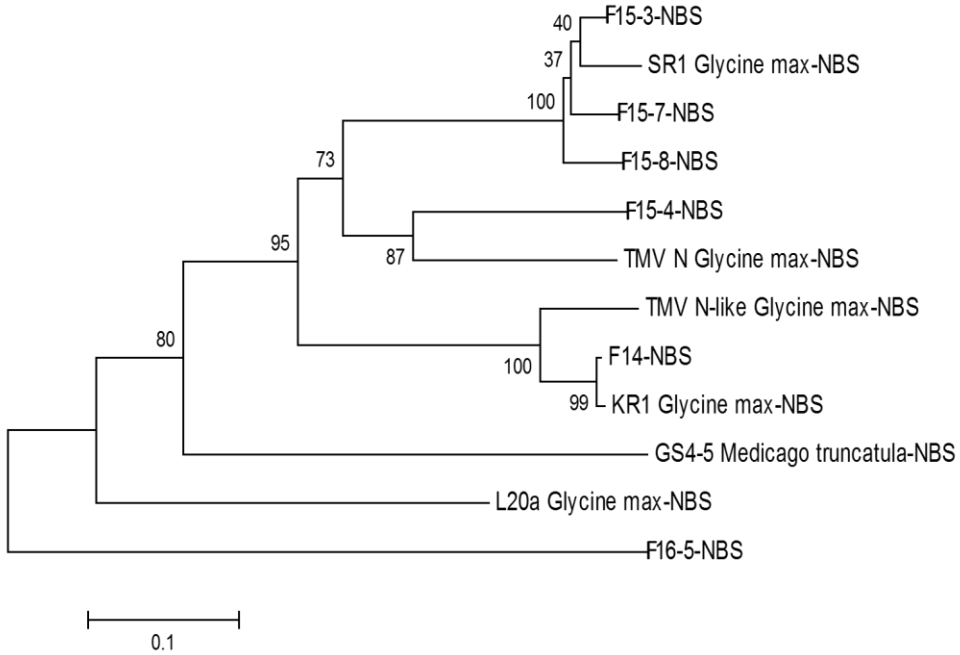


Figure 4.4 Conserved motifs identified in the TIR, NBS and LRR domains in soybean RGH proteins encoded by RGHs from fosmid 7, 14, 15 and 16 and reference proteins N, SR1 and KR1. TNL represents TN linker motif and NL represents NBS-LRR linker motif. There are four TIR motif identified in TIR domain while over 11 LRR motifs (L1-L11) identified in LRR region.

(A)


F15-3-TIR
F15-7-TIR
F15-8-TIR
SR1 Glycine max-TIR
F15-4-TIR
TMV N Glycine max-TIR
GS4-5 Medicago truncatula-TIR
TMV N-like Glycine max-TIR
F14-TIR
KR1 Glycine max-TIR
L20a Glycine max-TIR
F16-5-TIR
56
54
95
99
98
77
62
100
100
0.1

(B)


F15-3-NBS
SR1 Glycine max-NBS
F15-7-NBS
F15-8-NBS
F15-4-NBS
TMV N Glycine max-NBS
TMV N-like Glycine max-NBS
F14-NBS
KR1 Glycine max-NBS
GS4-5 Medicago truncatula-NBS
L20a Glycine max-NBS
F16-5-NBS
40
37
100
73
95
87
80
100
99
0.1

(C)

