12-2011

Functional Genomic Studies of Soybean Defenses against Pests and Soybean Meal Improvement

Jingyu (Lynn) Lin
jlin11@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Jingyu (Lynn) Lin entitled "Functional Genomic Studies of Soybean Defenses against Pests and Soybean Meal Improvement." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Plants, Soils, and Insects.

Feng Chen, Major Professor

We have read this dissertation and recommend its acceptance:

Vincent R. Pantalone, Charles N. Stewart, Jr., Arnold M. Saxton

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Functional Genomic Studies of Soybean Defenses against Pests and Soybean Meal Improvement

A Dissertation Presented for
the Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Jingyu Lin
December 2011

Dedication

I dedicate my dissertation work to my parents, Baozhong Lin and Shulan Bian for their unconditional love and support through the years. They always encourage me to continue my education and devote to my dream. Without their continued understanding and support, I could not have completed this process. The dissertation is also dedicated to my husband, Nan Zhao, who accompanied me through all the hardships and shared my uncertainties, challenges, frustrations and joys for completing this dissertation.

Acknowledgement

I want to express my deepest gratitude to my advisor Dr. Feng Chen for his guidance, understanding and encouragement through my Ph.D. study. Dr. Chen has been training me to be an independent researcher with creative thinking and an attitude toward strenuous research efforts. I also sincerely appreciate my co-advisor, Vincent R. Pantalone. His mentorship and encouragement helped me build self-confidence in research and life. I also want to thank Dr. C. Neal Stewart Jr. Without his funding support, I would not have such a wonderful opportunity to complete my Ph.D. study. I would also like to give thanks to another committee member, my minor advisor, Dr. Arnold M. Saxton, who offered me constructive suggestions for the dissertation. Besides my committee members, I also want to express my deep gratitude toward Drs. Mitra Mazarei, Mohammad R. Hajimorad, and Prakash R. Arelli for giving me guidance, and sharing their experimental experiences and kind cooperation during my study. I would also like to thank Lori D. Osburn for her efforts of proofreading my dissertation. Moreover, it has been much fun to work with the researchers at the University of Tennessee, where I appreciate all of my lab mates for their enlightening discussion of science and life. I also appreciate all of the genuine help from people in the department.

Abstract

Soybean [*Glycine max* (L.) Merr.] is an important crop worldwide. It has been widely consumed for protein, oil and other soy products. To develop soybean cultivars with greater resistance against pests and improved meal quality, it is important to elucidate the molecular bases of these traits. This dissertation aims to investigate the biochemical and biological functions of soybean genes from four gene families, which are hypothesized to be associated with soybean defense against pests and soybean meal quality. There are three specific objectives in this dissertation. The first one is to determine the function of components in the salicylic acid (SA) signaling pathway in soybean resistance against soybean cyst nematode (*Heterodera glycines,* SCN). The second one is to determine whether insect herbivory induce the emission of volatiles from soybean, and if so, how these volatiles are biosynthesized. The third objective is to identify and characterize soybean mannanase genes that can be used for the improvement of soybean meal quality. The soybean genome has been fully sequenced, which provides opportunities for cross-species comparison of gene families of interest and identification of candidate genes in soybean. The cloned cDNAs of putative genes were expressed in *Escherichia coli* to produce recombinant enzymes. Through biochemical assays, these proteins were proved to be soybean salicylic acid methyltransferase (GmSAMT1), methyl salicylate esterase (GmSABP2-1), α[alpha]-farnesene synthase (GmTPS1) and E-β[beta]-caryophyllene synthase

(GmTPS2), and endo-β[beta]-mannanase (GmMAN1). Through a transgenic hairy root system harboring overexpression of *GmSAMT1* and *GmSABP2-1*, both of these two genes were evaluated for their biological function related to resistance against SCN. The results showed that the over-expression of *GmSAMT1* and *GmSABP2-1* in the susceptible soybean background lead to enhanced resistance against SCN. Among four putative soybean mannanase genes, one gene was cloned and characterized. *GmMAN1* showed the endo-β[beta]-mannanase hydrolyse activity and can hydrolyze cell walls isolated from soybean seeds. In summary, using comparative and functional genomics, a number of genes involved in soybean defense and meal quality were isolated and characterized. This study provides novel knowledge and molecular tools for the genetic improvement of soybean for enhanced resistance and improved meal quality.

Table of Contents

Tables

Figures

Chapter I. Literature Review: Functional Genomic Studies of Soybean Defense Against Pests and Soybean Meal Improvement

1. Economic Importance of Soybean

Soybean [*Glycine max* (L.) Merr.] is the most widely grown legume in the world, providing an important source of protein and oil. Soybean can be utilized in many ways such as ingredients in the formulation of a multitude of human foods, animal feed, and industrial products. As the dominant oil-seed in world trade, soybean contributes about 56% of global oilseed production. The estimated value of soybean products is about $48.6 billion world-wide and about $18.7 billion in the U.S. alone (Wilson 2008). Therefore, soybean is considered as one of the most important economic crops. Although global soybean production has increased steadily over the past century, future demand for soybean still cannot be satisfied due to a growing world population and limited land resources (Wilson 2008).

2. Soybean Disease-causing Pathogens and Pests

During its entire life cycle, soybean may be attacked by many pathogens, such as fungi, bacteria, viruses, and nematodes (**Table 1.1**), and suffer many diseases in any tissue. For instance, the bacterial pathogen *Pseudomonas syringae* pv. *glycinea*, can cause bacterial blight, overwinter in crop residue, and be transmitted by seed. In laboratory studies, *P. syringae* is usually chosen as the model of bacterial pathogen (Alvarez et al. 1995). Fungal pathogens can also cause soybean diseases on both above- and below-ground tissues, accounting for approximately 50% of all soybean disease losses in the United States and around the world (Wrather and Koenning 2006). Soybean mosaic virus (SMV) detected on seeds and on plants in the field, is the major virus

causing soybean disease. Since this viral disease is seed-borne, it may cause problems in the next growing season. Its effect on seed quality is the main concern. Virus infected soybean may produce fewer, smaller, and often mottled seed. Another concern is that SMV infected soybean tends to be dually infected by other viruses, which magnifies the risk of yield loss and reduces seed quality even further (Giesler et al. 2002). In addition, nematodes and soybean insect pests such as beetles, caterpillars, aphids, and spider mites can also cause yield loss of soybean.

2.1. Soybean Cyst Nematode and Soybean Resistance

The soybean cyst nematode (*Heterodera glycines*, SCN), is the pest that causes the most economic damage of soybean in the United States, causing nearly $1 billion in yield losses (Wrather and Koenning 2006).

SCN is a small plant-parasitic roundworm and most stages of SCN cannot be seen by unaided eyes. SCN feeds on the soybean roots and robs nutrients from the soybean. One important feature, genetic heterogeneity of SCN populations, endows this pathogen to readily overcome resistance (Niblack et al. 2002). As a sedentary obligate endoparasite of roots, SCN is a microscopic, worm-like organism ranging in length from 0.04 cm to 0.16 cm. To complete their life cycle, infective second-stage juveniles (J2) enter host roots and migrate intracellularly within the cortical tissue to the vascular cylinder. The juveniles then initiate formation of specialized feeding sites called syncytia, which function as metabolic sinks to nourish the nematodes (Jones 1981). The nematodes feed exclusively from their syncytia as they develop into adult males and

females. Once fertilized, the female soybean nematode produces up to several hundred eggs that, for the most part, are retained within the nematode uterus. After the female's death, her body develops into a protective cyst around the eggs, giving the nematode its name. In susceptible soybean cultivars, nematodes depend entirely on functional syncytia to acquire nutrients to develop into reproductive adult males or females. J2 nematodes also penetrate roots of the resistant soybean cultivars and initiate syncytia. However, the resistance soon manifests itself by degrading the young syncytia and restricting the further development of the juveniles into adults (Endo 1991). The migration process, feeding cell formation, and maintenance of nematodes are proposed to be mediated through nematode secretions and are accompanied by changing plant genes expression (Davis and Mitchum 2005). Some of these genes encode a variety of cell wall-degrading enzymes and proteins involved in altering the normal host plant physiology. Parasitism proteins are injected by SCN into host root tissues through a stylet, a protrusible hollow mouth spear that is connected to the SCN's gland cells (Elling et al. 2007).

SCN infestations are difficult to identify in the field because the above-ground symptoms of SCN infestations are easily confused with nutrient deficiency symptoms. When soybean plants are severely damaged by nematodes, they become stunted and turn chlorotic (Kulkarni et al. 2008). These symptoms are mistaken for nutrient deficiency and water stress. Until the population of SCN reaches a certain number, its damaging effect can not be clearly identified by farmers. Currently, the most accurate way of diagnosing SCN infection is to analyze a soil sample and observe the pest

directly (Opperman and Bird 1998). Controlling SCN in commercial soybean productions still remains difficult because SCN has a short life cycle and populations can build rapidly. Frequent changes in population virulence of SCN also contribute to the difficulty in the management of this pest. In addition, the cysts of SCN can survive in the soil for up to nine years and then break to release the eggs under proper conditions, increasing the probability of the nematodes' dispersing via infested soil. The methods used to control and manage SCN in soybean production include crop rotation, the use of SCN-resistant cultivars, and the application of nematicides, which are often used in an integrated manner (Winter et al. 2006). However, these approaches always face economic restrictions, are time-consuming, and use of nematicides can result in environmental problems. Additionally, for some areas, economic factors may limit the use of crop rotation. Considering the long-term demands of soybean, it is critical to develop resistant soybean cultivars through breeding for managing soybean diseases. Even for crop rotation, resistant soybean cultivars are also required.

To facilitate breeding of SCN-resistant soybean, the race test for SCN populations was proposed in 1969 and based on comparative development of females on four differential soybean lines, including PI 548988 (Pickett), PI 548402 (Peking), PI 88788, and PI 90763 (Golden et al. 1970). A new classification scheme was proposed in early 2000. This new system, named the HG type designation, includes seven indicator lines: PI 548402 (Peking), PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, and PI 548316 (Cloud) (Niblack et al. 2002).

2.1.1. Current Understanding of Genetics of Soybean Resistance against SCN

Currently, three major genetic sources for SCN resistance genes have been found in commercial soybean cultivars, PI 88788, PI 548402 (Peking), and PI 437654. Five loci including *rhg1*, *rhg2*, *rhg3*, *Rhg4*, and *Rhg5*, have been reported to be involved in SCN resistance (Concibido et al. 2004). Researchers have tried to identify genes showing high correlation with the QTLs of soybean resistance against SCN. For example, results showed that one soybean gene exhibits a strong linkage to *rhg1* on linkage group G and encodes phosphoribosylformyl-glycinamidine (FGAM) synthase, which catalyzes an important reaction in *de novo* purine biosynthesis and is active in rapidly dividing cells. The promoter-reporter study showed that the expression of this gene was localized within syncytia and induced by beet cyst nematode (*Heterodera schachtii*) (Vaghchhipawala et al. 2004). Another candidate SCN resistant gene *Glyma18g02680.1* located at the Rhg1 locus, was shown to encode an apparent leucine-rich repeat transmembrane receptor-kinase (LRR-kinase). The function of this gene was further studied through transgenic roots generated with *Agrobacterium rhizogenes*. Unexpectedly, the reduced expression of this gene through the use of amiRNA did not show a significant effect on SCN resistance (Melito et al. 2010).

2.1.2. Molecular Biology and Genomics of Soybean Resistance Against SCN

The availability of the full genome sequence and large dataset of expressed sequence tags (ESTs) of soybean make it possible to identify candidate genes of interest at the genome scale. In addition, the construction of whole genome Affymetrix DNA chip

and EST-based cDNA microarray provide novel tools for gene identification by comparing gene expression changes during SCN parasitism.

A number of microarray studies have been performed to study gene expression changes of soybean when infected by SCN. Considering the complexity of soybean resistance, some different soybean cultivars (Williams 82, Peking, and Kent) and germplasm lines (PI 88788, PI 467312, and PI 507354), and SCN populations (NL1-RHg, TN8, TN10, TN20, and PA3) have been analyzed (Hamayun et al. 2010; Klink et al. 2007b, 2009c, 2010, 2011; Lu et al. 2006; Puthoff et al. 2007). To better understand the responses of SCN feeding on the soybean roots, multiple time points were analyzed for transcript profiling. For instance, the time points of 12 hours, 3, 6, 8, and 9 days post infection were chosen to study the gene expression of nematode-infected roots (Klink et al. 2009a). The syncytial cells, which serve as the SCN feeding sites and provide the nutrients to SCN, were collected as a specific tissue using laser capture microdissection (LCM). The precise dissection of syncytia made it possible for studying the genes specifically expressed in syncytia during a resistant and susceptible reaction (Klink et al. 2007a, 2009b, 2010, 2011). Gene expressions were also compared between incompatible and compatible responses for one single soybean germplasm line (Klink et al. 2007b). More recent studies compared the rapid and potent localized resistant reaction of two different soybean germplasm lines at the syncytia for different time points (Klink et al. 2011).

The results of these microarray analyses have provided an abundance of genetic information helping us to understand the mechanisms of soybean defense against SCN.

For example, induced or suppressed gene clusters for soybean response to SCN parasitism were identified. Two major groups of genes were identified during the interaction of soybean and SCN. One large proportion of these genes was associated with cell wall structure. Other functional categories consisted of a large number of up-regulated transcripts involved in defense and metabolism. With the completion of the soybean genome sequencing and the progress on plant defense mechanisms from other plants, many more candidate genes were annotated. For instance, Klink et al. (2009c) claimed that the genes encoding phenylalanine ammonia lyase, chalcone isomerase, isoflavone reductase, cinnamoyl-CoA reductase, and caffeic acid O-methyltransferase were SCN-resistance related genes.

Recently, soybean pathway analyses were performed for the interaction between soybean and SCN. For example, the jasmonic acid (JA) defense pathway was found to be involved in the localized resistance of *G. max* (PI 88788) against SCN (NL1-RHg/HG-type 7). In addition, the comparison of 9 to 6 days post infection of the resistant syncytium indicated that three genes were induced in the salicylic acid (SA) biosynthetic pathway (Klink et al. 2010). The phenylpropanoid pathway was also suggested to play a role at the local site of the syncytium during the resistance phase (Klink et al. 2009c).

Genetic engineering has also been utilized to study the functions of candidate SCN defense genes. For example, a stable transgenic soybean with an RNAi expression vector containing inverted repeats of a cDNA for the SCN major sperm protein (MSP) gene was shown to have a 68% reduction in eggs g^{-1} root tissue (Steeves et al. 2006).

To efficiently screen more functional genes from a large pool of candidate genes identified by microarray analyses, the transgenic hairy root system has been employed to study gene functions. For instance, a number of genes were shown to play a role in effective suppression of SCN infection through the RNAi strategy. These genes were the SCN homologs encoding the small ribosomal protein 3a and 4 (Hg-rps-3a and Hg-rps-4), synaptobrevin (Hg-snb-1), spliceosomal SR protein (Hg-spk-1) (Klink et al. 2009b), and the proteins related to nematode reproduction or fitness (Cpn-1, Y25, Prp-17, Fib-1 and Y25C1A.5) (Li et al. 2010a, 2010b).

2.1.3. The Salicylic Acid Pathway and its Role in Plant Defense Against Nematodes

The infection of plant tissues by biotrophic pathogens results in systemic induction of a long-lasting and broad-spectrum disease resistance, which is referred to as systemic acquired resistance (SAR) (Ryals et al. 1996). SA has been shown to play a critical role in the activation of local and systemic disease resistance following pathogen attack, such as potentiating host cell death and suppressing the dispersion of the pathogen (Dempsey et al. 1999). As described in the previous section, some microarray studies showed that the expression of some genes in the SA pathway was significantly changed during SCN parasitism.

To date, two pathways leading to the production of SA have been described. In one pathway, the biosynthesis of SA is derived from chorismate via isochorismate with a two-step process involving isochorismate synthase (ICS) and isochorismate pyruvate lyase (IPL). In the second pathway, SA is derived from phenylalanine, which first is

converted to cinnamic acid by phenylalanine ammonia lyase (PAL). Two possible routes have been proposed in the PAL pathway of SA biosynthesis. In the first route, cinnamic acid can be 2-hydroxylated to form O-coumaric acid and then decarboxylated to form SA. In the second route, the side chain of cinnamic acid is first decarboxylated to produce benzoic acid, which then is hydroxylated to form SA. The modification of SA can be continued to form a number of conjugated products including salicyloyl glucose ester (SGE), SA O-β-glucoside (SAG), methyl salicylate (MeSA), and SA-amino acid conjugates, which may function as the inactive storage forms to regulate the level of free SA *in vivo* (Vlot et al. 2009).

The downstream signaling pathway of SA, associated with pathogen defense, has been discovered. Non-expressor of pathogenesis-related genes 1 (NPR1) acts downstream of SA and is the critical component in the NPR1-dependent pathway. After a plant is infected, the SA level of the local infected sites increases. This change of SA level induces the dissociation of the NPR1 oligomers into monomers. The monomeric NPR1 molecules then can translocate from the cytosol into the nucleus, where they interact with the transcription factors of the TGA and WKRY families (especially WRKY70) to regulate expression of pathogenesis related (PR) genes in response to infection. Among these *PR* genes, *PR-1* gene expression is believed to be induced by SA accumulation (Pieterse and Van Loon 2004). Recently, methyl salicylate (MeSA), the methyl ester of SA, was shown to function as a critical SAR signal for tobacco (*Nicotiana tabacum*) defense against tobacco mosaic virus (TMV). The result of grafting experiments suggested that in inoculated leaves SA was converted to MeSA,

which was then transported to distant leaves via the phloem, and subsequently reconverted to active SA (Park et al. 2007).

There are two enzymes involved in the conversion between SA and MeSA: salicylic acid methyltransferase (SAMT) and methyl salicylate esterase. SAMT, belonging to the plant SABATH family of methyltransferases, catalyzes the transfer of a methyl group from S-adenosyl-L-methionine (SAM) to SA to form MeSA. *SAMT* genes have been isolated from a variety of plant species, such as *Clarkia breweri*, snapdragon (*Antirrhinum majus*), and *Stephanotis floribunda* (Negre et al. 2002; Pott et al. 2002; Ross et al. 1999). Later a *SAMT*, named *AtBSMT1*, was identified from *Arabidopsis thaliana* and shown to play a defense role (Chen et al. 2003). SA-binding protein 2 (SABP2) was first identified from tobacco as a protein (NtSABP2) of extremely low abundance yet containing a high affinity for SA. *NtSABP2* was cloned and its deduced protein sequence showed high similarity to α/β fold hydrolase. Further biochemical study demonstrated that NtSABP2 is a MeSA esterase whose function is to convert biologically inactive MeSA in the systemic tissue to active SA. The accumulated SA then activates or primes defenses leading to SAR. SABP2's esterase activity can be inhibited in the primary infected tissue by a high concentration of SA binding in its active site, which might benefit the accumulation of MeSA. *SABP2* was indicated to be vital to plant innate immunity, since the silencing of this gene restrained both local resistance to TMV and SAR development (Kumar and Klessig 2003).

Despite the numerous reports about SA signaling pathway in plant defense, the knowledge of the SA signaling pathway in soybean is still limited. The development of

SAR in soybean was first reported following infection with *Colletotrichum trancatum*, which causes anthracnose disease (Wrather and Elrod 1990). Application of a SAR inducer, either 2,6-dichloroisonicotinic acid (INA) or benzothiadiazole (BTH), was proven to be partially effective in reducing white mold disease incidence on field grown soybean (Dann et al. 1998). A soybean *PR-1* homolog, *GmPR-1*, was identified and its expression was induced by both SA treatment and SMV infection (Hajimorad and Hill 2001). Recently, Sandhu et al. (2009) reported SAR in soybean was regulated by two proteins, orthologous to Arabidopsis NPR1. As mentioned above, NPR1 is an important downstream signaling component from SA to activate SA-dependent defense (Pieterse and Van Loon 2004). Therefore soybean may share some conserved SA signaling pathway that confers resistance against pathogens.

In Arabidopsis, it was shown that successful cyst nematode parasitism may involve a local suppression of SA signaling in roots (Wubben et al. 2008). In tomato, the SA pathway was also studied for root knot nematode (RKN) resistance. The residual level of free SA was shown to be sufficient for basal resistance to RKN (Bhattarai et al. 2008). These results strongly suggested that the SA-mediated signaling pathway may play a critical role in limiting nematode parasitism during a compatible interaction.

2.1.4. Transgenic Hairy Root System for Studying Soybean-SCN Interaction

The transgenic hairy root system has been demonstrated to be useful for studying the interaction between root pathogen and soybean, since the traditional stable transformation of soybean is a time-consuming work. In the past ten years or so,

procedures for rapid establishment of transgenic hairy roots have been developed using *Agrobacterium rhizogenes*. This method produces “composite plants” comprising a transgenic hairy root system attached to non-transformed shoot system. *A. rhizogenes*-mediated transformation also makes it possible to co-transform plant cells with more than one T-DNA (transferred DNA) at the same time. The T-DNA containing the transgene of interest in a disarmed binary vector is generally co-transformed with the resident *A. rhizogenes* Ri (root-inducing) T-DNA containing the root locus (rol) genes (responsible for root proliferation). Compared with the traditional stable transformation, one major advantage of the transgenic soybean hairy root system is its shorter time of establishment. It takes about 10 months to produce a stable transgenic soybean line. In contrast, an ample number of independent soybean plants with transgenic hairy roots can be generated within only one month for many soybean cultivars. The hairy root system is especially useful for studying SCN since the SCN can complete its entire life cycle in the hairy root (Cho et al. 2000). Additionally, the transgenic hairy roots exhibit similar growth characteristics and morphology comparable to normal roots, which also benefits the study of plant-nematode interaction. Therefore, the high-throughput chimeric hairy root system can be used for rapid assessment of target genes in plants. Most of the studies using transgenic soybean for controlling SCN (**Table 1.2)** showed the reduction of SCN in transgenic hairy roots by RNAi-induced suppression of numerous genes essential for nematode development, reproduction, or parasitism. In this dissertation, I propose to explore the function of soybean candidate resistance genes using the transgenic hairy root system.

2.2. Soybean Insect Pests and Soybean Induced Defense

During growth, plants also suffer damage from herbivores in addition to pathogens. There is a large array of insect pests that attack soybean and they vary with regards to their strategies of attack (**Table 1.3)**. Usually based on the mouthpart type of the insect, their feeding habits can be divided into chewing and piercing-sucking types. These two feeding habits have different damage patterns. Chewing is a single instance of mechanical damage, while piercing-sucking is a continuous damage. Green stink bug (*Acrosternum hilare)*, soybean aphid (*Aphis glycines*), two-spotted spider mite (*Tetranychus urticae*), and garden fleahopper (*Halticus bractatus)* are all piercing-sucking insect pests (**Table 1.3**). Stink bug is considered to be the most important insect pest that attacks soybean. There are also some beneficial insects of soybean such as pirate bug (*Orius sp.*), spotted lady beetle (*Coleomegilla maculata*), which are actually predators of some soybean pests.