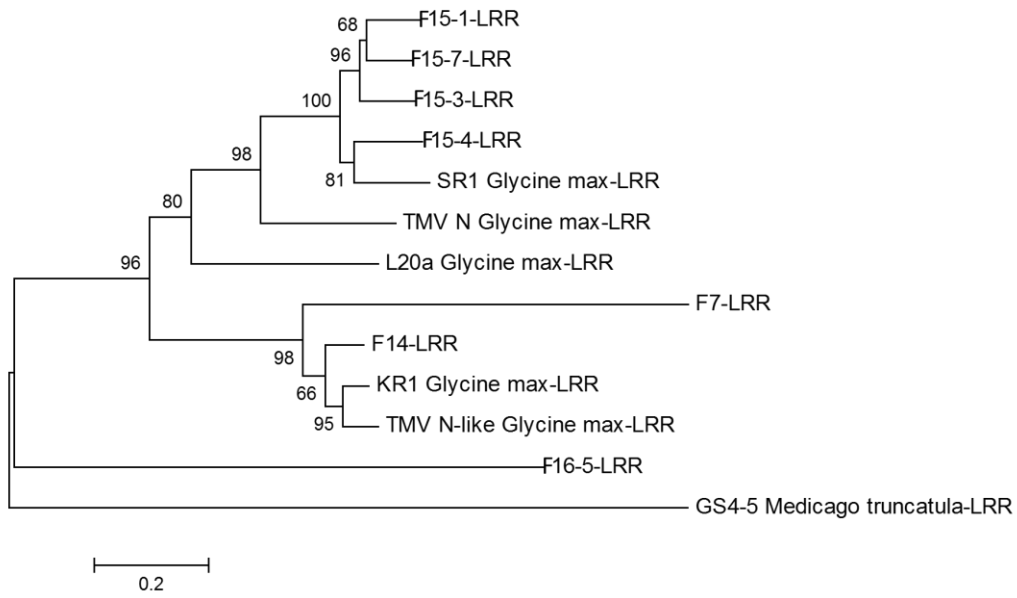
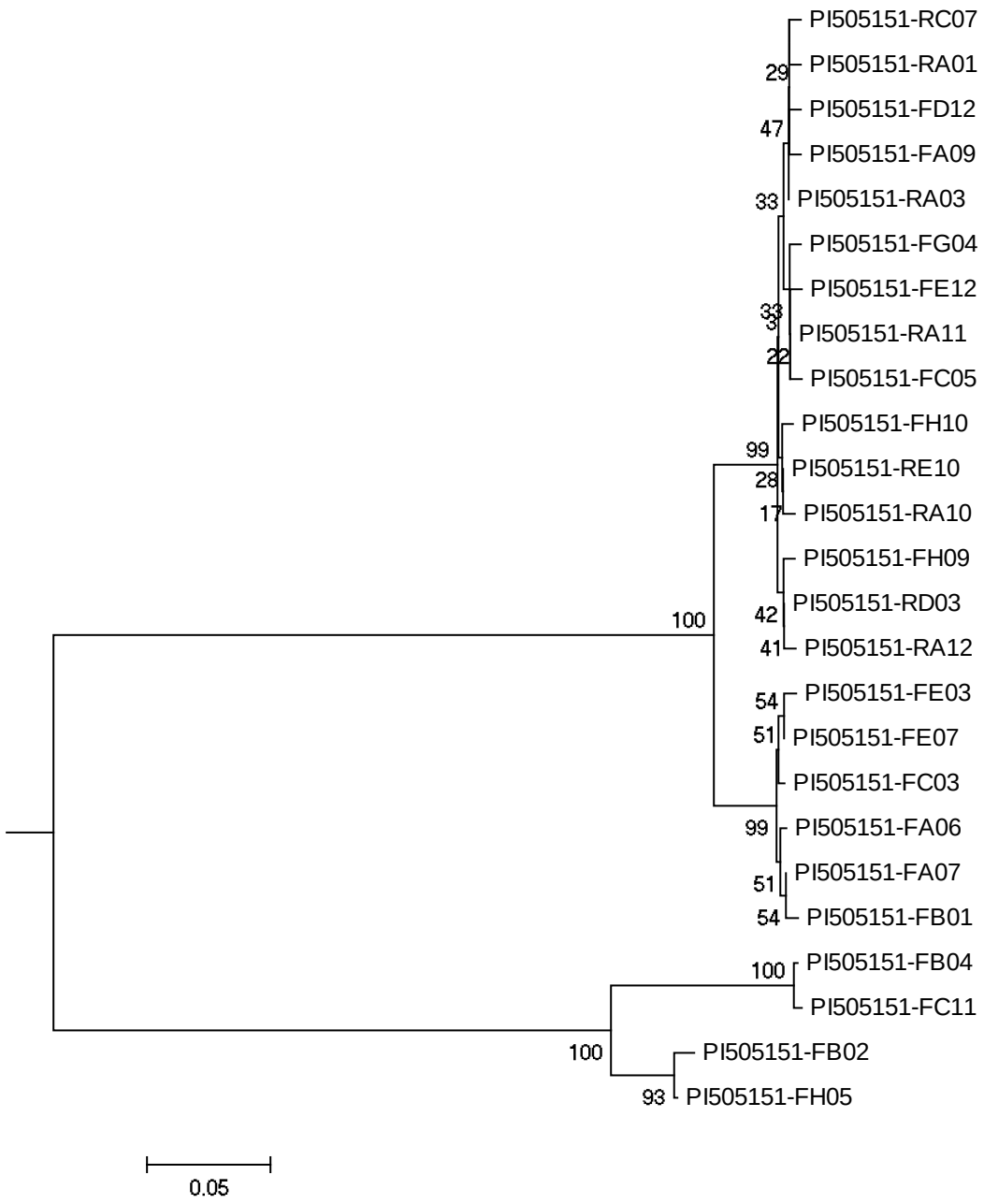


Figure 4.5 Neighbor-Joining tree of the different domains from RGHs trees generated based on the alignment of amino acid sequences TIR regions (A), NBS regions (B) and LRR regions (C) of soybean RGHs and reference R-genes.

Table 4.5 Number of NBS and LRR nucleotide and amino acid sequences cloned from each soybean wild relative.

PIs	No. of clones being sequenced	No. of NBS nucleotide sequences	No. of NBS aa sequences with ORF	No. of TIR-NBS aa sequences	No. of Non-TIR-NBS aa sequences	No. of LRR nucleotide sequences	No. of LRR aa sequences with ORF
483218	96	48	37	11	26	92	41
373983A	96	29	20	9	11	32	5
440996	96	74	17	17	0	70	5
446988	96	35	34	31	3	36	22
446995	96	14	13	5	8	18	10
483199	96	71	24	22	2	64	17
505151	96	32	25	21	4	30	18
339661	96	26	0	0	0	33	1
Total	768	329	170	116	54	375	119

(A)


PI505151-RC07
PI505151-RA01
PI505151-FD12
PI505151-FA09
PI505151-RA03
PI505151-FG04
PI505151-FE12
PI505151-RA11
PI505151-FC05
PI505151-FH10
PI505151-RE10
PI505151-RA10
PI505151-FH09
PI505151-RD03
PI505151-RA12
PI505151-FE03
PI505151-FE07
PI505151-FC03
PI505151-FA06
PI505151-FA07
PI505151-FB01
PI505151-FB04
PI505151-FC11
PI505151-FB02
PI505151-FH05
29
47
33
33
3
22
99
28
17
100
42
41
54
51
99
51
54
100
100
93
0.05

(B)