Plant defense against herbivore attack generally has two strategies: direct defense with repellents, toxins, and deterrents, and indirect defense through attracting the natural enemies of the feeding insects. Plants are able to produce and emit a mixture of volatile organic compounds (VOCs) as both constitutive and induced response to herbivore feeding. The induced defense usually requires a complex recognition and signaling system to control and coordinate, in which the plant hormones including JA and SA are involved (Arimura et al. 2005).

The study of herbivore-induced plant volatiles (HIPVs) suggested that HIPVs can attract carnivorous enemies of herbivores to work as the agents of plant induced defense.

It is generally assumed that the benefit of inducible defense over constitutive defense is cost reduction, especially in terms of biosynthetic costs (Dicke and Baldwin 2010). HIPVs derive from diverse biosynthetic pathways, including (1) octadecanoid pathway producing fatty-acid derived green leaf volatiles (GLVs) (e.g., hexanal, (*Z*)-3-hexen-1-ol and (*Z*)-3-hexenyl acetate); (2) methyl-erythritol phosphate (MEP) pathway for synthesis of monoterpenes and diterpenes; (3) mevalonate pathway for production of sesquiterpenes; (4) shikimate pathway giving rise to various aromatic metabolites (e.g., methyl salicylate and indole). Among the HIPVs, volatile terpenoids consist of monoterpenes (containing 10 carbons, C10), sesquiterpenes (C15) and homoterpenes (C11 or C16) (e.g., (E)-β-ocimene, (E,E)-α-farnesene, and (E)-4,8-dimethyl-1,3,7-nonatriene) (Lucas-Barbosa et al. 2011; Mumm and Dicke 2010). In my dissertation, I propose to study the herbivore-induced soybean volatiles, mainly focused on volatile terpenoids.

.

3. Soybean Meal and its Problems

Besides producing edible oil, soybeans have many other uses, one of which is soybean meal. Soybean meal as a livestock feed also needs some improvement, given it contains anti-nutritional factors (Hsiao et al. 2006). A number of secondary compounds produced in soybean meal are classified into heat-labile and heat-stable types. The heat-labile factors, including trypsin inhibitor and lectins, can be removed or decreased by heat treatment during processing soybean meal. Heat-stable secondary compounds include non-starch polysaccharides, saponins, phytate, phytoestrogens, and protein antigens,

which must be removed or reduced through solvent extraction, fractionation, or use of exogenous enzymes. These anti-nutritional factors reduce nutrient absorption, thereby influencing the growth of livestock and poultry at different levels (Drew et al. 2007).

Soybean breeders have been working on developing soybean cultivars with the characters of disease resistance, high yield, or improved quality of soybean products (Hartwig and Epps 1973; Jin et al. 2010; Keisuke 1993). To resolve the above problems, genetic modification is one option. For example, genetically modified soybeans are being used in an increasing number of products. Roundup Ready (RR) soybeans genetically modified to be resistant to the herbicide Roundup was introduced by Monsanto in 1995 through substitution of the *Agrobacterium sp.* (strain CP4) EPSP (5-enolpyruvyl shikimic acid-3-phosphate) synthase gene (Padgette et al. 1995). To genetically improve soybean, it is important to fully understand the genetic and molecular basis controlling the various soybean traits related with defense, yield, and quality. In this dissertation, the soybean natural defense mechanism against diseases is investigated, as well as the genes used for soybean meal improvement.

4. The Soybean Gene Discovery with Functional Genomics Approach

4.1. Genomic Study of Soybean

Large-scale shotgun sequencing of soybean (*G. max* cv. Williams 82) was completed in 2008. It was reported that the soybean genome contained 950 megabases of sequence with predicted 46,430 protein coding genes distributed over twenty chromosomes. It was speculated that the ancestors of soybean underwent two whole genome duplications

that were estimated to occur 59 million years ago and 13 million years ago (Schmutz et al. 2010). After the two duplication events the evolution of the soybean genome continued in forms of gene diversification, gene loss, and abundant chromosome rearrangements. In addition, long-term agricultural soybean breeding increases the complexity of the soybean genome. Therefore, to gain a more comprehensive physical map, it is necessary to combine the present genetic map and locate the genes on the soybean chromosomes.

The next step of a genomic study is to identify the functions of the genes, since the first step of sequencing has been finished. Recently, researchers verified the applicability of transposons in studying gene functions of soybean (Mathieu et al. 2009). Transposons are mobile genetic elements, which lead to mutations and disruption in gene functions. By understanding transposons, researchers can screen plants with important traits, such as seed composition and resistance. For instance, considering the feature of soybean nitrogen fixation, it is believed that soybeans might contain a certain number of distinct genes, not in rice or maize, whose functions are allowing the rhizobium microbes to enter the soybean roots to form symbiosis and fix nitrogen.

Therefore, functional genomic studies of soybean will speed up the study of soybean-pathogen interaction and soybean meal improvement, from both the practical as well as the fundamental points of view. For instance, identification of molecular markers for marker-assisted breeding and isolation of resistance genes are some of the most important and immediate practical benefits of genomics. To this end, soybean

genome information should help pave the way for improving soybean and other legume plants.

4.2. Soybean Genomics and its Implication to Soybean Genetic Improvement

Various genomic tools and database information are available to help researchers study genes function. For example, National Centre for Biotechnology Information (NCBI) (http:// www.ncbi.nlm.nih.gov) is a primary source for DNA and protein sequences of all organisms. Sequence similarity searches, such as with the Basic Local Alignment Search Tool (BLAST), can be conducted using a known sequence to find the closest matches to all other organisms. SoyBase (http://soybase.ncgr.org) provides the latest news about the soybean and links to other soybean and legume websites. Another database, the Soybean Genomics and Microarray Database (SGMD), provides an integrated view of the interaction of soybean with the SCN to help us to establish the correlation of soybean ESTs with their gene expression profiles (Matthews et al. 2009). All these database and genomic tools allow researchers to screen, validate, and identify putative genes.

A comparative functional genomics approach has been demonstrated to be a powerful method to identify target genes. Since functionally important regions in a genome sequence tend to evolve slower than less important regions, functional regions are believed to be under stabilizing selection and should be preferentially conserved during the evolutionary process. Comparative functional genomics seeks to compare the genomic structure of evolutionarily related and unrelated species for predicting gene

structures and the number of active genes in the species. Comparative analysis can also be carried out by searching for conservation among genome sequences in both the protein coding sequences and the functional non-coding sequences in a genome (Ureta-Vidal et al. 2003). Availability of genome and proteome sequences from many species has sped up the study of comparative genomics.

Understanding of the molecular basis is vital for genetic improvement of soybean. In my dissertation, I propose to use a functional genomics approach to identify soybean genes involved in soybean resistance against pests as well as soybean genes for improving soybean meal quality.

5. Goal and Specific Objectives of the Dissertation

The main goal of this dissertation is to investigate and elucidate the soybean defense mechanisms. In this dissertation, I explore the function of the SA-dependent defense pathway in the resistance against SCN, a below-ground obligate root pathogen which causes severe damage on soybean. The specific objective for this part of the work was to identify candidate SCN resistance genes and evaluate their biological functions. Particularly, I use a comparative functional genomics approach to identify and characterize the soybean salicylic acid methyltransferase gene (*GmSAMT1*) and the salicylic acid binding protein 2 gene (*GmSABP2-1*). Since soybean also suffers from insect pests, I also explore the role of the herbivore-induced plant volatiles in soybean defense against above-ground herbivore. My research objective was to discover which volatiles were induced in soybean under attack of a piercing-sucking insect pest

(two-spotted spider mite). I then identify terpene synthase genes that are responsible for producing herbivore-induced terpenoid volatiles.

Besides the soybean defense mechanism, an additional goal of this dissertation is to identify and characterize the soybean mannanase (GmMAN1) gene that can be used for soybean meal improvement.

REFERENCES

Alvarez E, Braun EJ, McGee DC (1995) New assays for detection of *Pseudomonas syringae* pv. *glycinea* in soybean seed. Plant Dis. 79: 12-14

Arimura G-i, Kost C, Boland W (2005) Herbivore-induced, indirect plant defences. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids 1734: 91-111

Bhattarai KK, Xie Q-G, Mantelin S, Bishnoi U, Girke T, Navarre DA, Kaloshian I (2008) Tomato susceptibility to root-knot nematodes requires an intact jasmonic acid signaling pathway. Mol. Plant Microbe In. 21: 1205-1214

Chen F, D'Auria JC, Tholl D, Ross JR, Gershenzon J, Noel JP, Pichersky E (2003) An *Arabidopsis thaliana* gene for methylsalicylate biosynthesis, identified by a biochemical genomics approach, has a role in defense. Plant J. 36: 577-588

Cho HJ, Farrand SK, Noel GR, Widholm JM (2000) High-efficiency induction of soybean hairy roots and propagation of the soybean cyst nematode. Planta 210: 195-204

Concibido VC, Diers BW, Arelli PR (2004) A decade of QTL mapping for cyst nematode resistance in soybean. Crop Sci. 44: 1121-1131

Dann E, Diers B, Byrum J, Hammerschmidt R (1998) Effect of treating soybean with 2,6-dichloroisonicotinic acid (INA) and benzothiadiazole (BTH) on seed yields and the level of disease caused by *Sclerotinia sclerotiorum* in field and greenhouse studies. Eur. J. Plant Pathol. 104: 271-278

Davis EL, Mitchum MG (2005) Nematodes. Sophisticated parasites of legumes. Plant Physiol. 137: 1182-1188

Dempsey D, Shah J, Klessig D (1999) Salicylic acid and disease resistance in plants. Crit. Rev. Plant Sci. 18: 547-575

Dicke M, Baldwin IT (2010) The evolutionary context for herbivore-induced plant volatiles: beyond the 'cry for help'. Trends Plant Sci. 15: 167-175

Drew MD, Borgeson TL, Thiessen DL (2007) A review of processing of feed ingredients to enhance diet digestibility in finfish. Anim. Feed Sci. Tech. 138: 118-136

Elling AA, Davis EL, Hussey RS, Baum TJ (2007) Active uptake of cyst nematode parasitism proteins into the plant cell nucleus. Int. J. Parasitol. 37: 1269-1279

Endo B (1991) Ultrastructure of initial responses of resistant and susceptible soybean

roots to infection by *Heterodera glycines*. Rev. Nematol. 14: 73-94

Giesler LJ, Ghabrial SA, Hunt TE, Hill JH (2002) Bean pod mottle virus: A threat to U. S. soybean production. Plant Dis. 86: 1220-1289

Golden A, Epps J, Riggs R, Duclos L, Fox J, Bernard R (1970) Terminology and identity of intraspecific forms of the soybean cyst nematode (*Heterodera glycines*). Plant Dis. Rep. 54: 544-546

Hajimorad MR, Hill JH (2001) Rsv1-mediated resistance against soybean mosaic virus-N is hypersensitive response-independent at inoculation site, but has the potential to initiate a hypersensitive response-like mechanism. Mol. Plant Microbe In. 14: 587-598

Hamayun M, Khan SA, Khan AL, Shinwari ZK, Hussain J, Sohn EY, Kang SM, Kim YH, Khan MA, Lee IJ (2010) Effect of salt stress on growth attributes and endogenous growth hormones of soybean cultivar Hwangkeumkong. Pak J. Bot. 42: 3103-3112

Hartwig E, Epps J (1973) Registration of 'Forrest' soybeans. Crop Sci. 13 : 287

Hsiao H-Y, Anderson DM, Dale NM (2006) Levels of beta-mannan in soybean meal. Poult. Sci. 85: 1430-1432

Jin J, Liu X, Wang G, Mi L, Shen Z, Chen X, Herbert SJ (2010) Agronomic and physiological contributions to the yield improvement of soybean cultivars released from 1950 to 2006 in Northeast China. Field Crop Res. 115: 116-123

Jones MGK (1981) Host cell responses to endoparasitic nematode attack: structure and function of giant cells and syncytia. Ann. Appl. Biol. 97: 353-372

Keisuke K (1993) Breeding trials for improving the food-processing quality of soybeans. Trends Food Sci. Tech. 4: 64-67

Klink V, Hosseini P, MacDonald M, Alkharouf N, Matthews B (2009a) Population-specific gene expression in the plant pathogenic nematode *Heterodera glycines* exists prior to infection and during the onset of a resistant or susceptible reaction in the roots of the *Glycine max* genotype Peking. Bmc Genomics 10: 111

Klink V, Kim K-H, Martins V, MacDonald M, Beard H, Alkharouf N, Lee S-K, Park S-C, Matthews B (2009b) A correlation between host-mediated expression of parasite genes as tandem inverted repeats and abrogation of development of female *Heterodera glycines* cyst formation during infection of *Glycine max*. Planta 230: 53-71

Klink VP, Hosseini P, Matsye P, Alkharouf NW, Matthews BF (2009c) A gene

expression analysis of syncytia laser microdissected from the roots of the *Glycine max* (soybean) genotype PI 548402 (Peking) undergoing a resistant reaction after infection by *Heterodera glycines* (soybean cyst nematode). Plant Mol. Biol. 71: 525-567

Klink VP, Hosseini P, Matsye PD, Alkharouf NW, Matthews BF (2010) Syncytium gene expression in *Glycine max*([PI 88788]) roots undergoing a resistant reaction to the parasitic nematode *Heterodera glycines*. Plant Physiol. Bioch. 48: 176-193

Klink VP, Hosseini P, Matsye PD, Alkharouf NW, Matthews BF (2011) Differences in gene expression amplitude overlie a conserved transcriptomic program occurring between the rapid and potent localized resistant reaction at the syncytium of the *Glycine max* genotype Peking (PI 548402) as compared to the prolonged and potent resistant reaction of PI 88788. Plant Mol. Biol. 75: 141-165

Klink VP, Overall CC, Alkharouf NW, MacDonald MH, Matthews BF (2007a) Laser capture microdissection (LCM) and comparative microarray expression analysis of syncytial cells isolated from incompatible and compatible soybean (*Glycine max*) roots infected by the soybean cyst nematode (*Heterodera glycines*). Planta 226: 1389-1409

Klink VP, Overall CC, Alkharouf NW, MacDonald MH, Matthews BF (2007b) A time-course comparative microarray analysis of an incompatible and compatible response by *Glycine max* (soybean) to *Heterodera glycines* (soybean cyst nematode) infection. Planta 226: 1423-1447

Kulkarni SS, Bajwa SG, Rupe JC, Kirkpatrick TL (2008) Spatial correlation of crop response to soybean cyst nematode (*Heterodera glycines*). T. Asabe 51: 1451-1459

Kumar D, Klessig DF (2003) High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity. Proc. Natl. Acad. Sci. U.S.A. 100: 16101-16106

Li J, Todd T, Oakley T, Lee J, Trick H (2010a) Host-derived suppression of nematode reproductive and fitness genes decreases fecundity of *Heterodera glycines* Ichinohe. Planta 232: 775-785

Li J, Todd T, Trick H (2010b) Rapid in planta evaluation of root expressed transgenes in chimeric soybean plants. Plant Cell Rep. 29: 113-123

Lu P, Shannon JG, Sleper DA, Nguyen HT, Cianzio SR, Arelli R (2006) Genetics of cyst nematode resistance in soybean PIs 467312 and 507354. Euphytica 149: 259-265

Lucas-Barbosa D, van Loon JJA, Dicke M (2011) The effects of herbivore-induced plant volatiles on interactions between plants and flower-visiting insects. Phytochemistry 72: 1647-1654

Mathieu M, Winters E, Kong F, Wan J, Wang S, Eckert H, Luth D, Paz M, Donovan C, Zhang Z, Somers D, Wang K, Nguyen H, Shoemaker R, Stacey G, Clemente T (2009) Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository. Planta 229: 279-289

Matthews D, Lazo G, Anderson O (2009) Plant and crop databases. Methods Mol. Biol. 513: 243-262

Melito S, Heuberger AL, Cook D, Diers BW, MacGuidwin AE, Bent AF (2010) A nematode demographics assay in transgenic roots reveals no significant impacts of the Rhg1 locus LRR-Kinase on soybean cyst nematode resistance. Bmc Plant Biol. 10: 104

Mumm R, Dicke M (2010) Variation in natural plant products and the attraction of bodyguards involved in indirect plant defense. Can. J. Zool. 88: 628-667

Negre F, Kolosova N, Knoll J, Kish CM, Dudareva N (2002) Novel S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase, an enzyme responsible for biosynthesis of methyl salicylate and methyl benzoate, is not involved in floral scent production in snapdragon flowers. Arch. Biochem. Biophys. 406: 261-270

Niblack T, Arelli P, Noel G, Opperman C, Orf J, Schmitt D, Shannon J, Tylka G (2002) A revised classification scheme for genetically diverse populations of *Heterodera glycines*. J. Nematol. 34: 279-288

Opperman CH, Bird DM (1998) The soybean cyst nematode, *Heterodera glycines*: a genetic model system for the study of plant-parasitic nematodes. Curr. Opin. Plant Biol. 1: 342-346

Padgette S, Kolacz K, Delannay X, Re D, LaVallee B, Tinius C, Rhodes W, Otero Y, Barry G, Eichholz D, Peschke V, Nida D, Taylor N, Kishore G (1995) Development, identification, and characterization of a glyphosate-tolerant soybean line. Crop Sci. 35: 1451-1461

Park SW, Kaimoyo E, Kumar D, Mosher S, Klessig DF (2007) Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. Science 318: 113-116

Pieterse CMJ, Van Loon LC (2004) NPR1: the spider in the web of induced resistance signaling pathways. Curr. Opin. Plant Biol. 7: 456-464

Pott MB, Pichersky E, Piechulla B (2002) Evening specific oscillations of scent emission, SAMT enzyme activity, and SAMT mRNA in flowers of *Stephanotis floribunda*. J. Plant Physiol. 159: 925-934

Puthoff DP, Ehrenfried ML, Vinyard BT, Tucker ML (2007) GeneChip profiling of transcriptional responses to soybean cyst nematode, *Heterodera glycines*, colonization of soybean roots. J Exp Bot 58: 3407-3418

Ross JR, Nam KH, D'Auria JC, Pichersky E (1999) S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase, an enzyme involved in floral scent production and plant defense, represents a new class of plant methyltransferases. Arch. Biochem. Biophysi. 367: 9-16

Ryals J, Neuenschwander U, Willits M, Molina A, Steiner H, Hunt M (1996) Systemic acquired resistance. Plant Cell 8: 1809-1819

Sandhu D, Tasma IM, Frasch R, Bhattacharyya MK (2009) Systemic acquired resistance in soybean is regulated by two proteins, orthologous to Arabidopsis NPR1. Bmc Plant Biol. 9: 105

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X-C, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178-183

Steeves RM, Todd TC, Essig JS, Trick HN (2006) Transgenic soybeans expressing siRNAs specific to a major sperm protein gene suppress *Heterodera glycines* reproduction. Funct. Plant Biol. 33: 991-999

Ureta-Vidal A, Ettwiller L, Birney E (2003) Comparative genomics: genome-wide analysis in metazoan eukaryotes. Nat. Rev. Genet. 4: 251-262

Vaghchhipawala ZE, Schlueter JA, Shoemaker RC, Mackenzie SA (2004) Soybean FGAM synthase promoters direct ectopic nematode feeding site activity. Genome 47: 404-413

Vlot AC, Dempsey DA, Klessig DF (2009) Salicylic acid, a multifaceted hormone to combat disease. Annu. Rev. Phytopathol. 47: 177-206

Wilson R (2008) Soybean: market driven research needs. In: Stacey G (ed) Genetics and genomics of soybean. New York, NY : Springer Science+Business Media, LLC.

Winter SMJ, Rajcan I, Shelp BJ (2006) Soybean cyst nematode: Challenges and opportunities. Can. J. Plant Sci. 86: 25-32

Wrather J, Elrod J (1990) Apparent systemic effect of *Colletotrichum truncatum* and *C. lagenarium* on the interaction between soybean and *C. truncatum*. Phytopathology 80: 472-447

Wrather JA, Koenning SR (2006) Estimates of disease effects on soybean yields in the United States 2003 to 2005. J. Nematol. 38: 173-180

Wubben MJE, Jin J, Baum TJ (2008) Cyst nematode parasitism of *Arabidopsis thaliana* is inhibited by salicylic acid (SA) and elicits uncoupled SA-independent pathogenesis-related gene expression in roots. Mol. Plant Microbe In. 21: 424-432

Appendix

Table 1.1 Soybean pathogens and their resulting diseases.

Soybean Disease Category	**Disease**	**Pathogen Category**	**Pathogen Name**
Leaf and Stem Diseases	Bacterial Blight	Bacterium	*Pseudomonas syringae* pv. *glycinea*
	Brown Spot	Fungus	*Septoria glycines*
	Pod and Stem Blight	Fungi	*Diaporthe phaseolorum* var. *sojae* and *Phomopsis longicola*
	Downy Mildew	Fungus	*Peronospora manshurica*
Root and Lower Stem Diseases	Phytophthora Rot	Fungus	*Phytophthora sojae*
	Rhizoctonia stem rot	Fungus	*Rhizoctonia solani*
	Sclerotinia Stem Rot	Fungus	*Sclerotinia sclerotiorum*
	Brown Stem Rot	Fungus	*Phialophora gregata*
	Sudden Death Syndrome	Fungus	*Fusarium solani* race A
	Charcoal Root Rot	Fungus	*Macrophomina phaseolina*
	Soybean Cyst Nematode	Nematode	*Heterodera glycines*
Virus Diseases	Bud Blight	Virus	Tobacco Ringspot Virus
	Soybean Mosaic	Virus	Soybean Mosaic Virus

Table 1.1 was developed from the website
http://www.btny.purdue.edu/extension/pathology/cropdiseases/soybean/Soybean.html

Table 1.2 Transgenic soybeans for controlling SCN.