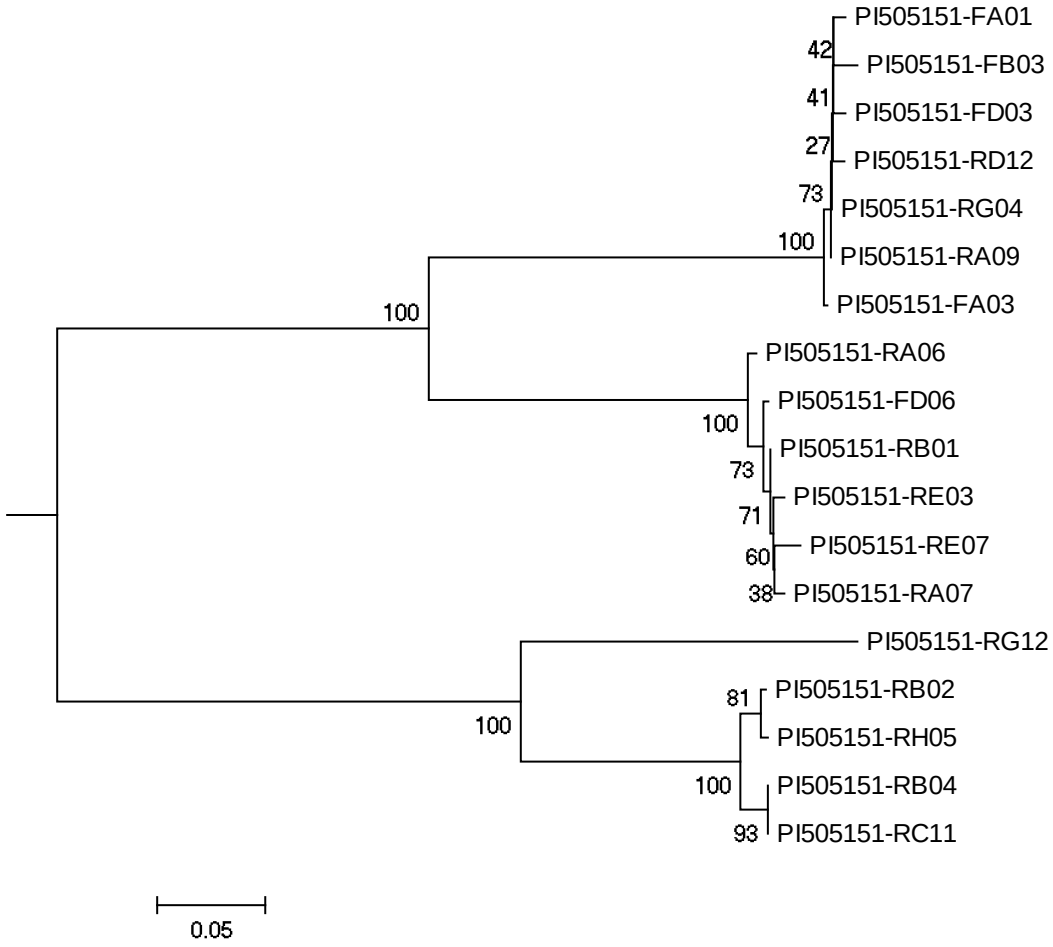


Figure 4.6 Two phylogenetic trees were generated based on the alignment of amino acid sequences of NBS regions (A) and LRR regions (B) of RGH clones of PI 505151.

CHAPTER 5

CONCLUSIONS

Plants have evolved a complex mechanism to protect themselves from invading pathogens. Significant efforts have been made in the past few years to better understand plant disease resistance gene evolution. This has begun to provide some insight into the source of R-gene specificity and the creation of new resistance specificities. Continued characterization of R-genes, *Avr* genes and genes coding for interacting proteins, and analysis of allelic diversity in these genes, should allow us to establish clearer trends as to how resistance gene loci evolve.

This dissertation describes the parallel analysis of *P. pachyrhizi* variability, candidate soybean resistance genes and soybean effectors for the study and improvement of soybean resistance. In regards to the pathogen, we collected *P. pachyrhizi* field isolates from different locations of Georgia, and analyzed the population structure utilizing phenotypic and population genetic analysis. At least three pathotypes were identified in the current Georgia *P. pachyrhizi* population based on the phenotypic data. They are likely carrying three different combinations of avirulence effector proteins. Multiple genotypes were found in Georgia *P. pachyrhizi* field isolates were corresponding to genotypes from different countries and southern states in the USA, demonstrating a high genetic diversity in current *P. pachyrhizi* population in Georgia. Two predominate genotypes representing two major groups of genotypes were identified suggesting they may be associated with the ancestral genotypes that all of the other genotypes were derived. This also suggested that the two genotypes may be the most problematic genotypes in the Georgia field isolates. For example, if a genotype from the pathogen population becomes