Gene Name／ Genbank Accession no.	Putative Function	Transgenic Form	Reduction Level of SCN	Reference
MSP	Major sperm protein	Stable transgenic soybean plant	Up to 68% reduction in nematode eggs	(Steeves et al. 2006)
Hg-rps-3a, CB379877	Ribosomal protein 3a	Transgenic hairy roots	87% reduction in female cysts	(Klink et al. 2009)
Hg-spk-1, BI451523.1	Spliceosomal SR protein	Transgenic hairy roots	88% reduction in female cysts	(Klink et al. 2009)
Hg-snb-1, BF014436	Synaptobrevin	Transgenic hairy roots	93% reduction in female cysts	(Klink et al. 2009)
Y25, CB824330	Beta subunit of the COPI complex	Transgenic hairy roots	81% reduction in nematode eggs	(Li et al. 2010)
Prp-17, AF113915	Pre-mRNA splicing factor	Transgenic hairy roots	79% reduction in nematode eggs	(Li et al. 2010)
Cpn-1, GU074018	Unknown protein	Transgenic hairy roots	95% reduction in nematode eggs	(Li et al. 2010)
LRR-kinase, Glyma18g02680.1	leucine-rich repeat transmembrane receptor-kinase	Transgenic hairy roots	No detectable change	(Melito et al. 2010)

Table 1.3 The growth stages and infected tissues of soybean subject to attack by pests listed by feeding habit.

	Soybean Growth Stage			Soybean Injury Tissue					Feeding Habit	
	S	**V**	**R**	**Leaf**	**Pod**	**Root**	**Seed**	**Stem**	**Chewing**	**Piercing-Sucking**
Bean leaf beetle	+	+	+	+	+				+	
Cutworms and armyworms	+	+		+					+	
Imported longhorned weevil	+	+		+					+	
Seedcorn maggot	+						+		+	
Soybean leaf miner	+	+		+					+	
White grubs	+					+			+	
Blister beetles		+	+	+					+	
Colaspis beetles		+		+					+	
Dectes stem borer		+						+	+	
Garden fleahopper		+		+						+
Green cloverworm		+	+	+					+	
Japanese beetle		+	+		+				+	
Soybean aphid		+	+	+						+
Soybean looper		+		+					+	
Stalk borer		+	+					+	+	
Thistle caterpillar		+		+					+	
Stink bugs			+		+					+
Two-spotted spider mite			+	+						+
Grasshoppers		+	+		+				+	

S, V, and R represent the seedling, vegetative, and reproductive stages of soybean growth.

The information in Table 1.3 is adapted from the website http://www.ent.iastate.edu/soybeaninsects/.

Chapter II. Salicylic Acid Methyltransferase Confers Resistance to Soybean Cyst Nematode in Soybean

Jingyu Lin, Mitra Mazarei, Nan Zhao, Mohammad R. Hajimorad, Junwei Zhu, Vincent R. Pantalone, Prakash R. Arelli, Charles Neal Stewart, Jr, Feng Chen. Salicylic Acid Methyltransferase Confers Resistance to Soybean Cyst Nematode in Soybean. Drafted

ABSTRACT

Soybean cyst nematode (*Heterodera glycines*, SCN) is the most severe pest of soybean [*Glycine max* (L.) Merr.]. Through prior microarray analysis on the gene expression of two genetically-related soybean lines (recombinant inbred lines, RILs) with different responses to SCN, the soybean gene *Glyma02g06070*, was identified to be SCN-induced in the resistant line but not in the susceptible line. From sequence data this gene was predicted to encode an S-adenosyl-L-methionine (SAM) dependent carboxyl methyltransferase. Based on biochemical activity of *Escherichia coli*-expressed protein encoded by this gene, *Glyma02g06070* was identified to be salicylic acid methyltransferase gene (*GmSAMT1*). GmSAMT1 catalyzed the conversion of salicylic acid to methyl salicylate (MeSA). To determine the biological role of GmSAMT1 in soybean resistance against SCN, we generated transgenic hairy roots in the susceptible soybean background with a binary vector harboring cassettes for the over-expression of *GmSAMT1* and red fluorescent protein (RFP) gene. After two weeks post inoculation by SCNs the demographics of SCN in over-expressed *GmSAMT1* transgenic hairy roots and control hairy roots were compared. The result showed the transgenic hairy roots with over-expression of *GmSAMT1* in the susceptible soybean line exhibited increased resistance, similar to that of the resistant soybean line. To further understand its defense role in soybean, we also compared the expression level of *GmSAMT1* in the compatible and incompatible reactions between soybean and soybean mosaic virus. The expression of *GmSAMT1* was found to be up-regulated in the incompatible reaction, and not the compatible reaction. Taken together, the results

presented in this study suggest that *GmSAMT1* play a critical role in soybean defense against SCN and soybean mosaic virus.

KEYWORDS: salicylic acid methyltransferase, soybean cyst nematode, defense, transgenic hairy root

INTRODUCTION

Soybean cyst nematode (*Heterodera glycines*, SCN) is the most devastating pest of soybean [*Glycine max* (L.) Merr.] and is a sedentary obligate endoparasite of roots. When soybean plants are severely damaged by SCN, they exhibit stunted and chlorotic phenotype (Kulkarni et al. 2008). These symptoms are easy to be confused with nutrient deficiency and water stress. Until the population of SCN reaches a critical level, its damaging effect cannot be clearly identified by farmers. Currently, the most accurate way of diagnosing SCN infection is to analyze soil samples and observe the pest directly. Controlling SCN in commercial soybean production still remains difficult. Because the SCN life cycle is short, its population can build up rapidly. Frequent changes in population virulence also contribute to the difficulty in management. In addition, the cysts can survive in the soil for up to nine years and then break to release the eggs under proper conditions, which also enables easy spread via infested soil. The methods used to control and manage SCN in soybean production include crop rotation, the use of SCN-resistant cultivars, and the application of nematicides (Winter et al. 2006).

A number of microarray analyses have been performed to study gene expression changes of soybean when infected by SCN. Considering the complexity of soybean resistance, many different soybean lines (such as PI 88788, PI 467312, PI 507354, Williams 82, Peking, and Kent) and SCN populations (NL1-RHg, TN8, TN10, TN20, and PA3) have been analyzed (Khan et al. 2004; Klink et al. 2007, 2009b, 2010, 2011; Lu et al. 2006; Puthoff et al. 2007). These microarray analyses have given abundant gene regulatory information to help researchers to understand the mechanisms of

soybean defense against SCN. For example, some induced or suppressed soybean gene clusters that respond to SCN infection have been identified (Klink et al. 2009). One study identified a soybean methyltransferase gene that was up-regulated in the resistant soybean line whereas it was not significant changed in the susceptible soybean line during compatible interaction between SCN and soybean (Mazarei et al. 2011).

In *Arabidopsis thaliana*, it was shown that the successful beet cyst nematode (*Heterodera schachtii*) parasitism may involve the local suppression of salicylic acid (SA) signaling in roots (Wubben et al. 2008). In tomato, the SA pathway was also studied for root knot nematode (RKN) resistance. The residual level of free SA was shown to be sufficient for basal resistance to RKN (Bhattarai et al. 2008). These results strongly suggest that the SA-mediated signaling pathway plays a critical role in limiting nematode parasitism during a compatible interaction. However its function in soybean is not clear. Studying the SA pathway in soybean defense should provide more comprehensive information to understand the soybean defense mechanism.

The development of systemic acquired resistance (SAR) in soybean was first reported following infection with *Colletotrichum trancatum*, which causes anthracnose disease (Wrather and Elrod 1990). Application of SAR inducers, 2,6-dichloroisonicotinic acid (INA) and benzothiadiazole (BTH), were shown be partially effective to reduce white mold disease incidence on field grown soybean (Dann et al. 1998). A soybean pathogenesis related (PR) gene, *GmPR-1*, was identified and its expression was induced by both SA treatment and soybean mosaic virus (SMV) infection (Hajimorad and Hill 2001). This gene was used as a SAR marker in our study.

Recently, Sandhu et al. (2009) reported SAR in soybean was regulated by non-expressor of pathogenesis-related genes 1 proteins, GmNPR1-1 and GmNPR1-2. NPR1 was identified to be an important signaling component downstream of SA to activate SA-dependent defense (Pieterse and Van Loon 2004). Through soybean (*G. max* cv. Williams 82) cell cultures, the function of SA was also studied during the soybean's response to *Pseudomonas syringae*, which was expressing the corresponding avirulence gene *AvrA*. In the presence of low concentrations of SA (50 μM), the cells were able to go through a programmed cell death and induce many related genes in the hypersensive reaction (Ludwig and Tenhaken 2000). Taken together, these findings indicate that soybean might share a conserved SA signaling pathway that confers resistance against pathogens.

Methyl salicylate (MeSA), a volatile ester of SA, has been shown to be a critical mobile signal for SAR in tobacco (*Nicotiana tabacum*) (Park et al. 2007). In prior studies, the emission of MeSA, was not detectable in soybean leaf tissue under normal conditions, which was the same at that found in the previous study in Arabidopsis (Chen et al. 2003a), Under the infestation of certain insects, such as soybean aphid (*Aphis glycines*) and herbivorous pentatomid bug *Euschistus heros*, the emission of methyl salicylate was found to be induced in soybean (Zhu and Park, 2005; Michereff et al. 2011). The association of MeSA emission and infestation from soybean pests provides a clue for understanding soybean defense mechanism as well.

The biosynthesis of MeSA is catalyzed by an enzyme, S-adenosyl-L-methionine (SAM)-dependent salicylic acid carboxyl methyltransferase (SAMT), which catalyzes

the reaction of SA and the methyl donor SAM to form MeSA. Some *SAMT* genes have been cloned and characterized in some species and their functions have been shown to be related with plant defense. For instance, tobacco plant with a silenced *SAMT* gene, *NtSAMT1*, was shown to have reduced level of MeSA. *NtSAMT1* was also shown to be required for generating the mobile signal of SAR (Park et al. 2007).

Here, we identify and characterize one soybean salicylic acid methyltransferase gene, *GmSAMT1*. To verify its role of resistance in soybean, a modified vector system for rapid screening *Agrobacterium rhizogenes*-mediated transgenic hairy roots with over-expression of *GmSAMT1* in the susceptible soybean line was developed. A nematode demographic assay was used to compare the resistant response between transgenic line and control line by monitoring the infecting population for progression through nematode life stages. To further understand function of *GmSAMT1* in defense against virus, its expression in soybean mosaic virus infected soybean was also examined.

MATERIALS AND METHODS

Plant, Nematode, Bacterium, and Chemical Sources. In this study, two genetically-related soybean lines, TN02-226 and TN02-275, were used for gene cloning and generating soybean hairy roots. The breeding details for these two soybean lines have been described elsewhere (Mazarei et al. 2011). These two F6-derived sister lines, TN02-226 and TN02-275, are resistant and susceptible, respectively, to the SCN race 2 (HG type 1.2.5.7), which has been confirmed in USDA Southern Regional Tests from

2004 through 2007 (Paris and Shelton 2005, 2006; Gillen and Shelton 2007, 2008). In this study, the SCN race 2 eggs were utilized as the pathogens in the bioassay of the transgenic hairy roots since the above two soybean lines exhibited different responses to SCN race 2. The maintenance of SCN followed the method described before (Arelli et al. 2000). *Escherichia coli* strain BL21 (DE3) CodonPlus (Stratagene, La Jolla, CA) was used for the expression of recombinant protein. *Agrobacterium* cultures, *A. tumefaciens* GV3101 and *A. rhizogenes* strain K599, were utilized for tobacco transient transformation and the generation of the soybean hairy roots. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

Viruses, Soybean Genotypes, Inoculation, and SMV Detection. In this study, a mutant of soybean mosaic virus SMV-N (SMV-N25) was used for virus infection on soybean. Two soybean genotypes Williams 82 (Bernard et al. 1991), susceptible to SMV-N25, and L78-375, resistant to SMV-N25 (Hajimorad and Hill, 2001; Eggenberger et al. 2008), were used in this study. The mechanically inoculated plants were maintained in a growth chamber operating at 22 °C with a photoperiod of 16 h. SMV inoculated soybean seedlings were grown for another 21 days and their trifoliate leaves were harvested, pulverized in the presence of liquid nitrogen and stored at -80 °C until processed for quantitative reverse-transcription PCR analysis.

Database Search and Sequence Analysis. The sequence of Affymetrix probe Gma.12911.1.A1_S_AT, which identified our target gene in microarray experiment, was

used as a query to BLAST search against the soybean database GmGDB/GMtranscript (http://www.plantgdb.org/GmGDB/cgi-bin/blastGDB.pl). Program blastn was performed with E-value of 1e-20 (Altschul et al. 1997). Five soybean genes showed high similarity (**Figure 2.1)**. A multiple sequence alignment of selected SABATH proteins and Glyma02g06070 was performed using ClustalX program and viewed using TreeView software (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html). The neighbor-joining unrooted phylogenetic tree was constructed (**Figure 2.2)**.

Isolation of Glyma02g06070 Gene from Soybean. Total RNA was isolated from the SCN-infected the root tissue of soybean line TN02-226, using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) and DNA contamination was removed with an on-column DNase (Qiagen, Valencia, CA, USA) treatment. Then cDNA was synthesized from total RNA in a 15 µl reaction volume using the First-Strand cDNA Synthesis Kit (GE Healthcare, Piscataway, NJ, USA) as previously described (Chen et al. 2003b). Based on the previous microarray analysis, one soybean candidate gene *Glyma02g06070* was chosen for further study. The full length cDNA sequence of Glyma02g06070 was gained from public soybean database (http://www.phytozome.net/soybean), and primers were designed as following 5'-ATGGAAGTAGCACAGGTACTCCACATG-3' and 5'-TGCTTTTCTAGTCAATAATATGGTAAC-3'. The PCR conditions were as follows: 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 1 min 30 s, and a final extension at 72 °C for 10 min. PCR products were cloned into

vector pEXP5/CT-TOPO. By sequencing, the PCR product from TN02-226 soybean was the same as the one predicted in the soybean database.

To verify its biochemical function, the candidate gene was then cloned into a protein expression construct vector, pET100/D-TOPO vector (Invitrogen, Carlsband, CA) and expressed in the *E. coli* strain BL21 (DE3) CodonPlus (Stratagene, La Jolla, CA). Protein expression was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) at a concentration of 500 μM for 18 h at 25 °C, with cells lysed by sonication. *E. coli*-expressed Glyma02g06070 with a His-tag was purified from the *E. coli* cell lysate using Ni-NTA agarose following the manufacturer instructions (Invitrogen). Protein purity was verified by SDS–PAGE and protein concentrations were determined by the Bradford assay (Bradford, 1976).

Radiochemical SAMT Activity Assay. Radiochemical SAMT assays were performed according to Zhao et al. (2007). A 50 μl volume containing 50 mM Tris-HCl, pH 7.5, 1 mM SA, and 3 μM ^{14}C-SAM with a specific activity of 51.4 mCi/mmol (Perkin-Elmer, Boston, MA) was used. The assay was initiated by addition of SAM, maintained at 25 °C for 30 min, and stopped by the addition of ethyl acetate (150 μl). After phase separation by 1 min centrifugation at 14,000 g, the upper organic phase was counted using a liquid scintillation counter (Beckman Coulter, Fullerton, CA) as previously described (D'Auria et al. 2002). The radiation counts in the organic phase indicated the amount of synthesized MeSA. Substrate specificity assay was also performed for recombinant GmSAMT1, with a range of substrates including SA, benzoic acid,

anthranilic acid, 3-hydroxybenzoic acid, 4-hydroxybenzoic acid, nicotinic acid, p-coumaric acid, caffeic acid, cinnamic acid, vanallic acid, jasmonic acid (JA), gibberellic acid, indole-3-acetic acid, and 2,4-dichlorophenoxyacetic acid. The activity with SA was set at 100%. Three independent assays were performed for each compound.

Determination of Kinetic Parameters of GmSAMT1. Appropriate enzyme concentrations and incubation time were determined in time course assays to make sure that the reaction velocity was linear during the assay period. Lineweaver-Burk plots yielded apparent Km values as previously described (Chen et al. 2003a). To determine the Km for SA, concentrations of SA were independently varied from 2 to 150 μM, while SAM was held constant at 200 μM. Assays were conducted at 25 °C for 30 min, as described in SAMT activiy assay. The optimal pH for GmSAMT1 was assayed from pH 6.0 to pH 9.0, and optimal reaction of temperature was assayed ranging from 0 °C to 60 °C. Effects of metal ions on activity of GmSAMT1 were assayed with a range of metal ions, including K^+, Cu^{2+}, Zn^{2+}, Fe^{2+}, Fe^{3+}, Mn^{2+}, Ca^{2+}, Na^+, NH_4^+, and Mg^{2+}, in the form of chloride salts at 5 mM final concentrations.

Creating the Constructs with Over-expressed *GmSAMT1* and RFP Reporter Genes. The pCAMBIA 1305.2 vector was used as a primary binary vector. The following is the procedures of constructing the vector harboring over-expressed *GmSAMT1* (**Figure 2.5)**. As a screening tool, a red fluorescent protein (RFP) gene (Alieva et al. 2008), *ppor RFP*

with primers containing *Nco*I fragment, was introduced into *Nco*I site of pCAMBIA 1305.2 vector. This modified pCAMBIA 1305.2 vector was called pJL-RFP vector and used to test RFP expression through transient transforming tobacco. Then the cassette of 35S::GUS/NOS terminator was taken from a pBI121 vector by digestion with *EcoR*I and *Hind*III and inserted into the *EcoR*I and *Hind*III sites of pJL-RFP vector to form pJL-RFP-35S::GUS vector, which was used for testing the RFP and GUS gene expression in one binary vector. *GmSAMT1* gene was amplified from the above mentioned pEXP5/CT-TOPO vector with primers containing *BamH*I and *Sac*I fragment and replaced with GUS gene in pJL-RFP-35S::GUS vector by digestion and ligation. Finally the over-expressed *GmSAMT1* vector was constructed and called pJL-RFP-35S::GmSAMT1 vector. The vectors of pJL-RFP-35S::GmSAMT1 and pJL-RFP were transformed into *A. rhizogenes* as target gene and vector control, respectively. *A. tumefaciens* GV3101 and *A. rhizogenes* strain K599, were transformed by freeze-thaw method. The cells with no binary vector were plated on LB media with no antibiotic. The cells with the binary vectors were plated on selective LB media with kanamycin and incubated at 28°C for 2 days.

Testing the Reporter Gene Expression Through Tobacco Transient Transformation. Before generating hairy roots, the constructs containing RFP gene were tested by transient expression assays in tobacco plants according to Sparkes et al (2006). The plasmids, pJL-RFP and pJL-RFP-35S::GmSAMT1, were tested for RFP function by agroinfiltration of *A. tumefaciens* GV3101 into tobacco leaves, along with *A.*

tumefaciens GV3101 without binary plasmid as a negative control. Bacterial suspensions carrying target plasmids were pressure-infiltrated using a syringe with no needle into the 5-week tobacco leaves through the abaxial surfaces over leaf segments lying between the major veins that branch off from the mid-rib of the leaf. RFP fluorescence was investigated 2 days after infiltration in excised leaf sections that covered the infiltrated zone part of surrounding non-infiltrated tissue, using an Olympus SZX12 fluorescence microscope and an RFP filter.

Generation of Transgenic Soybean Hairy Roots. The soybean hairy root transformation system by infection with *A. rhizogenes* harboring the target genes were modified and used to study soybean resistance against SCN (Cho et al. 2000; Kereszt et al. 2007). *A. rhizogenes* K599 harboring the target gene, vector control and *A. rhizogenes* K599 alone were used to inoculate the TN02-275 soybeans, which are susceptible to SCN race 2. To better evaluate the resistance of the target gene, *A. rhizogenes* harboring the vector control and *A. rhizogenes* K599 alone were also used to inoculate the TN02-226 soybeans, which are resistant to SCN race 2. Generation of soybean hairy roots was modified as follows: First, soybean seeds were sterilized with chlorine gas. Second, a single layer of mature soybean seeds with intact seed coats were placed in open Petri dish, which was put in a bell jar desiccator within a fume hood. A beaker with 100 ml bleach was also placed in the desiccator and added 5 ml concentrated (12N) HCl along the side of the beaker. Then the seeds were kept in the closed desiccator immediately and maintained in the chlorine gas atmosphere

overnight (about 16 h). After sterilization, seeds were transferred to the autoclaved filter paper moistened with sterilized distilled water in the Petri dishes for germination in a laminar flow hood. The germinated soybean seeds were sowed in sterilized vermiculite and six seeds were grown in each cell of 18-cell germination tray. The overnight cultured bacteria lawn of *A. rhizogenes* K599 with binary vector pJL-RFP, pJL-RFP-35S::GUS or pJL-RFP-35S::GmSAMT1 from LB medium plate containing 50 mg ml^{-1} kanamycin, or a culture of *A. rhizogenes* K599 lacking the binary vector grown in LB medium without any antibiotics was collected and suspended in 1 ml of sterile distilled water, respectively. Soybean cotyledonary nodes and upper hypocotyls of 5-day-old seedlings with unfolded cotyledons were stabbed three times by the syringe containing the bacterial suspension (**Figure 2.7A**). In the first week, the trays were covered with the sterile transparent lids to retain high humidity (**Figure 2.7B, C**). In the subsequent three weeks, plants were transferred into pots and the *A. rhizogenes* wounding sites were covered by wet sterile vermiculite to continually keep high humidity (**Figure 2.7D,E**). Each day soybean plants were watered using sterile B&D solution containing 1 mM $CaCl_2$, 0.5 mM KH_2PO_4, 10 µM Fe-citrate, 0.25 mM $MgSO_4$, 0.25 mM K_2SO_4, 1 µM $MnSO_4$, 2 µM H_3BO_4, 0.5 µM $ZnSO_4$, 2 µM $CuSO_4$, 0.1 µM $CoSO_4$, 0.1 µM Na_2MoO_4, and 1 mM KNO_3. The growth conditions for soybean plants were 12 h light/12 h dark at 28 °C /25 °C with irradiance from fluorescent bulbs at 150-200 µmol $m^{-2} \cdot s^{-1}$. After about four weeks, the hairy roots grew to approximately 10 cm in length. Soybean transgenic roots were detected based on RFP expression, using a fluorescence stereomicroscope (Olympus SZX12 fluorescence microscope and

RFP filter). The tap root and non-transgenic roots were excised under the wounding site (**Figure 2.7F**). Transgenic roots harboring pJL-RFP-35S::GUS screened by RFP signal were immersed in GUS solution (1 mM X-Gluc), 0.5 mM potassium ferrocyanide, 0.1% (v/v) triton X-100, 100 mM sodium phosphate buffer pH 7.0), and incubated at 37 °C overnight. In the following day, the X-Gluc solution was discarded. The roots were washed repeatedly with 70% ethanol to remove chlorophyll. Transgenic roots with visual GUS stain were recorded.

Nematode Infection and Demographics Assay. To easily compare the SCN infection, one transgenic root was left for the SCN inoculation. Other transgenic roots were collected for RNA extraction and detection of the gene expression. The hairy root parts of twenty soybean plants harboring the same construct were loaded horizontally in a 13x 9 x 2 cm sterilized inoculating tray containing one thin layer mixture of sterile sand and top soil (1:1) and the shoot parts were left out of inoculating tray (**Figure 2.9A**). About 0.75 ml of inoculum, which contained about 5600 SCN eggs (**Figure 2.9B**), was added to each root system. Nematodes were allowed to infect roots for 7 days under humid conditions. Then all the roots were washed to remove extra SCN eggs and juvenile nematodes that had not penetrated the root tissue. Sterile vermiculite within cone-tainers was used for growing the infected chimeras in the growth chamber (**Figure 2.9C, D, E**). In two weeks post inoculation, infected root samples were washed to remove vermiculite by tap water and cleared by 20% (v/v) bleach for no more than 4 min, and then stained by acid fuchsin according the procedure of Byrd et al. (1983).

Nematodes of different stages within each infected root sample were recorded. Since the responses between the susceptible and resistant soybean lines occur on the J3 stage of SCN, the ratio of the number of J3+J4 to total number of nematodes per plant was used as an index to determine the resistance difference among the hairy roots harboring respective constructs. A minimum of 10 transformed plants were analyzed for each vector.

Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR). Quantitative reverse transcription PCR was performed to check *GmSAMT1* and some defense-related genes' transcription level of transgenic soybean hairy roots according to Mazarei et al. (2007). In order to conduct qRT-PCR assays, gene specific primers of *GmSAMT1*, defense-related genes, *GmPR-1* (GenBank accession BU577813), *GmPR-3* (GenBank accession AF202731), and reference gene, soybean ubiqutin3 gene (*GmUBI-3,* GenBank accession D28123), were designed (**Table 2.2**). Total RNA was isolated from the respective root tissues of transgenic hairy roots and control hairy roots using RNeasy columns (Qiagen) following the manufacturer's instructions for plants. RNA was treated with RNase-Free DNase set (Qiagen) to remove genomic DNA and total RNA (about 1.5 µg) was reverse-transcribed to synthesize cDNA using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) for qRT-PCR according to manufacturer's instructions. First-strand cDNA was diluted and placed in each qPCR reaction. DNA accumulation was measured using SYBR Green as the reference dye. Genomic DNA contamination was monitored by PCR using RNA

sample as template. Only one product was present in each reaction as indicated by the reference dye's dissociation curve of amplified products. The following PCR conditions were used: 50 °C for 2 min; 95 °C for 10 min; followed by 40 cycles of 95 °C for 15 s; 60 °C for 1 min; 72 °C for 30 s. All qRT-PCR assays were conducted in triplicate. PCR efficiencies for target and reference genes were equal between the target and control samples. Ct values and relative abundance were calculated using software supplied with the Applied Biosystems 7900 HT Fast Real-Time PCR system.

Statistical Analysis. The resistance difference among the transgenic soybean hairy roots with over-expression of *GmSAMT1* and control soybean hairy roots was analyzed by one way ANOVA using the mixed model of SAS (SAS 9.2 version). Least squares means of the resistance index were separated using Fisher's least significant difference (LSD) test. The expression difference of target gene and defense-related genes examined by qRT-PCR was also analyzed by one way ANOVA using the mixed model of SAS. Least squares means of relative expression level were separated using Fisher's LSD test. A $P<0.05$ was considered significant.

RESULTS AND DISCUSSION

Glyma02g06070 Is a Member of the SABATH Family. A recent microarray analysis on soybean recombinant inbred lines (RILs), TN02-226 and TN02-275, revealed some genes with significant expression changes during resistant and susceptible reactions to SCN. One of these genes showed significant up-regulation only in SCN inoculated

resistant soybean roots while no significant change of its gene expression occurred in SCN-inoculated susceptible soybean roots (Mazarei et al. 2011). To obtain the complete open reading frame of this resistance candiadate gene, the sequence of Affymetrix SoyChip probe Gma.12911.1.A1_S_AT, which identified our target gene from microarrays, was used as a query to BLAST search against the soybean database GmGDB/GMtranscript (http://www.plantgdb.org/GmGDB/cgi-bin/blastGDB.pl) and SoyBase (http://soybase.org/gbrowse/cgi-bin/gbrowse/gmax1.01/), the soybean gene, *Glyma02g06070* was found to have the highest similarity (99%) with the sequence of Gma.12911.1.A1_S_AT and was annotated for encoding SAM-dependent carboxyl methyltransferase (**Figure 2.1**).

By further comparison with the known methyltransferases from other species, this SAM-dependent carboxyl methyltransferase encoded by *Glyma02g06070* was determined to be a member of the SABATH family. The SABATH family is named after the first identified enzymes (SAMT, BAMT and theobromine synthase) in this family (Chen et al. 2003a). To investigate the evolutionary relationships between Glyma02g06070 and the known SABATH members from other species, a phylogenetic tree of selected functionally characterized SABATH members with Glyma02g06070 was constructed. The result showed that Glyma02g06070 belonged to the cluster containing the enzymes with salicylic acid methyltransferase activity (**Figure 2.2)**.

Glyma02g06070 Encodes Soybean SAMT. A protein BLAST search reveals that Glyma02g06070 contains 370 amino acids and is 40.1% identical to the previously

identified rice SAMT, OsBSMT1, at the amino acid sequence level. Glyma02g06070 was also predicted to contain an SAM-binding domain. When the sequence of Glyma02g06070 was searched in EST database, it also showed this gene is induced by SCN and SA. To confirm the prediction from the databases, we amplified and sequenced the complete open reading frame (ORF) of this gene from cDNA attained from SCN-infected root tissue of soybean TN02-226 line.

Since *Glyma02g06070* encodes a member of the SABATH family, whose member enzymes can transfer a methyl group from methyl-donor SAM to a range of substrates, the enzyme activity of the recombinant protein of *Glyma02g06070* was examined using various potential substrates (**Table 2.1**). Increases in reaction rate by increasing concentrations of SAM and SA were found to obey Michaelis-Menten kinetics. Among the 14 examined substrates, SA, benzoic acid, anthranilic acid and 3-hydroxybenzoic acid served as *in vitro* substrates in the reactions catalyzed by the recombinant protein encoded by *Glyma02g06070*, although they showed obvious differences in activities. At substrate concentrations of 1 mM, the recombinant protein showed the highest activity with SA (100%) rather than benzoic acid (16.9%), anthranilic acid (9.3%) and 3-hydroxybenzoic acid (1.8%). For other substrates including JA, indole-3-acetic acid, and gibberellic acid, no activity was detected under the same reaction conditions. Glyma02g06070 catalyzed the methylation of SA with an apparent Km value of 32.9±1.7 μM. Final values represent the average of three independent measurements. Due to SA serving as its sufficient substrate, Glyma02g06070 was designated as *GmSAMT1*.

Prior to this study, SAMT has been cloned and characterized from a number of plant species including *Clarkia breweri, Arabidopsis thaliana*, rice (*Oryza sativa*), tobacco (Ross et al. 1999; Chen et al. 2003a; Koo et al. 2007; Park et al. 2007). These studies contributed to our understanding SA-dependent defense mechanism. In soybean, however, the study of SA-dependent defense has been very limited. The isolation and characterization of *GmSAMT1* makes it possible to investigate the role of SAMT in soybean defense.

Biochemical Properties of Soybean SAMT. Recombinant GmSAMT1 with a His-tag expressed in *E. coli* was purified to electrophoretic homogeneity and subjected to detailed biochemical characterization. The molecular mass of the GmSAMT1 with His-tag protein estimated on the gel was 43 kDa, which was close to the expected size (**Figure 2.3**). Purified recombinant GmSAMT1 exhibited an apparent Km value of 32.9 ±1.7 μM for salicylic acid. When enzyme assays were performed in buffers of different pH values, GmSAMT1 showed the highest level of catalytic activity at pH 7.5. (**Figure 2.4A**). The optimal temperature for GmSAMT1 activity was 25 ºC (**Figure 2.4B**). The effects of various metal ions on GmSAMT1 activity were measured. Compared with control, GmSAMT1 exhibited double activity in presence of K^+, while it showed no activity in presence of Cu^{2+}, Zn^{2+}, Fe^{2+} and Fe^{3+}. GmSAMT1 activity was moderately inhibited by Mn^{2+} and Ca^{2+}. Other metal ions, including Na^+, NH_4^+, and Mg^{2+}, had minimal effect on GmSAMT1 activity (**Figure 2.4C**).

All the above results suggested the recombinant GmSAMT1 showed salicylic acid methyltransferase activity. Taken together, the phylogenetic and biochemical evidence suggested that GmSAMT1 is a conserved enzyme among a number of plant species.

Reporter Gene Expression through Tobacco Transient Transformation. Before performing *A. rhizogenes*-mediated transformation, transient expression of pJL-RFP-35S::GUS in tobacco leaf through *A. tumefaciens*-mediated transformation was conducted to check RFP signal under RFP filter. Transgenic tobacco leaves harboring the target gene displayed RFP fluorescence, while the mock plant showed no RFP signal (**Figure 2.6A**). This result helped to rapidly confirm the RFP expression.

After *A. rhizogenes*-mediated transformed soybean hairy roots with target genes were generated to be about 10 cm long (**Figure 2.7**), transgenic hairy roots harboring over-expressed *GmSAMT1* were screened under RFP filter (**Figure 2.6B**). This method was demonstrated by checking GUS expression by staining the transgenic hairy roots harboring plasmid pJL-RFP-35S::GUS. All the positive transgenic hairy roots screened by RFP signal showed uniform GUS expression in the root tissue, while negative control hairy roots screened by RFP signal did not show GUS expression.

Gene Expression of *GmSAMT1* and Several Pathogenesis-Related (PR) Genes. The gene expression level of *GmSAMT1* and selected PR genes in transgenic hairy roots were examined by qRT-PCR (**Figure 2.8**). The expression level of *GmSAMT1* in over-expressed *GmSAMT1* transgenic hairy roots was about 4 times higher than the

control plants (**Figure 2.8A**). *GmPR-1* showed significant higher expression in over-expressed *GmSAMT1* transgenic hairy roots compared with some control hairy root lines. But there was no significant difference of the gene expression of *GmPR-1* between the transgenic hairy roots with over-expression of *GmSAMT1* and vector control in the susceptible soybean background (**Figure 2.8B**). This may suggest that the susceptible soybean can respond to exogenous proteins when SCN infection does not occur. When SCN parasitism happens, the defense system of the susceptible soybean may be changed or inhibited by SCNs, which produce and secrete parasitism proteins into soybean to aid themselves in the suppression of host plant defense. Chorismate mutase was found to be one of the SCN parasitism proteins, which can manipulate the plant's shikimate pathway (Bekal et al. 2003). Since the SA biosynthesis is closely associated with the shikimate pathway, the SA signaling pathway of the susceptible soybean would be suppressed by SCN. It was already found that a local suppression of SA signaling in Arabidopsis roots was required for beet cyst nematode parasitism (Wubben et al. 2008). In future, the expression of PR genes needs to be examined for the SCN-infected transgenic hairy roots. Since *PR-1* gene was considered as a marker gene of SA signaling pathway, over-expressed *GmSAMT1* might increase the signal transduction of SA pathway by providing abundant mobile molecule MeSA for responding SCN infection. Interestingly, the expression of *PR-3*, which encodes soybean chitinase, was observed to be higher in over-expressed *GmSAMT1* transgenic hairy roots than other control lines (**Figure 2.8C**). JA was found to induce *PR-3* expression in Arabidopsis (Thomma et al. 1998, Seo et al. 2008).

Biological Roles of Soybean SAMT in SCN Resistance. Since the different resistant responses to SCN race 2 between the susceptible (TN02-275) and resistant (TN02-226) soybean lines occur on the J3 stage of SCN, the demographics of SCN in all the transgenic hairy roots with respective constructs were analyzed after two-week inoculation by SCNs. The nematodes from J2 stage to J4 stage were observed (**Figure 2.10**). The index of J3+J4/total SCN number was used to evaluate the resistance among the different hairy roots in the assay. The susceptibility level of negative vector control was arbitrarily set at 100%. The results showed the transgenic soybean hairy roots with over-expressed *GmSAMT1* had significantly fewer female juvenile SCN developing to J3 stage, compared with vector control in the susceptible soybean background (85.4% reduction). And there is no significant difference on the resistance level between the transgenic hairy roots with over-expressed *GmSAMT1* and vector control in the resistant soybean background, which also had significantly fewer female SCN juvenile developed to J3 stage, compared with vector control in the susceptible soybean background (94.8% reduction) (**Figure 2.11**).

SCN is the most economically damaging pest of soybean in the U.S., so the mechanism of the interaction between soybean and SCN has been studied to better control this pest. With the development of genomics, a number of microarray analyses have been performed to study gene expression changes of soybean when infected by SCN. However, the various experiments are not congruent owing to different types of soybean and SCN samples (Khan et al. 2004; Klink et al. 2007, 2009, 2010, 2011; Lu et al. 2006; Puthoff et al. 2007). In the present study, the candidate gene was identified

based on microarray analysis of two genetically-related soybean sister lines TN02-226 and TN02-275, which are resistant and susceptible, respectively, to the SCN race 2 infection (Mazarei et al. 2011). To our knowledge, it is the first time that recombinant inbred lines have been used to study gene profile changes in response to a novel SCN population, race 2 (HG type 1.2.5.7). *GmSAMT1* was one of the identified candidate genes, which showed up-regulated in the resistant line TN02-226 but not in the susceptible line TN02-275.

GmSAMT1 may play a critical role in soybean defense against SCN. The induced expression of *GmSAMT1* may be required for soybean defense against SCN, considering up-regulation of *GmSAMT1* was only observed in the resistant soybean line. In susceptible soybean lines, SCN may avoid or suppress the recognition of soybean defense system (Bekal et al. 2003) and block the SA signaling pathway through not trigging the expression of *GmSAMT1*. Low level of expression of *GmSAMT1* may lead to the deficient production of mobile MeSA that can be transported from the infected sites. In this study, the over-expression of *GmSAMT1* in the susceptible soybean line showed similar resistance level with the resistant line, which may be caused by complementing the low expression of *GmSAMT1* in the susceptible soybean.

During SCNs' inoculation, GmSAMT1 activity may also be manipulated through changing the concentration of potassium of soybean cells. The most common aboveground symptom of SCN infection is a potassium deficiency symptom with a bright yellowing of the margins of leaves. However, such SCN-induced potassium deficiency symptoms are the result of the plant's inability to raise sufficient potassium

due to nematode feeding. It was demonstrated that potassium concentration in soybean roots grown at the medium level of potassium fertility condition was decreased by SCN (Smith et al. 2001). The result indicated that *in vitro* GmSAMT1 showed double activity in the presence of potassium (**Figure 2.4**), which suggests that SCN may also inhibit the activity of GmSAMT1 by decreasing the concentration of potassium in addition to suppressing the transcriptional level of *GmSAMT1*.

GmSAMT1 may have a role in crosstalk between SA and JA signaling pathway in soybean resistance against SCN. It was demonstrated that the gene expression of rice *SAMT* can be induced by JA (Zhao et al. 2010). MeSA was also detected to be highly elevated after the wild-type Arabidopsis was treated by MeJA (Koo et al. 2007). In our study, the over-expression of *GmSAMT1* in the susceptible soybean line caused up-regulation of *GmPR-1* and *GmPR-3*, which may be the integrative effect of SA and JA signaling pathways. Microarray analyses also suggested that the JA pathway was activated in the soybean defense response against SCN (Klink et al. 2007; Mazarei et al. 2011). Therefore, the SA-dependent pathway, combined with JA pathway through the regulatory points, including NPR1, WRKY family transcription factor, and LOX (Klink et al. 2007), may play a role of restraining the development of SCN. Our results suggested that GmSAMT1 may be another regulatory point in the crosstalk between SA and JA pathways. In Chapter 3, the role of GmSAMT1 in SA-dependent soybean resistance against SCN will be further discussed in a model (**Figure 3.6**).

Biological Roles of *GmSAMT1* in Soybean Defense Against Soybean Mosaic Virus. To test whether *GmSAMT1* is also involved in resistance against SMV, qRT-PCR was also conducted on measuring the expression of *GmSAMT1* and *GmPR-1* during interaction between SMV and soybean. In this assay two soybean lines L78-379 (*Rsv1*) and Williams 82 (*rsv1*) were utilized. Soybean genotype L78-379 is an isoline of Williams, but contains the Rsv1 resistance gene from PI 96983 (*Rsv1*) (Hajimorad and Hill, 2001). Inoculation of L78-379 (*Rsv1*) with a mutant derived from SMV-N (N25) results in induction of delayed hypersensitive response, whereas infection of Williams 82 (*rsv1*) with the same virus causes mosaic.

When the resistant response occurred for L78 against soybean virus 25 (21 days post inoculation), *GmSAMT1* expression of the soybean top leaves is around 7.3 times that in the mock plant, and interestingly, a high expression level (44.6 times) of *GmPR-1* were also observed. While in the susceptible response of soybean Williams 82, the expression of *GmSAMT1* and *GmPR-1* was reduced to half of that in the mock plant (**Figure 2.12**). So the strongly positive correlation between the expression of *GmSAMT1* and *GmPR-1* suggested that *GmSAMT1* may also play a role in soybean defense against soybean mosaic virus.

Taken together, the gene expression differences of *GmSAMT1* and selected *PR* genes, were up-regulated in the two inoculated tissues of the resistant soybean line, roots and leaves but down-regulated or not changed in the susceptible soybean line, may explain, at least partially, the soybean responses against pathogens. This conclusion is drawn largely based on correlative data from this study. Additional

experiments, such as transgenic studies using stable transformants, are still needed to elucidate the specific roles of GmSAMT1 in soybean defense against pathogens.

REFERENCES

Alieva NO, Konzen KA, Field SF, Meleshkevitch EA, Hunt ME, Beltran-Ramirez V, Miller DJ, Wiedenmann J, Salih A, Matz MV (2008) Diversity and evolution of coral fluorescent proteins. Plos One 3: e2680

Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman D (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402

Arelli PR, Sleper DA, Yue P, Wilcox JA (2000) Soybean reaction to races 1 and 2 of *Heterodera glycines*. Crop Sci. 40: 824-826

Bekal S, Niblack TL, Lambert KN (2003) A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. Mol. Plant Microbe In. 16: 439-446

Bernard R, Nelson R, Creemens C (1991) USDA soybean genetic collection: isoline collection. Soybean Genet. Newsl. 18: 27-57

Bhattarai KK, Xie Q-G, Mantelin S, Bishnoi U, Girke T, Navarre DA, Kaloshian I (2008) Tomato susceptibility to root-knot nematodes requires an intact jasmonic acid signaling pathway. Mol. Plant Microbe In. 21: 1205-1214

Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 7: 248-254

Byrd D, Kirkpatrick T, Barker K (1983) An improved technique for clearing and staining plant tissue for detection of nematodes. J. Nematol. 15:142-143

Chen F, D'Auria JC, Tholl D, Ross JR, Gershenzon J, Noel JP, Pichersky E (2003a) An *Arabidopsis thaliana* gene for methylsalicylate biosynthesis, identified by a biochemical genomics approach, has a role in defense. Plant J. 36: 577-588

Chen F, Tholl D, D'Auria JC, Farooq A, Pichersky E, Gershenzon J (2003b) Biosynthesis and emission of terpenoid volatiles from Arabidopsis flowers. Plant Cell 15: 481-494

Cho HJ, Farrand SK, Noel GR, Widholm JM (2000) High-efficiency induction of soybean hairy roots and propagation of the soybean cyst nematode. Planta 210: 195-204

Dann E, Diers B, Byrum J, Hammerschmidt R (1998) Effect of treating soybean with

2,6-dichloroisonicotinic acid (INA) and benzothiadiazole (BTH) on seed yields and the level of disease caused by *Sclerotinia sclerotiorum* in field and greenhouse studies. Eur J. Plant Pathol. 104: 271-278

Eggenberger AL, Hajimorad MR, Hill JH (2008) Gain of virulence on rsv1-genotype soybean by an avirulent soybean mosaic virus requires concurrent mutations in both P3 and HC-Pro. Mol. Plant Microbe In. 21: 931-936

Gillen AM, Shelton GW (2007) Uniform soybean tests, southern states, 2006. USDA-ARS, Stoneville

Gillen AM, Shelton GW (2008) Uniform soybean tests, southern states, 2007. USDA-ARS, Stoneville

Hajimorad MR, Hill JH (2001) Rsv1-mediated resistance against soybean mosaic virus-N is hypersensitive response-independent at inoculation site, but has the potential to initiate a hypersensitive response-like mechanism. Mol. Plant Microbe In. 14: 587-598

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

Khan R, Alkharouf N, Beard H, MacDonald M, Chouikha I, Meyer S, Grefenstette J, Knap H, Matthews B (2004) Microarray analysis of gene expression in soybean roots susceptible to the soybean cyst nematode two days post invasion. J. Nematol. 36: 241-248

Klink VP, Hosseini P, Matsye P, Alkharouf NW, Matthews BF (2009) A gene expression analysis of syncytia laser microdissected from the roots of the *Glycine max* (soybean) genotype PI 548402 (Peking) undergoing a resistant reaction after infection by *Heterodera glycines* (soybean cyst nematode). Plant Mol. Biol. 71: 525-567

Klink VP, Hosseini P, Matsye PD, Alkharouf NW, Matthews BF (2011) Differences in gene expression amplitude overlie a conserved transcriptomic program occurring between the rapid and potent localized resistant reaction at the syncytium of the *Glycine max* genotype Peking (PI 548402) as compared to the prolonged and potent resistant reaction of PI 88788. Plant Mol. Biol. 75: 141-165

Klink VP, Overall CC, Alkharouf NW, MacDonald MH, Matthews BF (2007) A time-course comparative microarray analysis of an incompatible and compatible response by *Glycine max* (soybean) to *Heterodera glycines* (soybean cyst nematode) infection. Planta 226: 1423-1447

Klink VP, Overall CC, Alkharouf NW, MacDonald MH, Matthews BF (2010) Microarray detection call methodology as a means to identify and compare transcripts expressed within syncytial cells from soybean (*Glycine max*) roots undergoing resistant and susceptible reactions to the soybean cyst nematode (*Heterodera glycines*). J. Biomed. Biotechnol. 2010:491217

Koo Y, Kim M, Kim E, Song J, Jung C, Moon J-K, Kim J-H, Seo H, Song S, Kim J-K, Lee J, Cheong J-J, Choi Y (2007) Overexpression of salicylic acid carboxyl methyltransferase reduces salicylic acid-mediated pathogen resistance in *Arabidopsis thaliana*. Plant Mol. Biol. 64: 1-15

Kulkarni SS, Bajwa SG, Rupe JC, Kirkpatrick TL (2008) Spatial correlation of crop response to soybean cyst nematode (*Heterodera Glycines*). T. Asabe 51: 1451-1459