predominate, it has the ability to mutate and change where it is no longer recognized by the plant R-gene protein. This results in susceptibility. By identifying the potentially problematic genotypes in the pathogen population, effective R-genes can be selected and deployed. In regards to the host, the soybean lines, adapted cultivars and soybean wild relative accessions were tested for resistance against current ASR in Georgia. Japanese cultivar Hyuuga and several wild relatives conferred a high level of resistance. An extensive structural analysis targeting NBS-LRR resistance genes was performed in both soybean cultivar Bingnan and eight wild relatives and provided evidence on how the R-genes in each soybean genome are maintained and evolve. In this study, a parallel analysis of both host and pathogen was conducted by performing a dual-RNA-seq. This analysis separated expressed fungal genes in the *P. pachyrhizi* isolate Expo from the upregulated soybean genes and characterized soybean defense-related genes and *P. pachyrhizi* pathogen effector genes that are potentially responsible for Hyuuga-mediated resistance against ASR field isolates in Georgia. Nine putative candidate disease resistance genes were identified in Hyuuga. Five putative effectors were identified in the Expo isolate and may be the molecular determinates for pathogenesis of this isolate.

The idea of obtaining genetic resistance to ASR has been an area of intense research for nearly a decade (Hartman *et al*., 2005). The data presented here suggests that durable resistance is not possible with the current strategy and that successful control of ASR with host resistance cannot be achieved without knowledge of variation in ASR populations and soybean genotypes (Martin *et al*., 2003). Hence, knowledge of both the host and the pathogen is critical to effectively breed for durable resistance to ASR. Parallel analysis of soybean and the *P. pachyrhizi* population in Georgia has allowed us to answer important applied and basic questions that are needed to inform the breeding process including: 1. What resistance genes are in the

current released soybean cultivars? 2. Which host resistance genes are effective against the current ASR population? 3. What are the origins of new ASR isolates in Georgia and how do they spread? and 4. What are the molecular determinants of pathogenesis in ASR? This work will have major implications for the characterization of resistance mechanisms that may be more durable and the development of novel strategies to control ASR.

This work has also provides some insight regarding the evolution of R-genes in soybean. It seems that there is more than one way for a resistance gene to interact with a pathogen *Avr* gene and that there is also more than a single mode of evolution for an R gene, but the relationship between these two features of R-genes is not clear (Sharma *et al*., 2014). It may be that selective forces are more purifying in their effects on R-genes that guard other host components from modification by the pathogen (Bux *et al*., 2012). Perhaps these types of genes are more likely to exhibit a single resistance specificity and have more ancient alleles. Would it then be possible to predict how an R-gene interacts with an *Avr* gene by examining the extent and patterns in nucleotide substitutions between different alleles? Could this molecular signature of selection then be used to predict the utility of R-genes in agriculture? Perhaps ancient resistance genes from wild relatives are more likely to provide durable resistance. A better understanding of the mechanisms in which R-genes evolve may thus allow us to utilize them more efficiently in crop improvement programs and may ultimately allow us to design them for in vitro construction.

This work produced a number of functional tools (RNA-seq data from plant and pathogen; Genomic library) that will be useful for studying gene function in soybean, *P. pachyrhizi* and potentially other plant and pathogen species. These molecular tools were made

with the idea that they would be used to improve the genetic resistance of soybean to ASR and be adopted by breeders/researchers to facilitate their own work.

Literature cited

1. Bux H., Ashraf M., Rasheed A., Poudyal D. S., Kazi A. G., & Afzaal M. (2012) Molecular basis of disease resistance in cereal crops: an overview. In Crop Production for Agricultural Improvement (pp. 477-489). Springer Netherlands.
2. Dangl J. L. and Jones J. D. G. (2001) Plant pathogens and integrated defence responses to infection. Nature 411: 826-833.
3. Hartman G. L., Miles M. R. and Frederick R. D. (2005) Breeding for resistance to soybean rust. Plant Disease 89:664-666.
4. Martin G. B., Bogdanove A. J. and Sessa G. (2003) Understanding the functions of plant disease resistance proteins. Annual Review of Plant Biology 54:23–61.
5. Sharma T. R., Das A., Thakur S., & Jalali B. (2014) Recent Understanding on Structure, Function and Evolution of Plant Disease Resistance Genes. In Proc. Indian Nat. Sci. Acad 1:83-93.