Lu P, Shannon JG, Sleper DA, Nguyen HT, Cianzio SR, Arelli R (2006) Genetics of cyst nematode resistance in soybean PIs 467312 and 507354. Euphytica 149: 259-265

Ludwig A, Tenhaken R (2000) Defence gene expression in soybean is linked to the status of the cell death program. Plant Mol. Biol. 44: 209-218

Mazarei M, Elling AA, Maier TR, Puthoff DP, Baum TJ (2007) GmEREBP1 is a transcription factor activating defense genes in soybean and Arabidopsis. Mol. Plant Microbe In. 20: 107-119

Mazarei M, Liu W, Al-Ahmad H, Arelli P, Pantalone V, Stewart C (2011) Gene expression profiling of resistant and susceptible soybean lines infected with soybean cyst nematode. Theor. Appl. Genet. 123: 1193-1206

Michereff MFF, Laumann RA, Borges M, Michereff M, Diniz IR, Neto ALF, Moraes MCB (2011) Volatiles mediating a plant-herbivore-natural enemy interaction in resistant and susceptible soybean cultivars. J. Chem. Ecol. 37: 273-285

Paris RL, Shelton GW (2005) Uniform soybean tests, southern states, 2004. USDA-ARS, Stoneville

Paris RL, Shelton GW (2006) Uniform soybean tests, southern states, 2005. USDA-ARS, Stoneville

Park SW, Kaimoyo E, Kumar D, Mosher S, Klessig DF (2007) Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. Science 318: 113-116

Pieterse CMJ, Van Loon LC (2004) NPR1: the spider in the web of induced resistance signaling pathways. Curr. Opin. Plant Biol. 7: 456-464

Puthoff DP, Ehrenfried ML, Vinyard BT, Tucker ML (2007) GeneChip profiling of transcriptional responses to soybean cyst nematode, *Heterodera glycines*, colonization of soybean roots. J. Exp. Bot. 58: 3407-3418

Ross JR, Nam KH, D'Auria JC, Pichersky E (1999) S-Adenosyl-L-methionine: salicylic acid carboxyl methyltransferase, an enzyme involved in floral scent production and plant defense, represents a new class of plant methyltransferases. Arch. Biochem. Biophys. 367: 9-16

Sandhu D, Tasma IM, Frasch R, Bhattacharyya MK (2009) Systemic acquired resistance in soybean is regulated by two proteins, orthologous to Arabidopsis NPR1. Bmc Plant Biol. 9: 105

Seo PJ, Lee AK, Xiang F, Park CM (2008) Molecular and functional profiling of Arabidopsis pathogenesis-related genes: insights into their roles in salt response of seed germination. Plant Cell Physiol. 49: 334-344

Smith GJ, Wiebold WJ, Niblack TL, Scharf PC, Blevins DG (2001) Macronutrient concentrations of soybean infected with soybean cyst nematode. Plant Soil 235: 21-26

Sparkes IA, Runions J, Kearns A, Hawes C (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nat. Protocols 1: 2019-2025

Thomma BP, Eggermont K, Penninckx IA, Mauch-Mani B, Vogelsang R, Cammue BP, Broekaert WF (1998) Separate jasmonate-dependent and salicylate-dependent defense-response pathways in Arabidopsis are essential for resistance to distinct microbial pathogens. Proc. Natl. Acad. Sci. USA 95: 15107-15111

Winter SMJ, Rajcan I, Shelp BJ (2006) Soybean cyst nematode: challenges and opportunities. Can. J. Plant Sci. 86: 25-32

Wrather J, Elrod J (1990) Apparent systemic effect of *Colletotrichum truncatum* and *C. lagenarium* on the interaction between soybean and *C. truncatum*. Phytopathology 80: 472-447

Wubben MJE, Jin J, Baum TJ (2008) Cyst nematode parasitism of *Arabidopsis thaliana* is inhibited by salicylic acid (SA) and elicits uncoupled SA-independent pathogenesis-related gene expression in roots. Mol. Plant Microbe In. 21: 424-432

Zhao N, Guan J, Ferrer J-L, Engle N, Chern M, Ronald P, Tschaplinski TJ, Chen F (2010) Biosynthesis and emission of insect-induced methyl salicylate and methyl benzoate from rice. Plant Physiol. Bioch. 48: 279-287

Zhu JW, Park KC (2005) Methyl salicylate, a soybean aphid-induced plant volatile attractive to the predator *Coccinella septempunctata*. J. Chem Ecol. 31: 1733-1746

Appendix

Table 2.1 Relative activity of GmSAMT1 with salicylic acid and other related substrates

Substrate	**Relative activity [a] (%)**
Salicylic acid	100
Benzoic acid	16.9
Anthranilic acid	9.3
3-Hydroxybenzoic acid	1.8
4-Hydroxybenzoic acid	0
Nicotinic acid	0
p-Coumaric acid	0
Caffeic acid	0
Cinnamic acid	0
Vanallic acid	0
Indole-3-acetic acid	0
2,4-Dichlorophenoxyacetic acid	0
Jasmonic acid	0
Gibberellic acid	0

a: Values are averages of three independent measurements. All substrates were tested at 1 mM concentration. The activity of the recombinant GmSAMT1 with salicylic acid was set arbitrarily at 100%.

Table 2.2. Primers for qRT-PCR analysis of soybean SA methyltransferase gene and relevant pathogenesis-related genes.

Gene Name	**Orientation**	**Sequence (5'-3')**
GmSAMT1	Forward	AAAAGCATAAGCAAGAAGTTTGCA
	Reverse	CACAAACAAGCAAACAGTCACAAC
GmUBI-3	Forward	GTGTAATGTTGGATGTGTTCCC
	Reverse	ACACAATTGAGTTCAACACAAACCG
GmPR-1	Forward	AACTATGCTCCCCCTGGCAACTATATTG
	Reverse	TCTGAAGTGGTAGCTTCTACATCGAAACAA
GmPR-3	Forward	AACTACAATTACGGGCAAGCTGGCAA
	Reverse	TTGATGGCTTGTTTCCCTGTGCAGT

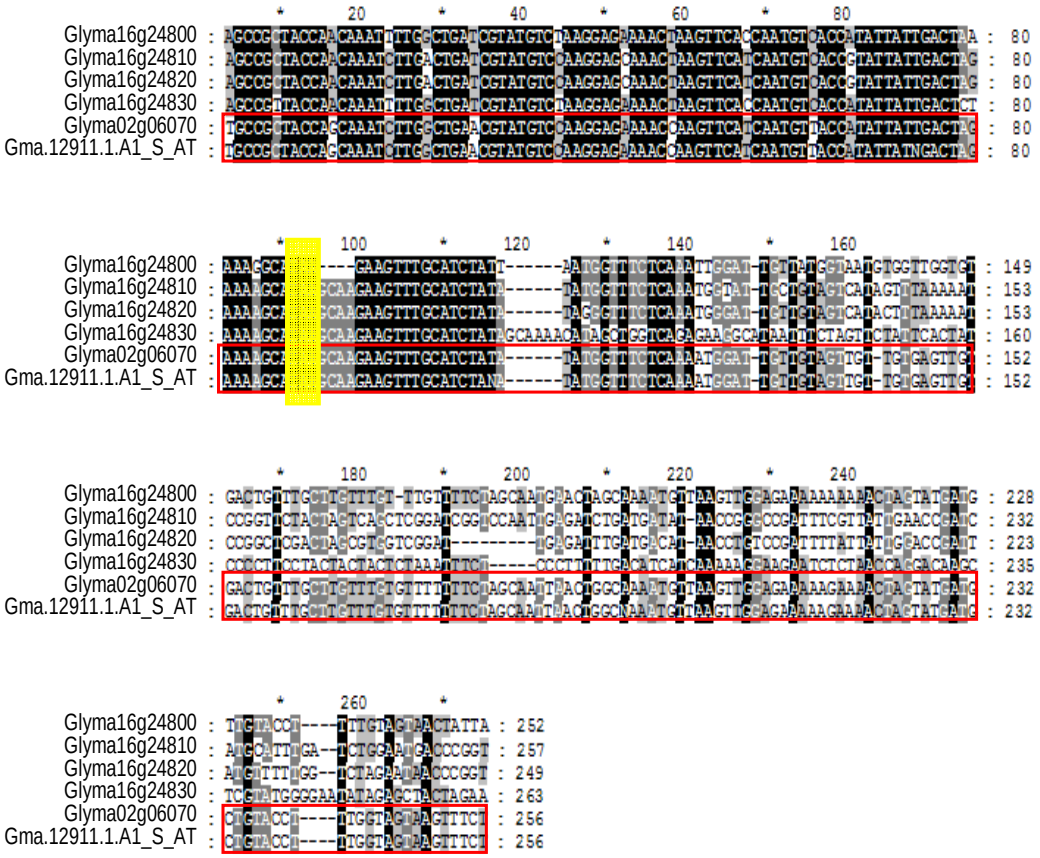

Figure 2.1. The alignment of Affymetrix probe Gma.12911.1.A1_S_AT sequence with the five soybean candidate genes. The sequence of Affymetrix probe Gma.12911.1.A1_S_AT exhibits at the end of open reading frame sequence of the soybean genes and covers the part of 3' UTR sequence (TAA in the yellow box is the stop codon). Based on the alignment, only *Glyma02g06070* has the identical sequence with the probe (indicated in the red box).

Figure 2.2. The phylogenetic tree of Glyma02g06070 and the functionally characterized SABATH family proteins from other plants. A neighbor-joining tree based on the degree of sequence similarity between Glyma02g06070 and selected functionally characterized SABATH from other plant species using the ClustalX program. Glyma02g06070 is in the SAMT clade shown in the oval. CbSAMT, *Clarkia breweri* SAMT (AF133053); AmSAMT, *Antirrhinum majus* (snapdragon) SAMT (AF515284); SfSAMT, *Stephanotis floribunda* SAMT (AJ308570); HcSAMT, *Hoya carnosa* SAMT (AJ863118); DwSAMT, *Datura wrightii* SAMT (EF472972); AmBAMT, *Antirrhinum majus* BAMT (AF198492); NsBSMT, *Nicotiana suaveolens* BSMT (AJ628349); AtBSMT, *Arabidopsis thaliana* BSMT (BT022049); AlBSMT, *A. lyrata* BSMT (AY224596); AbSAMT, *Atropa belladonna* SAMT (AB049752); OsBSMT1, *Oryza sativa* BSMT1 (XM467504); PhBSMT, *Petunia hybrida* BSMT (AY233465); OsIAMT1, *O. sativa* IAMT1 (EU375746); PtIAMT1, *Populus trichocarpa* IAMT1 (XP_002298843); AtIAMT, *A. thaliana* IAMT (AK175586); AtGAMT1, *A. thaliana* GAMT1 (At4g26420); AtGAMT2, *A. thaliana* GAMT2 (At5g56300); CaCaS1, *Coffea arabica* caffeine synthase 1 (AB086414); CaXMT1, *C. arabica* XMT1 (AB048793); CaDXMT1, *C. arabica* DXMT1 (AB084125). Branches were drawn to scale with the bar indicating 0.1 substitutions per site.

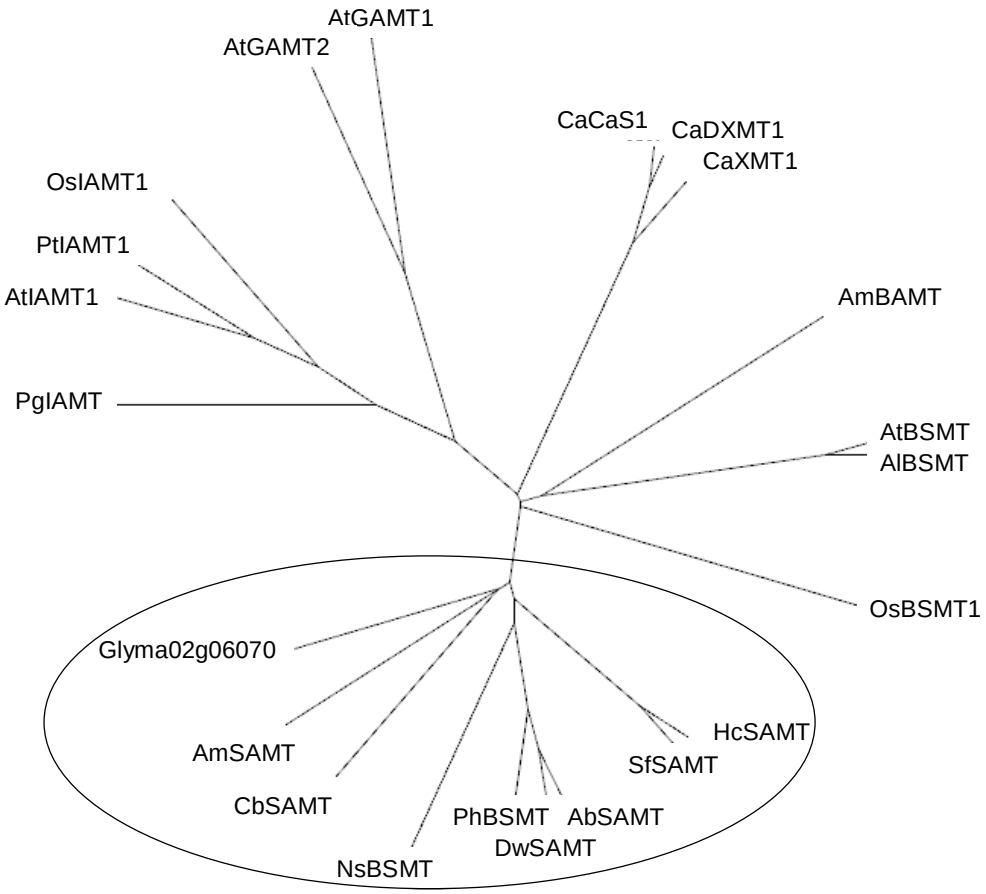

Figure 2.2. The phylogenetic tree of Glyma02g06070 and the functionally characterized SABATH family proteins from other plants.

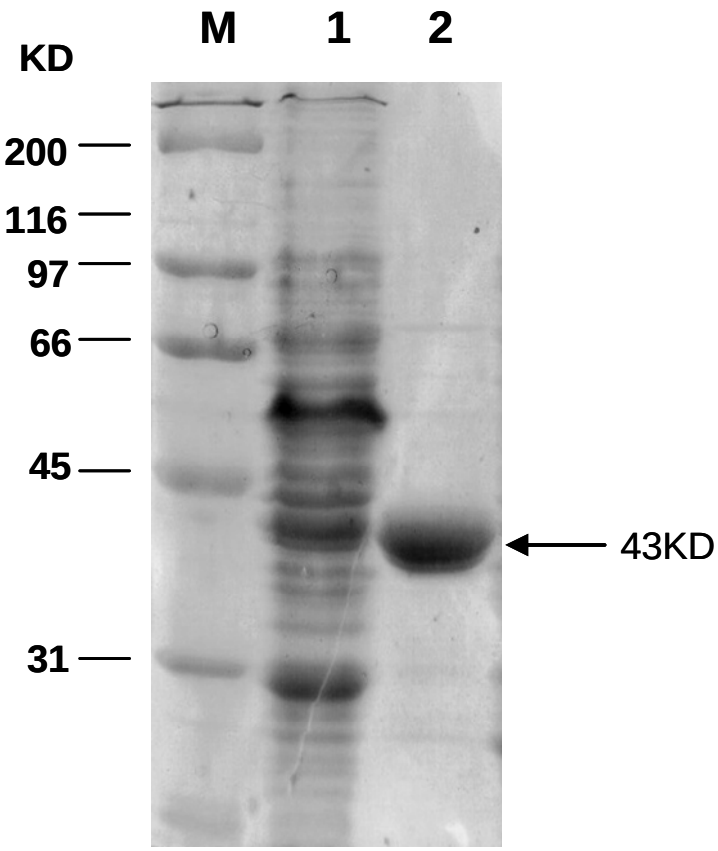

Figure 2.3. SDS-PAGE of purified recombinant GmSAMT1 protein. His-tagged GmSAMT1 expressed in *E. coli* was purified as described in the materials and methods. Lane M contained protein molecular weight markers. Lane 1 contained crude extract and lane 2 contained about 2 μg of purified GmSAMT1 protein. The gel was stained with Coomassie blue.

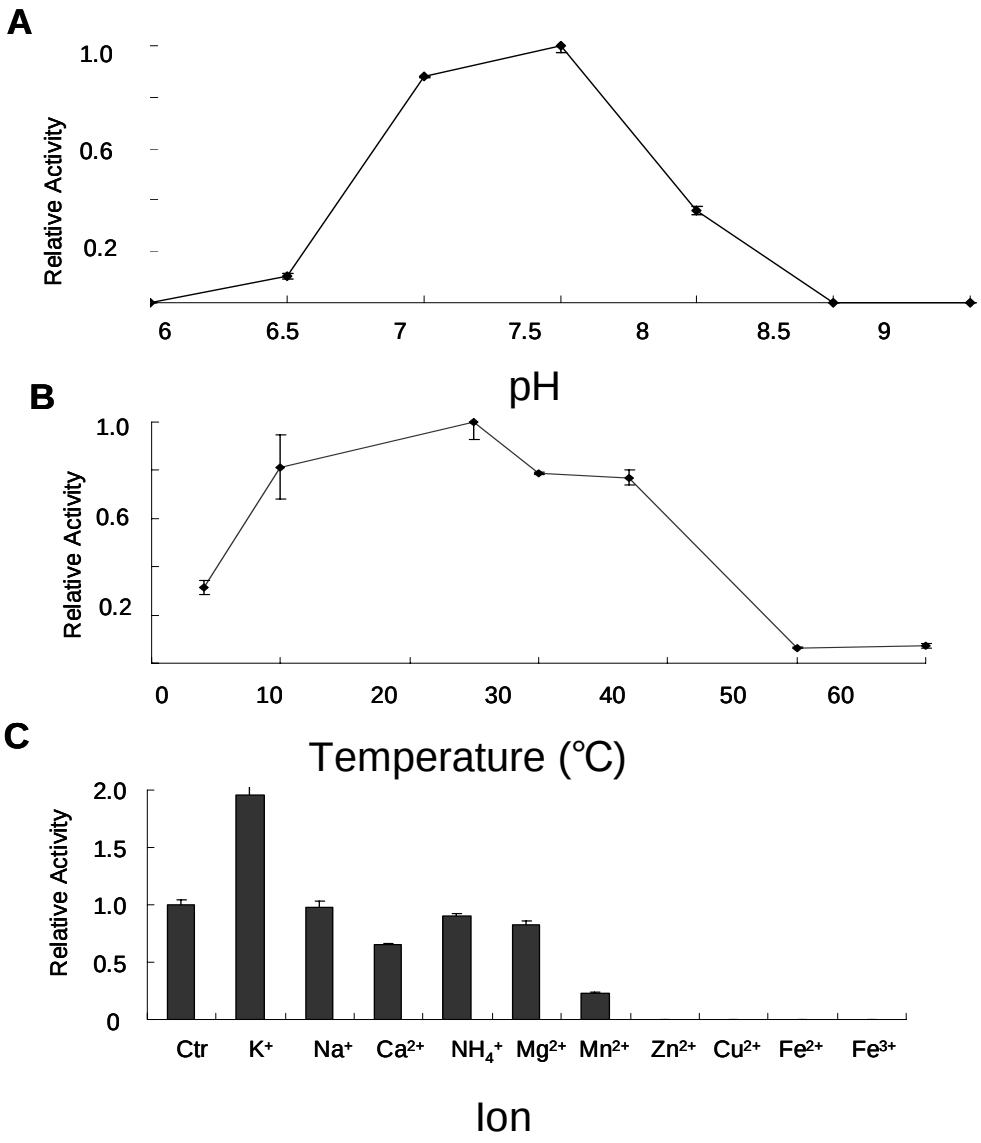

Figure 2.4. Biochemical properties of GmSAMT1. **A**, pH effect on GmSAMT1 activity. Level of GmSAMT1 activity in 50 mM Tris–HCl buffer, pH 7.5, was arbitrarily set at 1.0. **B**, GmSAMT1 activity in response to temperature ranging from 0 °C to 60 °C. The level of GmMAN1 activity at 25 °C was arbitrarily set at 1.0. **C**, Effects of metal ions on activity of GmSAMT1. Metal ions were added to reactions in the form of chloride salts at 5 mM final concentrations. The level of GmSAMT1 activity without any metal ion added as control (Ctr) was arbitrarily set at 1.0.

Figure 2.5. Construction of pJL-RFP-35S::GmSAMT1, a plasmid for over-expression of *GmSAMT1* gene in soybean roots. pJL-RFP is modificated from pCAMBIA 1305.2 by inserting a 35S promoter-*ppor RFP*-NOS terminator at *Noc*I site and *ppor RFP* encodes a red fluorescent protein. pJL-RFP is an Agrobacterium compatible binary plasmid for plant transformation and used for vector control plasmid. pJL-RFP-35S::GUS is an intermediate vector and used for testing whether *RFP* and *GUS* within two cassettes co-express in soybean roots. pJL-RFP-35S::GmSAMT1 is obtained by replacing *GUS* gene with *GmSAMT1* by digestion and ligation between *BamH*I and *Sac*I sites.

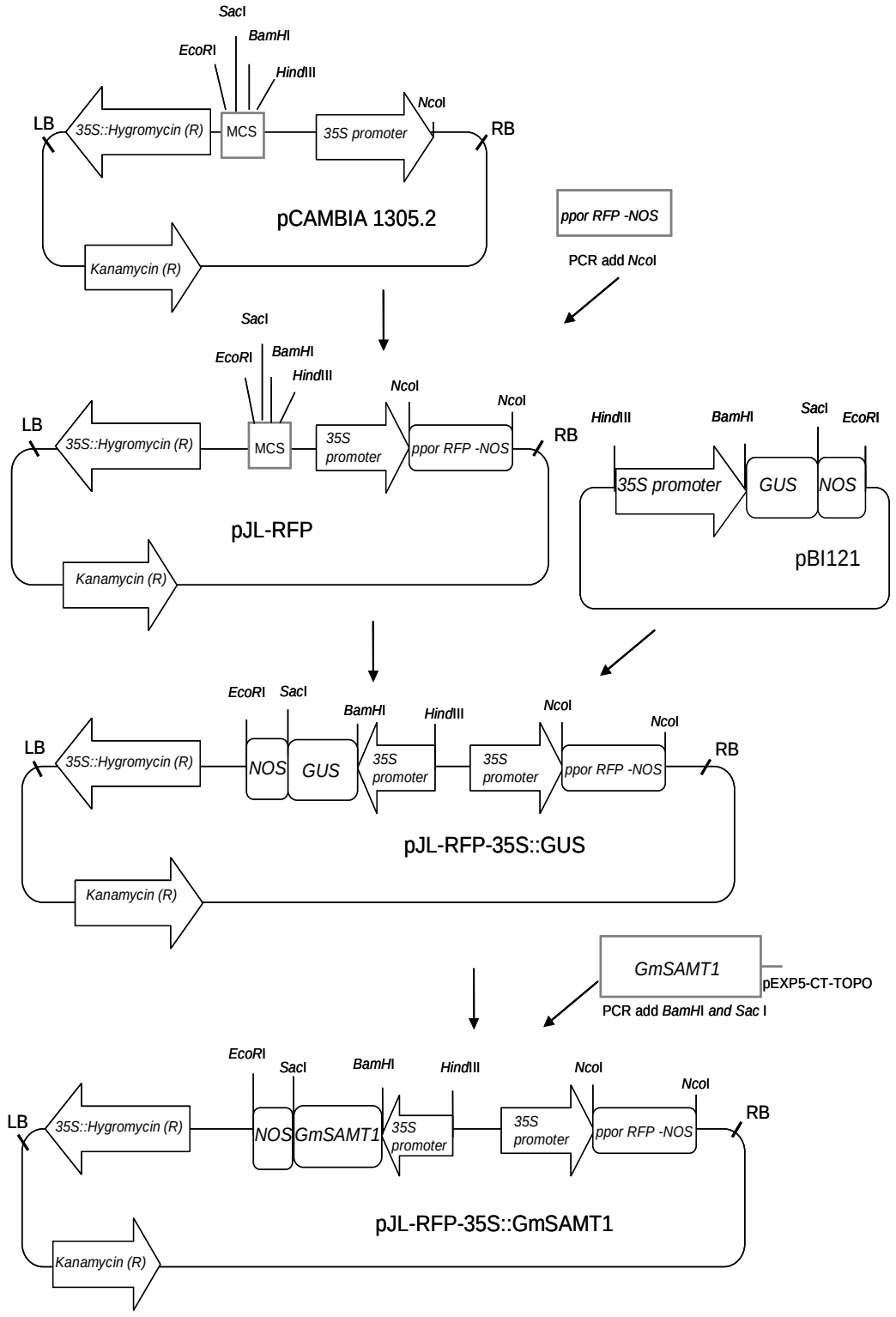

Figure 2.5. Construction of pJL-RFP-35S::GmSAMT1, a plasmid for over-expression of *GmSAMT1* gene in soybean roots.

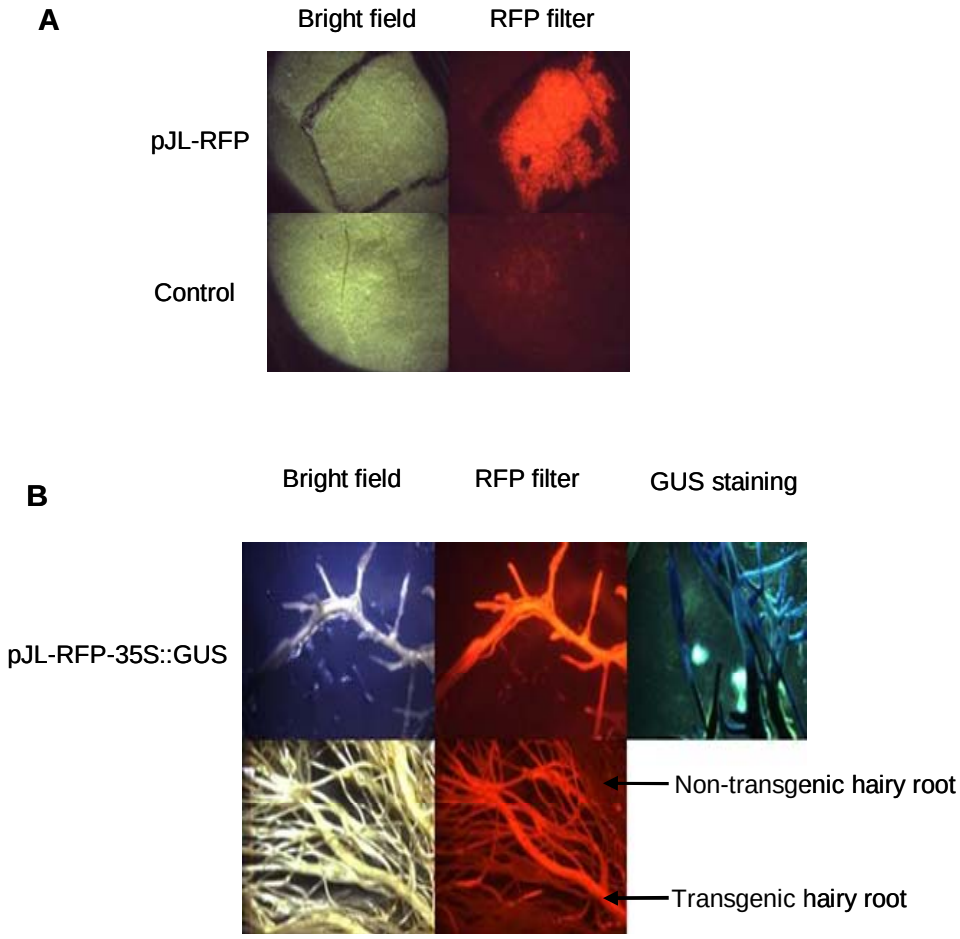

Figure 2.6. Test for reporter gene expression in tobacco leaves and soybean hairy roots. **A**, Transient expression of pJL-RFP and empty vector in tobacco leaves. **B**, Transgenic hairy roots harboring pJL-RFP-35S::GUS (top) and pJL-RFP-35S::GmSAMT1 (bottom). All the tested transgenic hairy roots with RFP signal showed GUS staining result. After inoculated by *A. rhizogenes* K599 harboring pJL-RFP-35S::GmSAMT1, transgenic hairy roots can be detected under a fluorescence stereomicroscope with the RFP filter.

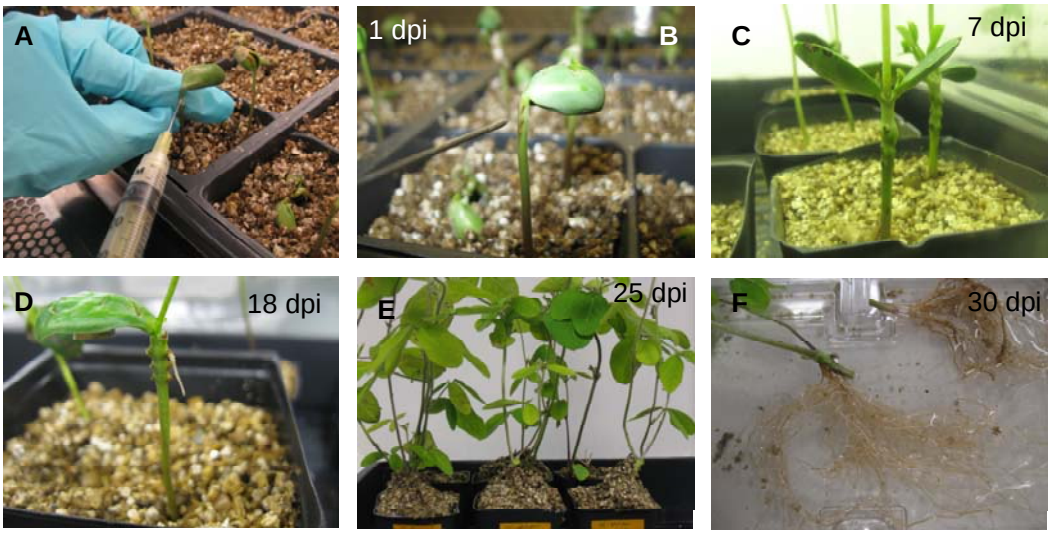

Figure 2.7. Different stages of soybean hairy root generation. **A**, Inoculation with *A. rhizogenes* cultures; **B**, Infection sites of soybean seedling at 1 day post inoculation (dpi); **C**, The swollen infection sites of soybean seedling at 7dpi; **D**, The hairy roots shown at 18 dpi; **E**, Hairy roots covered by sterile wet vermiculite; **F**, Tap root removed from soybean plant and healthy hairy roots left on soybean plant.

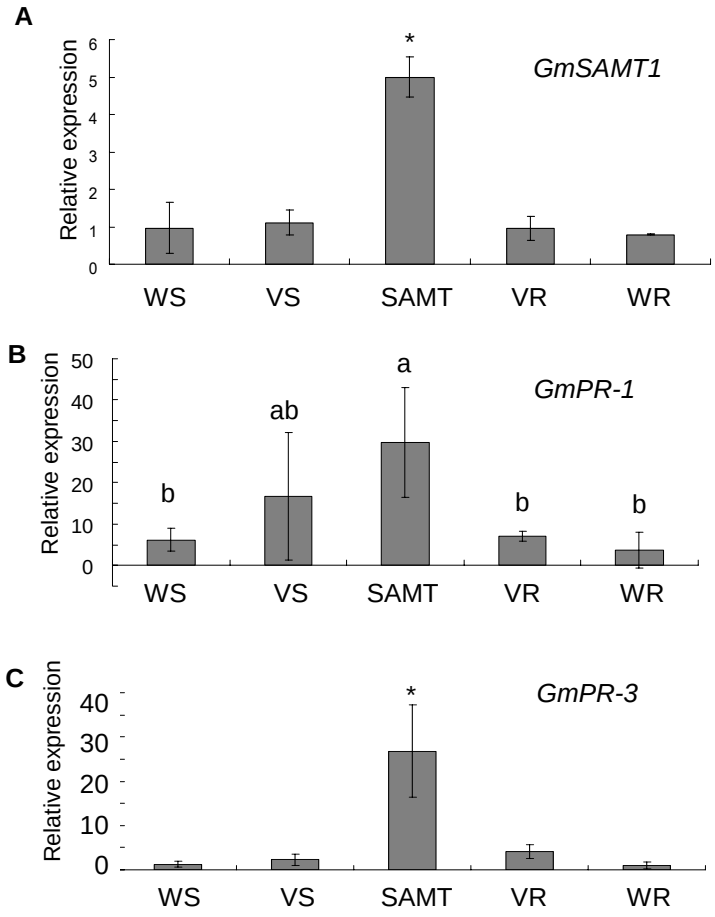

Figure 2.8. qRT-PCR analysis on transgenic hairy roots for *GmSAMT1*, *GmPR-1*, and *GmPR-3*. WS and WR represent the hairy root lines without any vector in TN02-275 and TN02-226 background, respectively; VS and VR represent the hairy root lines harboring plasmid pJL-RFP in TN02-275 and TN02-226 background, respectively; SAMT represents the hairy root line harboring plasmid pJL-RFP-35S::GmSAMT1 in TN02-275 background. Each bar represents the mean relative expression level of three independent experiments with the standard errors. In graphs **A** and **C**, the bars with stars represent the relative gene expression of hairy root lines that are significantly different compared with other soybean hairy root lines (LSD $p<0.05$). In graph **B**, bars with different letters are significantly different (LSD $p<0.05$).

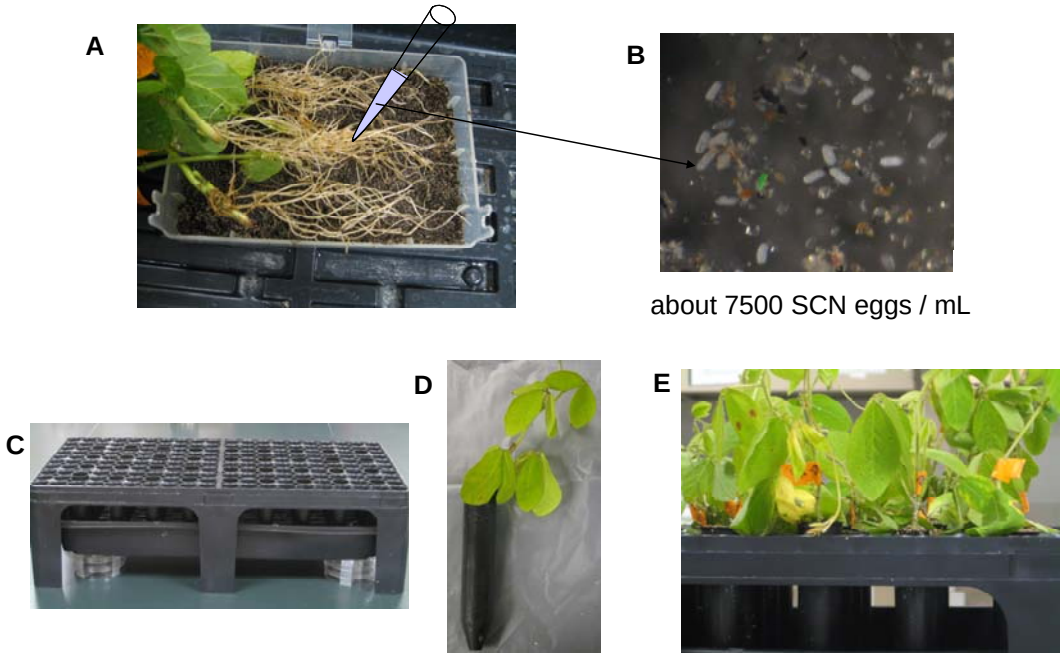

Figure 2.9. Inoculation of soybean hairy roots with SCN eggs. **A.** The hairy roots were inoculated by SCN eggs; **B.** SCN (race 2) eggs were used to inoculate soybean hairy roots; **C.** The rack holding cone-tainers was used to support soybean plants. **D.** Soybean plant was placed in the cone-tainer; **E.** The cone-tainters were loaded in the rack to grow soybean plants.

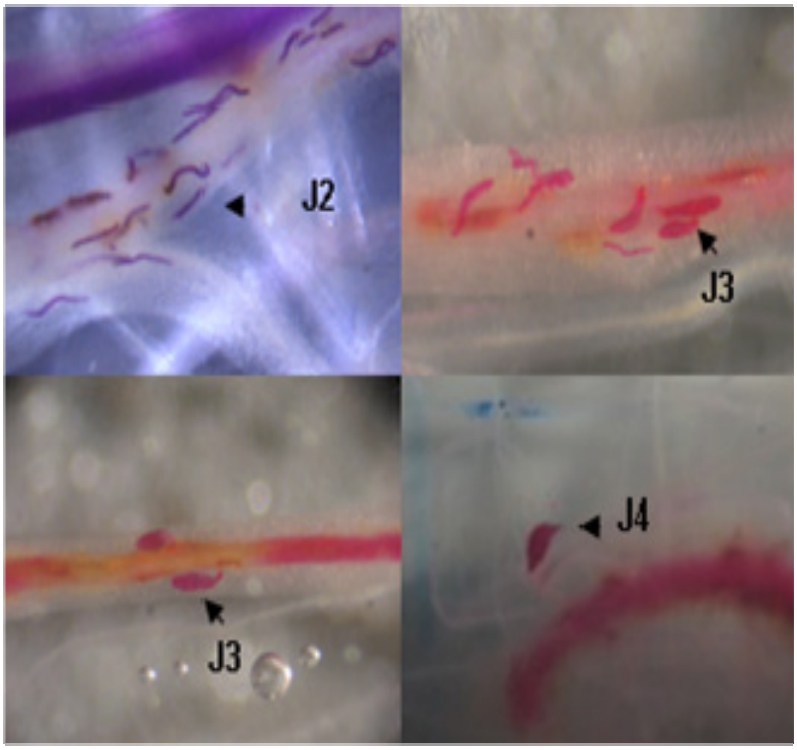

Figure 2.10. J2, J3, and J4 stages of soybean cyst nematodes shown in transgenic hairy roots by clearing and staining with acid fuchsin.

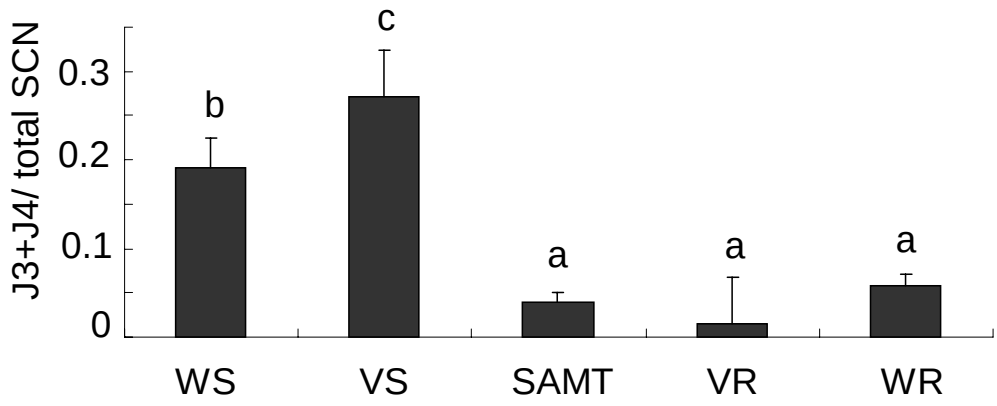

Figure 2.11. Nematode demographic assay for SCN resistance between transgenic hairy roots with over-expression of *GmSAMT1* and control soybean lines. Each bar represents the mean of resistance index with the standard error from 10 soybean hairy roots. The resistance index is the ratio of the sum of J3 and J4 nematodes to total infecting nematodes shown in *A. rhizogenes*-transformed hairy roots after two-week post inoculation by SCNs. WS and WR represent the hairy root lines without any vector in TN02-275 (susceptible to SCN race 2) and TN02-226 (resistant to SCN race 2) background, respectively; VS and VR represent the hairy root lines with over-expressing RFP reporter gene in TN02-275 and TN02-226 background, respectively as the vector control lines; SAMT represents the hairy root line with over-expression of *GmSAMT1* and RFP reporter gene in TN02-275 background. Bars with different letters are significantly different (LSD $p<0.05$).

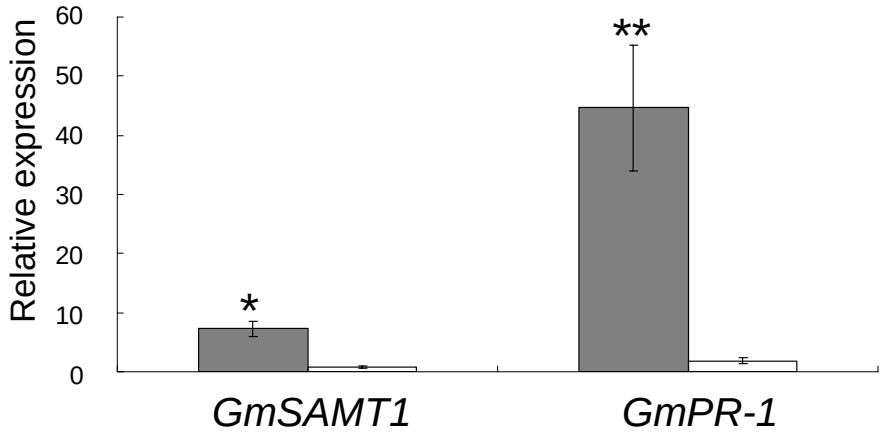

Figure 2.12. qRT-PCR analysis on soybean mosaic virus infected soybean leaves for *GmSAMT1* and *GmPR-1*. The grey and white bars represent the relative gene expression from the soybean plants with the resistant and susceptible responses, respectively. Each bar represents the mean relative expression level of three independent experiments with the standard errors. The bars with star(s) mean the gene expression are significantly different with the bars without star(s) (LSD $p<0.05$).

Chapter III. Methyl Salicylate Esterase Is Involved in Resistance against Soybean Cyst Nematode

Jingyu Lin, Mitra Mazarei, Nan Zhao, Vincent R. Pantalone, Prakash R. Arelli, Charles Neal Stewart, Jr, Feng Chen. Soybean Methyl Salicylate Esterase Is Involved in Resistance against Soybean Cyst Nematode. Drafted

ABSTRACT

There have been studies showing that methyl salicylate (MeSA) serves as a phloem-mobile systemic acquired resistance (SAR) signal in tobacco (*Nicotiana tabacum*) and *Arabidopsis thaliana*. However, it remains a question whether MeSA functions as a defense component in soybean [*Glycine max* (L.) Merr.]. In a previous microarray analysis on the gene expression of two genetically-related soybean sister lines under the attack of soybean cyst nematode (SCN), one candidate gene, *Glyma16g26060*, showed induced expression in the resistant line while suppressed expression in the susceptible line. *Glyma16g26060* was predicted to encode an α/β fold hydrolase. After cloning this gene and performing biochemical activity assay, this gene was identified to encode a methyl salicylate esterase. Since this gene was the ortholog gene of methyl salicylate esterase from tobacco, which was named *NtSABP2*, we designated the soybean gene *GmSABP2-1*. The recombinant enzyme encoded by *GmSABP2-1* can catalyze the conversion from MeSA to salicylic acid (SA). To explore the biological role of GmSABP2-1 in soybean resistance against SCN, we generated transgenic hairy roots with over-expression of *GmSABP2-1* and red fluorescent protein (RFP) gene in the susceptible soybean background. After two weeks post inoculation by SCN eggs, the demographics of SCN in over-expressed *GmSABP2-1* transgenic hairy roots and control hairy roots were compared. The results showed that the transgenic hairy roots with over-expressed *GmSABP2-1* in the susceptible soybean line had increased resistance, which was similar to the resistant soybean control line. Therefore, we conclude GmSABP2-1 plays a role in soybean resistance against SCN through

SA-dependent signaling pathway.

KEYWORDS: methyl salicylate esterase, soybean cyst nematode, defense

INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] is the most important legume crop in the world, providing both oil and protein products. Among the soybean pathogens, the soybean cyst nematode (*Heterodera glycines,* SCN) causes the most severe damage. It was reported that the global soybean yield losses due to SCN infection reached $2 billion a year (Niblack et al. 2004). SCN is a sedentary obligate endoparasite of soybean roots. Since most stages of nematodes are not detectable by the naked eye and its symptom is similar with nutrient deficiency (Kulkarni et al. 2008), the most accurate diagnosic method is to analyze the soil sample for SCN for visual observation. There are usually several ways to manage SCN, including crop rotation, applying nematicides, and growing SCN resistant soybean cultivars (Winter et al. 2006). However, if one resistant cultivar is grown in the same field for several years, SCN will adapt to the cultivar and conquer the resistance. To maintain resistance against SCN, soybean growers need to alternate use of soybean cultivars with different sources of SCN resistance, and grow a susceptible cultivar during the rotation.

To develop new soybean cultivars with enhanced resistance against SCN, researches to elucidate mechanism of soybean defense have been performed from different aspects. Until now, there have been five genes reported to be SCN resistance genes, including three recessive (*rhg1, rhg2* and *rhg3*) and two dominant genes (*Rhg4* and *Rhg5*) (Caldwell et al. 1960; Matson and Williams, 1965; Rao-Arelli 1994). Recently, to study the function of *rhg1,* Melito et al. (2010) introduced one putative SCN resistance gene located proximal to *rhg1* locus, *Glyma18g02680.1* encoding an

apparent leucine-rich repeat transmembrane receptor-kinase (LRR-kinase), into *rhg1*-SCN-susceptible plant lines through transgenic hairy roots. However, the result did not show the significant increase in SCN resistance, which suggested that one or more other genes at the *Rhg1* locus may be the primary determinant of differences in SCN resistance among different soybean types (Melito et al. 2010).

To identify the novel SCN resistance genes, a number of microarray studies have been performed to study gene expression changes during the interaction of soybean and SCN (Khan et al. 2004; Klink et al. 2007, 2009, 2010; Lu et al. 2006; Puthoff et al. 2007). A variety of genes had significant expression changes during the interaction of soybean and SCN. One large group of the genes was associated with cell wall structure. Other functional categories consist of genes involved in defense and metabolism.

The study of syncytium gene expression revealed some pathways playing roles in resistance against SCN, including jasmonic acid (JA) biosynthesis, 13-lipoxygenase pathway, S-adenosyl methionine pathway, phenylpropanoid biosynthesis, suberin biosynthesis, adenosylmethionine biosynthesis, ethylene biosynthesis from methionine, flavonoid biosynthesis, and the methionine salvage pathway (Klink et al. 2010). The authors also proposed the JA defense pathway as a factor involved in the localized resistant reaction.

In Chapter 2 of this dissertation, we have discussed the role of salicylic acid (SA) in SCN resistance and methyl salicylate (MeSA) may also work as a defense agent in soybean defense against pathogens. The over-expression of a soybean salicylic acid methyltransferase gene, *GmSAMT1*, in the susceptible soybean line showed increased

resistance. To further study whether MeSA functions as a signal molecule and which component works for receiving this signal, we aim to investigate other genes involved the MeSA metabolism. In recent years, an enzyme, SA-binding protein 2 (SABP2), has been demonstrated to catalyze the conversion of MeSA into SA. This enzyme, NtSABP2, was first identified from tobacco (*Nicotiana tabacum*) using a biochemical approach and showed specifically high affinity with SA, so it was named SABP2 (Kumar and Klessig, 2003). SABP2 belongs to α/β fold hydrolase gene family. Silencing of *SABP2* in tobacco by RNAi caused the loss of systemic acquired resistance (SAR) (Kumar and Klessig, 2003; Kumar et al. 2006) and a grafting experiment (Park et al. 2007), which suggested that SABP2 played a critical role in SAR of tobacco. Some SABP2 homologs genes have been identified and studied in several plant species including tobacco, *Arabidopsis thaliana*, poplar (*Populus trichocarpa*), and potato (*Solanum tuberosum*) (Kumar and Klessig, 2003; Manosalva et al. 2010; Vlot et al. 2008; Zhao et al. 2009), which suggested that MeSA may be a conserved SAR signal in plants. A phenomenon similar to SAR was reported for soybean-parasitic nematodes (Ibrahim and Lewis 1986). In their study, soybean cultivar 'Centennial', normally susceptible to root knot nematode *Meloidogyne arenaria*, showed increased resistance to this nematode when the plant was pre-inoculated with *M. incognita* that is another root knot nematode species closely related to *M. arenaria*. It is necessary to explore whether SAR plays an important role in soybean defense against SCN.

Recently, through comparison of gene expression profiling of two soybean recombinant inbred lines, resistant and susceptible to SCN race 2, under the infection of

SCN, Mazarei et al. (2011) found that some genes associated with SA pathway were highly induced in the resistant soybean line whereas they were suppressed in the susceptible soybean line. One of these genes was predicted to be α/β fold hydrolase gene and related to tobacco SABP2.

In this study, we cloned the above mentioned gene from soybean which was named *GmSABP2-1* and demonstrated that it encodes methyl salicylate esterase. To examine its role in soybean resistance, *Agrobacterium rhizogenes*-mediated transgenic hairy roots were generated that had over-expression of *GmSABP2-1* in the susceptible soybean line. The resistance levels between transgenic hairy root line and control lines were compared by monitoring the infecting population for progression through nematode life stages.

MATERIALS AND METHODS

Plant, Nematode, Bacteria, and Chemical Sources. In this study, two genetically-related soybean lines, TN02-226 and TN02-275, were used for gene cloning and generating soybean hairy roots. The breeding details for these two soybean lines and have been described in Chapter 2. SCN race 2 eggs were used as the pathogens in the bioassay of the hairy roots, with the same inoculation method described in Chapter 2. *Escherichia coli* strain BL21 (DE3) CodonPlus (Stratagene, La Jolla, CA) was used for expressing the recombinant protein. *A. rhizogenes* strain K599 was utilized for carrying the over-expressed *GmSABP2-1* and generating soybean hairy roots. The chemicals used in this study were purchased from Sigma-Aldrich (St. Louis, MO).

Database Search and Sequence Analysis. Both the sequences of Affymetrix probes GmaAffx.92649.1.S1_at and Gma.5867.1.A1_s_at showed 100% similarity with one soybean gene, *Glyma16g26060*. Based on the annotation in phytozome (http://www.phytozome.net/soybean), this gene encodes an α/β fold hydrolase. Alignment of Glyma16g26060 with two other characterized methyl ester esterase was performed to study the conserved domains of this candidate protein.

Isolation of Glyma16g26060 Gene from Soybean. Total RNA extraction and DNA contamination removal were performed as described in Chapter 2. To further study the function of Glyma16g26060, the full length cDNA sequence of Glyma16g26060 was obtained from public soybean database (http://www.phytozome.net/soybean). Using the specific primers 5'- ATGGGTTCACAAAATTGTATGGATAGG-3' and 5'-TCATGCATATTTAGTCGCTATCTGCTG-3', thermal cycling conditions were: 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 1 min, and a final extension at 72 °C for 10 min. The PCR product was cloned into vector pEXP5/CT-TOPO. The sequencing result of the PCR product from TN02-226 soybean was consistent with that predicted in the soybean database on Williams 82 soybean.

To verify its biochemical function, the candidate gene was then cloned into a protein purification vector, pET100/D-TOPO vector (Invitrogen, Carlsband, CA) and expressed in the *E. coli* strain BL21 (DE3) CodonPlus (Stratagene, La Jolla, CA). Protein expression was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) at a concentration of 500 μM for 18 h at 25°C, with cells lysed by sonication. *E.*

coli-expressed Glyma16g26060 with a His-tag was purified from the *E. coli* cell lysate using Ni-NTA agarose following the manufacturer instructions (Invitrogen). Protein purity was verified by SDS–PAGE and protein concentrations were determined by the Bradford assay (Bradford, 1976).

Methyl Ester Esterase Activity Assay. A two-step radiochemical esterase assay was performed to determine the activity of *E. coli*-expressed Glyma16g26060 following a protocol previously reported (Forouhar et al. 2005). The reaction of the first step was the reaction of substrate 10 µM MeSA and purified recombinant *E. coli*-expressed Glyma16g26060 for 30 min at 25°C. After that, the sample was boiled to stop the reaction and denature the enzyme. The second reaction step was started with the addition of 3 µM ^{14}C labeled S-adenosyl-L-methionine (SAM) with a specific activity of 51.4 mCi/mmol (Perkin Elmer, Boston, MA, USA), and respective known purified methyltransferase with high activity. Here, GmSAMT1 mentioned in previous chapter was used for SA methyltransferase enzyme. The reaction was preceded for 30 min at 25 °C. Then radiolabeled products were extracted and the radioactivity was counted using a liquid scintillation counter (Beckman Coulter, Fullerton, CA, USA) as previously described (D'Auria et al. 2002).

Determination of Kinetic Parameters of Soybean SABP2. The increase in reaction rate with increasing concentrations of MeSA was evaluated with the radiochemical assay described above and was found to obey Michaelis–Mention kinetics. Appropriate

enzyme concentrations and incubation time were determined in time-course assays so that the reaction velocity was linear during the assay period. To determine the Km for MeSA, the concentrations of MeSA were independently varied in the range from 5 μM to 100 μM. Lineweaver–Burk plots were made to obtain apparent Km value, as previously described (Chen et al. 2003). Optimal pH for GmSABP2-1 was assayed from pH 6.0 to pH 9.0. Final values are an average of three independent measurements.

Evaluation of SCN Resistance of Transgenic Hairy Roots with Over-expression of *GmSABP2-1*. In this study, transgenic soybean hairy roots with over-expressed *GmSABP2-1* were utilized to evaluate the biological function of GmSABP2-1. The steps including creating the construct with over-expressed *GmSABP2-1* and RFP reporter gene, generating the transgenic soybean hairy roots harboring over-expressed *GmSABP2-1*, the SCN inoculation, and demographics assay of nematodes were performed as described in Chapter 2, but *GmSABP2-1* gene was used in place of *GmSAMT1* gene.

Statistical Analysis. The resistance difference between soybean hairy roots with over-expressing *GmSABP2-1* and control soybean hairy roots was analyzed by one way ANOVA using the mixed model of SAS (SAS 9.2 version), and least squares means of the resistance index were separated using Fisher's LSD test. A $P<0.05$ was considered significant.

RESULTS

Identification and Sequence Analysis of One Putative SCN-Resistant Gene in Soybean. Recently, microarray analysis revealed some genes showing significant expression changes during resistant and susceptible reactions to SCN. One of these genes was significantly up-regulated only in SCN inoculated resistant soybean roots whereas it was significantly down-regulated in SCN inoculated susceptible soybean roots through the microarray analysis and quantitative real time PCR (Mazarei et al, 2011).

To obtain the complete open reading frames of this resistance related gene, the sequences of Affymetrix SoyChip probe GmaAffx.92649.1.S1_at and Gma.5867.1.A1_s_at which identified our target gene in microarray analysis (Mazarei et al, 2011), were used as a query to BLAST search against the soybean database GmGDB/GMtranscript (http://www.plantgdb.org/GmGDB/cgi-bin/blastGDB.pl), soybean gene *Glyma16g26060* was found to show the highest similarity (100%) with the sequence of the above two probes and predicted to belong to α/β hydrolase gene superfamily.

***Glyma16g26060* Encodes Soybean Methyl Ester Esterase.** A BLAST search revealed that Glyma16g26060 is 61.2% identical to the previously characterized NtSABP2 (GenBank accession AAR87711.1) at the amino acid sequence level, and its identity with the NtSABP2 is the highest among the soybean homolog genes. The alignment of peptide sequences among Glyma16g26060 and NtSABP2, At4g37150 was performed

(**Figure 3.1)**. Glyma16g26060 contains the catalytic triad (Ser-86, His-239, and Asp-211) that is highly conserved among SABP2 orthologs from tobacco (Forouhar et al. 2005), Arabidopsis (AtMES) (Vlot et al. 2008), poplar (PtSABP2-1 and -2) (Zhao et al. 2009) and potato (StMES1) (Manosalva et al. 2010), as well as the closely related members of the α/β hydrolase superfamily (Forouhar et al. 2005). Glyma16g26060 also contains all the 15 amino acids identified in NtSABP2 that interacts with SA (**Figure 3.1**) (Forouhar et al. 2005). In contrast, AtMES9 contains only 5 and StMES1 shares 12 of these 15 residues. The high sequence similarity and conservation of critical SA-binding residues between *Glyma16g26060* and NtSABP2 suggested that these two proteins shared similar biochemical properties. To confirm the prediction from the databases, the full-length cDNA of the Glyma16g26060 was cloned via reverse-transcription polymerase chain reaction (RT-PCR) from SCN-infected root tissue of soybean TN02-226 line and sequenced. The sequenced gene contained 786 bp, encoding a protein of 261 amino acids, which was the same as *Glyma16g26060* from database.

Biochemical Properties of Soybean SABP2. To further explore the biochemical properties of the enzyme encoded by *Glyma16g26060*, this gene was cloned into a protein expression vector. Recombinant Glyma16g26060 with a His-tag expressed in *E. coli* was purified to electrophoretic homogeneity and subjected to detailed biochemical characterization. The molecular mass of the Glyma16g26060 with His-tag protein estimated on the gel was 33 kDa (**Figure 3.2)**, which was close to the expected size of

our target enzyme. Purified recombinant Glyma16g26060 showed the activity of converting MeSA to SA. Under steady-state conditions, Glyma16g26060 hydrolyzed MeSA with an apparent Km value of 46.2±2.2 μM (**Figure 3.3**). So Glyma16g26060 was designated as soybean SABP2 (GmSABP2-1). When enzyme assays were performed in buffers with different pH conditions, GmSABP2-1 showed the highest level of catalytic activity at around pH 7.0 (**Figure 3.4**).

Biological Role of Soybean SABP2 in SCN Resistance. The study of the biological role of SABP2 in other plant species was performed through observing the leaves' resistance under virus or bacteria infection. Since the soybean resistance against SCN occurs in the root tissue, the number of SCNs that exist in the transgenic hairy roots at different stages were recorded to analyze the resistance level and evaluate the biological function of GmSABP2-1. As mentioned in Chapter 2, the different resistant responses to SCN race 2 between the susceptible (TN02-275) and resistant (TN02-226) soybean lines occurred on the J3 stage of SCN, the time point after two-week inoculation by SCNs was chosen for analyzing the demographics of SCN in all the hairy roots with respective constructs. The index of J3+J4/total SCN number was used to estimate the resistance between the susceptible line with over-expressed *GmSABP2-1* and other control lines, the same with the assay described in Chapter 2. The susceptibility level of vector control in the susceptible background was arbitrarily set at 100%. The results showed the transgenic soybean hairy roots with over-expressed *GmSABP2-1* had significantly fewer female SCN juvenile developed to J3 stage, compared with vector

control line in the susceptible soybean background (84.5% reduction). And there was no significant difference on the resistance level between the transgenic hairy roots with over-expressed *GmSABP2-1* and vector control in the resistant soybean background, which also had significant fewer female SCN juvenile developed to J3 stage, compared with negative vector control in the susceptible soybean background (94.8% reduction) (**Figure 3.5**).

DISCUSSION

In this chapter, we described the identification of GmSABP2-1 from soybean and demonstrated that it shared high similarity and conserved motifs with its tobacco and Arabidopsis orthologs. Through further biochemical assay, it was indicated that recombinant GmSABP2-1 exhibited esterase activity toward MeSA. The Km of GmSABP2-1 (46.2 μM) (**Figure 3.3**) was higher than the Km of NtSABP2 (8.6 μM) (Forouhar et al. 2005) PtSABP2-2 (24.6 μM) (Zhao et al. 2009), and lower than the Km of AtMES1 (57.3 μM), AtMES9 (147.1μM) (Volt et al. 2008), PtSABP2-1(68.2 μM) (Zhao et al. 2009), and StMES1 (57.9 μM) (Manosalva et al. 2010). The differences in these kinetic parameters may reflect endogenous MeSA levels of plant species.

In this study, the ortholog gene of known *SABP2s* in soybean was first identified and characterized. In other plant species, such as tobacco, Arabidopsis and potato, the function of SABP2 was documented to be required for SAR; e.g., in tobacco leaves and tobacco mosaic virus (TMV) system (Park et al. 2007; Manosalva et al. 2010; Vlot et al. 2008). MeSA, the methyl ester of SA, was recently demonstrated to be a mobile signal

for SAR in tobacco and be transported from the damaged tissue to the undamaged distal tissue through phloem (Park et al. 2007). Considering that MeSA is required to convert back to SA for plant's SAR, the identification of soybean SABP2 will be vital to study the mechanism of soybean SAR.

In this chapter, the transgenic hairy roots with over-expressed *GmSABP2-1* in the susceptible soybean line, showed increased resistance against SCN (**Figure 3.5**). This result suggested that a certain expression level of *GmSABP2-1* was required for soybean defense against SCN. In study of Tenhaken and Rübel (1997), a soybean -*Pseudomonas syringae* pv *glycinea* system was used to study SA's role in soybean defense. SA was shown to play a role in hypersensitive cell death. The authors also proposed that the function of SA is mainly for SAR rather than an important component of the HR, since the relatively late accumulation of SA occurred during the HR. In our study, GmSABP2-1 was suggested to be an important component of receiving MeSA signal to induce SAR in soybean defense against SCN, considering the temporal and spatial differences of SCN on reaching the feeding sites of soybean vascular tissues within the SCN populations. The increased level of GmSABP2-1 protein by up-regulating its gene expression under SCN infection may benefit for capture of the mobile MeSA. Demethylation of MeSA catalyzed by GmSABP2-1 may be able to increase the SA level in the distal undamaged soybean root tissues since its ortholog NtSABP2 has the role of inducing SAR in the undamaged tobacco leaf (Park et al. 2007).

Taken together, the two components, GmSAMT1 and GmSABP2-1 work for the transportation of mobile signal MeSA and play a role in soybean defense against SCN.

The functions of GmSAMT1 and GmSABP2-1 and the model of soybean resistance against SCN are proposed (**Figure 3.6**). In SCN-infected local root, GmSAMT1 plays a critical role of regulating the levels of basal SA and MeSA. In the root knot nematode resistance research, the residual level of free SA was found to be sufficient in the basal resistance (Bhattarai et al. 2008). This is suggestive that the level of SA required for basal defense against SCN may be low. Biosynthesized SA can be changed to other forms, such as MeSA, to induce SAR in the distal root tissue. Relative high level of basal SA may inhibit the activity of GmSABP2-1 and therefore inhibit the conversation of MeSA to SA at the local site. This notion was supported by the characterization of GmSABP2-1's orthologs in tobacco and Arabidopsis (Forouhar et al. 2005; Volt et al. 2008). In the distal root tissue, the relative high level of MeSA and low level of SA favors the activity of GmSABP2-1, leading to the conversion of MeSA to SA. Of course, this model is mainly based on these two enzymes' biochemical function of regulating SA and MeSA levels at different SCN infected locations. To test and refine this model, it is critical to measure the concentrations of SA and its possible conjugated forms, which includes MeSA, salicyloyl glucose ester (SGE), SA O-β-glucoside (SAG), and SA-amino acid conjugates, which may function together to adjust the level of free SA *in vivo* (Vlot et al. 2009), at the local SCN infected root and distal undamaged root tissues.

The complete elucidation of the soybean SA-dependent defense mechanism may facilitate the increase of durable soybean resistance against SCN by manipulating SA and MeSA levels through suitable specific and inducible promoters driving *GmSAMT1* and *GmSABP2-1* in transgenic plants.

REFERENCES

Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 7: 248-254

Caldwell BE, Brim CA, Ross JP (1960) Inheritance of resistance of soybeans to the cyst nematode, *Heterodera glycines.* Agron. J. 52:635-636

Chen F, D'Auria JC, Tholl D, Ross JR, Gershenzon J, Noel JP, Pichersky E (2003) An *Arabidopsis thaliana* gene for methylsalicylate biosynthesis, identified by a biochemical genomics approach, has a role in defense. Plant J. 36: 577-588

D'Auria JC, Chen F, Pichersky E (2002) Characterization of an acyltransferase capable of synthesizing benzylbenzoate and other volatile esters in flowers and damaged leaves of *Clarkia breweri*. Plant Physiol. 130: 466-476

Forouhar F, Yang Y, Kumar D, Chen Y, Fridman E, Park SW, Chiang Y, Acton TB, Montelione GT, Pichersky E, Klessig DF, Tong L (2005) Structural and biochemical studies identify tobacco SABP2 as a methyl salicylate esterase and implicate it in plant innate immunity. Proc. Natl. Acad. Sci. USA 102: 1773-1778

Ibrahim I, Lewis S (1986) Interrelationships of *Meloidogyne arenaria* and *M. incognita* on Tolerant Soybean. J. Nematol. 18: 106-111

Khan R, Alkharouf N, Beard H, MacDonald M, Chouikha I, Meyer S, Grefenstette J, Knap H, Matthews B (2004) Microarray analysis of gene expression in soybean roots susceptible to the soybean cyst nematode two days post invasion. J. Nematol. 36: 241-248

Klink VP, Hosseini P, Matsye P, Alkharouf NW, Matthews BF (2009) A gene expression analysis of syncytia laser microdissected from the roots of the *Glycine max* (soybean) genotype PI 548402 (Peking) undergoing a resistant reaction after infection by *Heterodera glycines* (soybean cyst nematode). Plant Mol. Biol. 71: 525-567

Klink VP, Hosseini P, Matsye PD, Alkharouf NW, Matthews BF (2011) Differences in gene expression amplitude overlie a conserved transcriptomic program occurring between the rapid and potent localized resistant reaction at the syncytium of the *Glycine max* genotype Peking (PI 548402) as compared to the prolonged and potent resistant reaction of PI 88788. Plant Mol. Biol. 75: 141-165

Klink VP, Overall CC, Alkharouf NW, MacDonald MH, Matthews BF (2007) A time-course comparative microarray analysis of an incompatible and compatible response by *Glycine max* (soybean) to *Heterodera glycines* (soybean cyst nematode)

infection. Planta 226: 1423-1447

Klink VP, Overall CC, Alkharouf N, MacDonald MH, Matthews BF (2010) Microarray detection calls as a means to compare transcripts expressed within syncytial cells isolated from incompatible and compatible soybean (*Glycine max*) roots infected by the soybean cyst nematode (*Heterodera glycines*). J. Biomed Biotechnol. 2010:491217 (1-30)

Kulkarni SS, Bajwa SG, Rupe JC, Kirkpatrick TL (2008) Spatial correlation of crop response to soybean cyst nematode (*Heterodera Glycines*). T. Asabe 51: 1451-1459

Kumar D, Gustafsson C, Klessig DF (2006) Validation of RNAi silencing specificity using synthetic genes: salicylic acid-binding protein 2 is required for innate immunity in plants. Plant J. 45: 863-868

Kumar D, Klessig DF (2003) High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity. Proc. Natl. Acad. Sci. USA 100: 16101-16106

Li J, Brader G, Palva ET (2004) TheWRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell 16: 319-331

Li J, Todd TC, Lee J, Trick HN (2011) Biotechnological application of functional genomics towards plant-parasitic nematode control. Plant Biotechnol. J. 10.1111/j.1467-7652.2011.00601.x

Lu P, Shannon JG, Sleper DA, Nguyen HT, Cianzio SR, Arelli R (2006) Genetics of cyst nematode resistance in soybean PIs 467312 and 507354. Euphytica 149: 259-265

Manosalva PM, Park S-W, Forouhar F, Tong L, Fry WE, Klessig DF (2010) Methyl esterase 1 (StMES1) is required for systemic acquired resistance in potato. Mol. Plant Microbe In. 23: 1151-1163

Matson AL, Williams LF (1965) Evidence of a fourth gene for resistance to the soybean cyst nematode. Crop Sci. 5:477

Mazarei M, Elling AA, Maier TR, Puthoff DP, Baum TJ (2007) GmEREBP1 is a transcription factor activating defense genes in soybean and Arabidopsis. Mol. Plant Microbe In. 20: 107-119

Mazarei M, Liu W, Al-Ahmad H, Arelli P, Pantalone V, Stewart C (2011) Gene expression profiling of resistant and susceptible soybean lines infected with soybean cyst nematode. Theor. Appl. Genet. 123: 1193-1206

Melito S, Heuberger AL, Cook D, Diers BW, MacGuidwin AE, Bent AF (2010) A nematode demographics assay in transgenic roots reveals no significant impacts of the Rhg1 locus LRR-Kinase on soybean cyst nematode resistance. Bmc Plant Biol. 10: 104
Park SW, Kaimoyo E, Kumar D, Mosher S, Klessig DF (2007) Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. Science 318: 113-116

Puthoff DP, Ehrenfried ML, Vinyard BT, Tucker ML (2007) GeneChip profiling of transcriptional responses to soybean cyst nematode, *Heterodera glycines*, colonization of soybean roots. J. Exp. Bot. 58: 3407-3418

Rao-Arelli AP (1994) Inheritance of resistance to *Heterodera glycines* race 3 in soybean accessions. Plant Dis. 78: 898-900

Spoel SH, Koornneef A, Claessens SM, Korzelius JP, Van Pelt JA, Mueller MJ, Buchala AJ, Metraux JP, Brown R, Kazan K, Van Loon LC, Dong X, Pieterse CM (2003) NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol. Plant Cell 15: 760-770

Tenhaken R, Rubel C (1997) Salicylic acid is needed in hypersensitive cell death in soybean but does not act as a catalase inhibitor. Plant Physiol. 115: 291-298

Vlot AC, Dempsey DA, Klessig DF (2009) Salicylic acid, a multifaceted hormone to combat disease. Annu. Rev. Phytopathol. 47: 177-206

Vlot AC, Liu P-P, Cameron RK, Park S-W, Yang Y, Kumar D, Zhou F, Padukkavidana T, Gustafsson C, Pichersky E, Klessig DF (2008) Identification of likely orthologs of tobacco salicylic acid-binding protein 2 and their role in systemic acquired resistance in *Arabidopsis thaliana*. Plant J.56: 445-456

Winter SMJ, Rajcan I, Shelp BJ (2006) Soybean cyst nematode: Challenges and opportunities. Can. J. Plant Sci. 86: 25-32

Zhao N, Guan J, Forouhar F, Tschaplinski TJ, Cheng Z-M, Tong L, Chen F (2009) Two poplar methyl salicylate esterases display comparable biochemical properties but divergent expression patterns. Phytochemistry 70: 32-39

Appendix

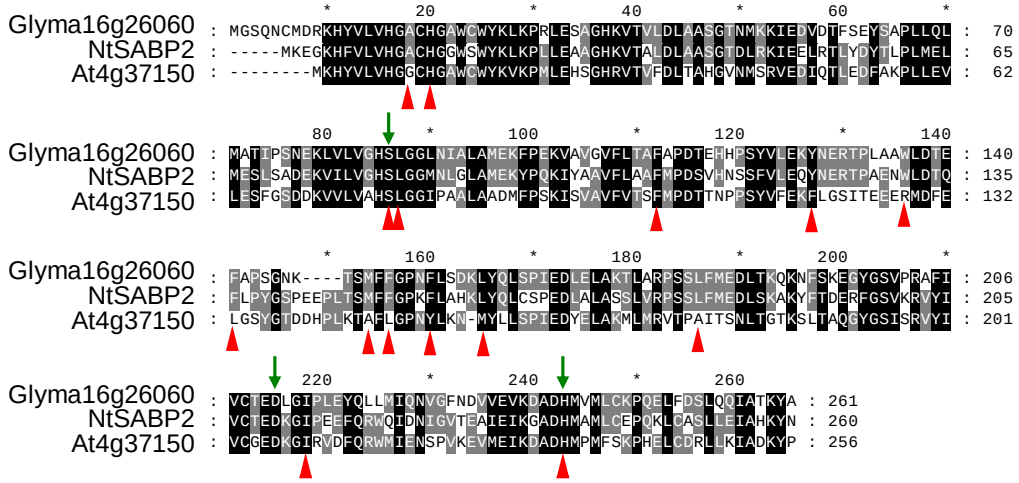

Figure 3.1. Multiple sequence alignment of soybean Glyma16g26060, tobacco SABP2 (NtSABP2), and an Arabidopsis ortholog of SABP2 (At4g37150). Identical residues are shaded in black and similar residues in gray. The catalytic triad residues are indicated by green arrows, and residues that contact salicylic acid are indicated with red triangles.

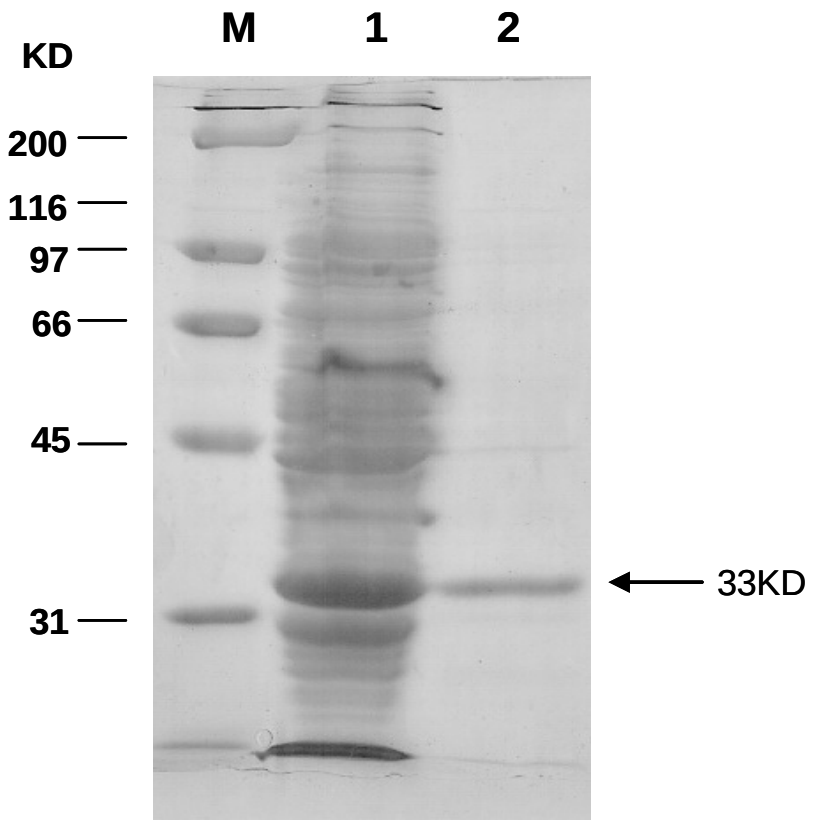

Figure 3.2. SDS-PAGE of purified recombinant GmSABP2-1 protein. His-tagged GmSABP2-1 expressed in *E. coli* was purified as described in the materials and methods. Lane M contained protein molecular weight markers. Lane 1 contained crude extract and lane 2 contained about 2 μg of purified GmSABP2-1 protein. The gel was stained with Coomassie blue.

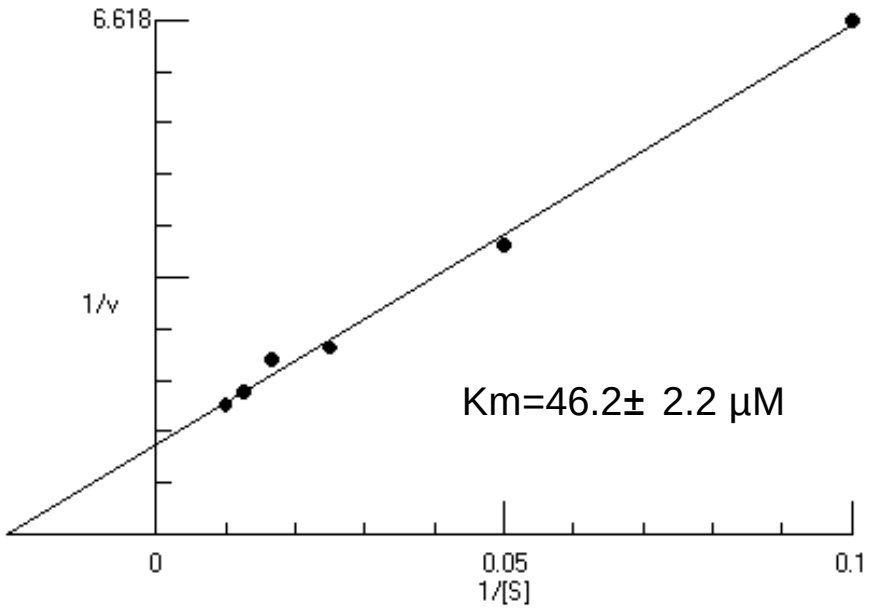

Figure 3.3. Km value of GmSABP2-1 using MeSA as substrate at concentrations of 5 µM to 100 µM. The apparent Km value value was obtained through Lineweaver-Burk plots.

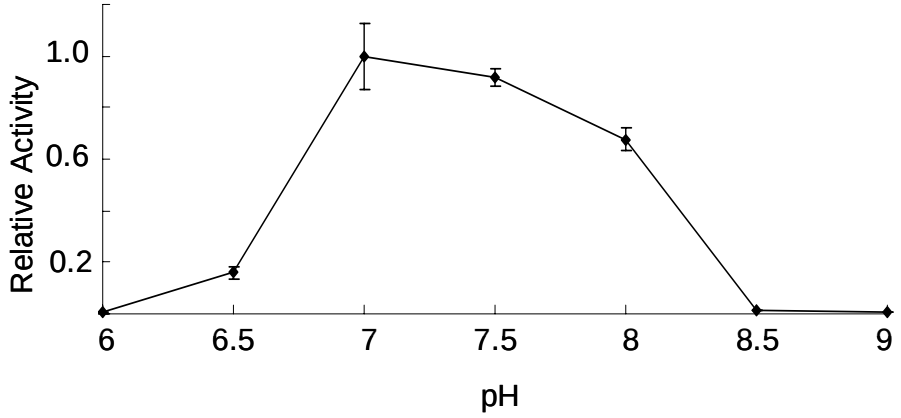

Figure 3.4. pH effect on GmSABP2-1 activity. Activity of the purified GmSABP2-1 was assayed from pH 6.0-9.0. Level of GmSABP2-1 activity in 50 mM Tris-HCl buffer, pH 7.0, was arbitrarily set at 1.0.

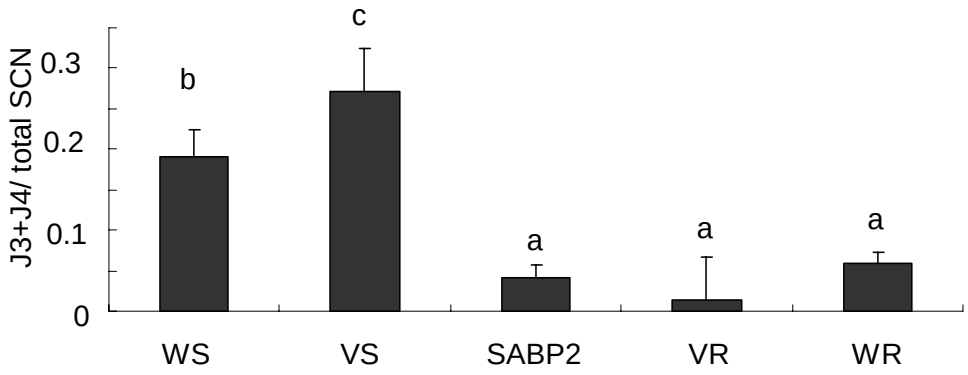

Figure 3.5. Nematode demographic assay for SCN resistance between transgenic hairy roots with over-expression of *GmSABP2-1* and control soybean lines. Ratio of the sum of J3 and J4 nematodes to the total infecting nematodes shown in hairy roots is used as the resistance index. The resistance index, WS, WR, VS and VR have been described in Chapter 2. WS and WR are wild type hairy root lines in the susceptible and resistant background, respectively. VS and VR are hairy root lines harboring RFP reporter gene in the susceptible and resistant background, respectively as the vector control lines. SABP2 represents the hairy root line with over-expressed *GmSABP2-1* and RFP reporter gene in TN02-275 background. Bars with different letters are significantly different (LSD $p<0.05$).

Figure 3.6. The model of soybean resistance against SCN and the functions of GmSAMT1 and GmSABP2-1. **A.** The reactions between SA and MeSA are catalyzed by GmSAMT1 and GmSABP2-1. SAM represents S-adenosyl-L-methionine functioning as the donor of methyl group (shown in red circle). **B.** In the resistant soybean root, the gene expression of *GmSAMT1* and *GmSABP2-1* is up-regulated when soybean is infected by SCN (Mazarei et al. 2011). The efficient transportation of mobile signal MeSA increases the soybean resistance against SCN. In the susceptible soybean root, the gene expression of GmSAMT1 does not change and the gene expression of GmSABP2-1 is down-regulated, when soybean is infected by SCN (Mazarei et al. 2011). The transportation of mobile signal MeSA is blocked, which decreases the soybean resistance against SCN. There are more SCNs completing their life cycle in the susceptible than the resistant soybean root. **C.** In SCN-infected local soybean root, SA is biosynthesized and the level of SA increases. The level of SA can be manipulated by GmSAMT1 through converting to mobile signal MeSA. The residual level of SA is sufficient for inducing *GmPR-1* gene expression. The decreased level of SA benefits jasmonic acid (JA)-dependent defense pathway, which induces *GmPR-3* gene expression. In undamaged soybean root, MeSA can be catalyzed by GmSABP2-1 to form SA, which plays a role of systemic acquired resistance against further SCN infection. Non-expressor of pathogenesis-related genes 1 (NPR1), WRKY family transcription factor, and lipoxygenase (LOX) may function as the regulatory points (Li et al. 2004; Klink et al. 2007; Spoel et al. 2003). The red dots represent the adult female SCNs within soybean roots.

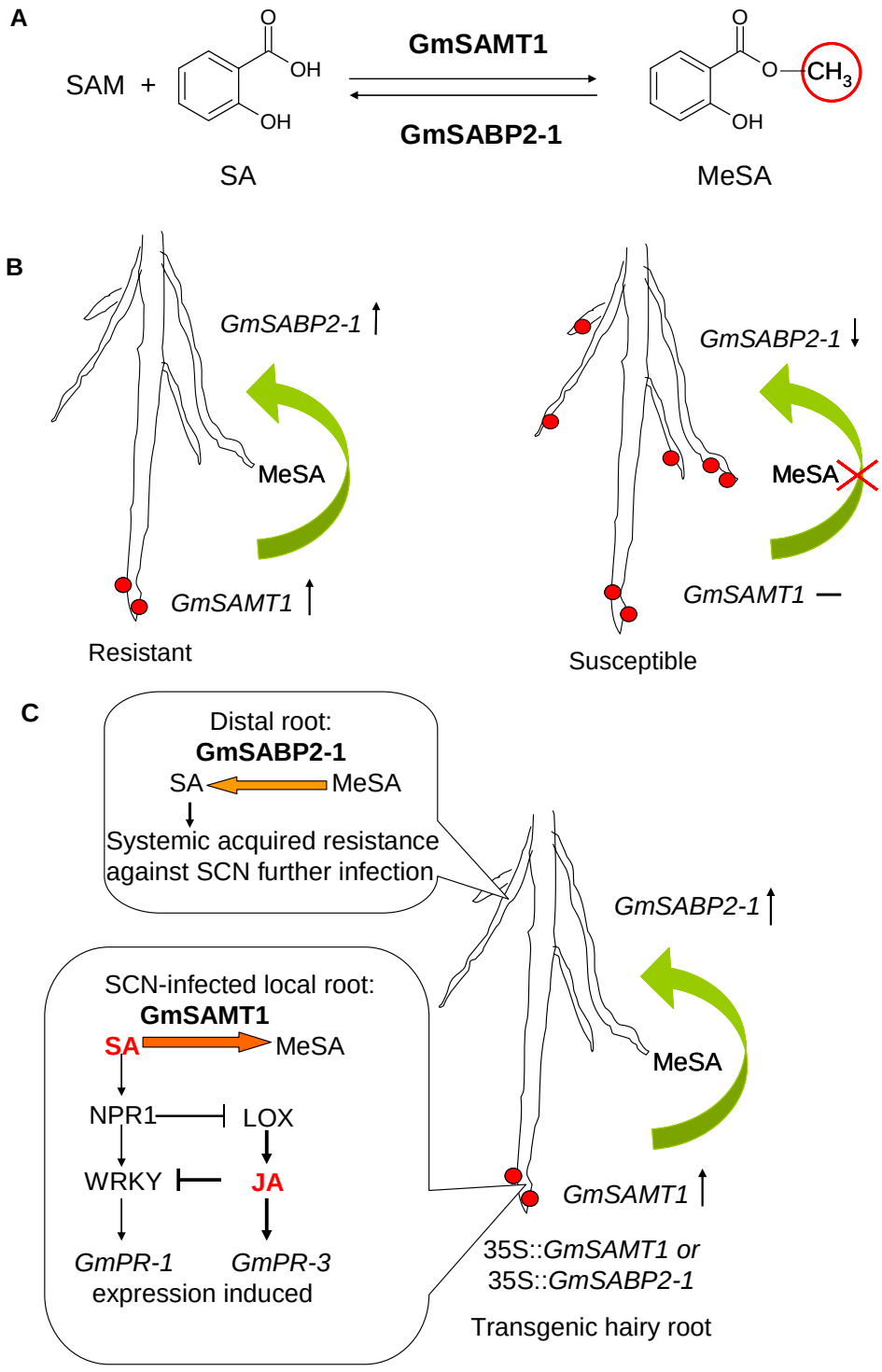

Figure 3.6. The model of soybean resistance against SCN and the functions of GmSAMT1 and GmSABP2-1.

Chapter IV. Identification of Soybean Terpene Synthase Genes Responsible for the Production of Herbivore-induced Volatiles

ABSTRACT

In response to insect herbivory, many plants release a mixture of volatile compounds, which can function as cues to attract the natural enemies of the attacking insects. In this study, we showed that soybean [*Glycine max* (L.) Merr.] plants, after being damaged by two-spotted spider mites, *Tetranychus urticae,* emitted significantly higher levels of volatiles than undamaged plants, with methyl salicylate and (E,E)-α-farnesene being the predominant components. α-farnesene is sesquiterpene and has been found to be herbivore-induced in many plants. Based on their relatedness to known terpene synthse (TPS) genes, two candidate terpene synthase (TPS) genes, *GmTPS1* and *GmTPS2,* for the production of herbivore-induced sesquiterpenes were identified from the sequenced soybean genome. Heterologous expression in a bacterial system indicated that these two genes encode active sesquiterpene synthases and are involved in the production of herbivore-induced sesquiterpenes in soybean. GmTPS1 formed (E,E)-α-farnesene as the major product, while GmTPS2 produced four sesquiterpenes dominated by E-β-caryophyllene. Phylogenetic analysis revealed that GmTPS1 and GmTPS2 fell into the conserved α-farnesene synthase and β-caryophyllene synthase clusters, respectively. Structural analysis indicated that GmTPS1 and GmTPS2 share the highly conserved motifs present in the known α-farnesene synthase and β-caryophyllene synthase from other plant species. Semi-quantitative reverse transcription PCR analysis was further performed to investigate the transcript abundance of these two soybean *TPS* genes in leaves. The result showed that the expression of *GmTPS1* was highly induced by insect damage, which is consistent with the volatile emission pattern. The identification of

TPS genes responsible for the production of herbivore-induced volatiles provides the molecular tools for the functional characterization of herbivore-induced volatiles in soybean.

KEYWORDS: herbivore-induced volatile, terpene, α-farnesene synthase, β-caryophyllene synthase

INTRODUCTION

When many plants are under herbivore attack, they can release elevated levels of volatiles, which have been demonstrated to be involved in plant indirect defense (Turlings and Wäckers, 2004; Yuan et al. 2008). Among these herbivore-induced volatiles, one important class is terpenoids. Terpenoids are built with the common five-carbon isoprene units. Based on the carbon skeletons of terpenes, they can be classified into monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), triterpene (C30), tetraterpenes (C40) and polyterpenes (>C40). For the biosynthesis of terpenes, isopentenyl diphosphate (IPP) is the common precursor of all terpenes. When isomerized, it gives off dimethylallyl pyrophosphate (DMAPP), which either serves as the substrate for hemiterpene biosynthesis or further forms geranyl diphosphate (GPP) by fusing with one IPP unit. GPP is the precursor for monoterpenes. By condensation of one GPP molecule with one IPP molecule, farnesyl diphosphate (FPP) can be formed and function as the precursor for sesquiterpenes. Similarly, the condensation of one FPP molecule with another IPP molecule will generate geranylgeranyl diphosphate (GGPP), which is the precursor for diterpenes. The enzymes that catalyze the biosynthesis of isoprene, monoterpenes, sesquiterpenes and diterpenes are isoprene synthase, monoterpene synthases, sesquiterpene synthases, and diterpene synthases respectively, using the respective substrates (Degenhardt et al. 2009; Chen et al. 2011). All these enzymes belong to a large protein family called terpene synthases (TPSs).

Since terpenoids are the largest group of plant secondary metabolites, the TPS gene family has been broadly studied in many plant species, such as *Arabidopsis thaliana*

(Aubourg et al. 2002), tomato (*Solanum lycopersicum*) (Falara et al. 2011), poplar (*Populus tremula*) (Danner et al. 2011), rice (*Oryza sativa*) (Yuan et al. 2008), sorghum (*Sorghum bicolor*) (Zhuang et al. 2011), and kiwifruit (*Actinidia deliciosa*) (Nieuwenhuizen, 2009).

However, the study of terpenoids in soybean [*Glycine max* (L.) Merr.] is very limited. Only a few studies have been reported. For example, when soybeans were damaged by herbivores, such as aphids and fall armyworm, a range of volatiles were released (Zhu and Park, 2005; Winter and Rostás, 2010). In the prior studies, α-farnesene was found to be the most abundant volatile emitted from the herbivore-damaged soybean.

The fully sequenced genome of soybean (*G. max* cv. Williams 82) (Schmutz et al. 2010) can facilitate the study of terpene biosynthesis in this plant on the genomic basis. Thorough characterization of soybean TPS genes (*GmTPSs*) will not only benefit the understanding of the molecular basis of terpene biosynthesis but also shed light on the ecological and biological roles of these genes. In this study, two TPS candidate genes were identified from the soybean genome. Enzyme assays showed that both of them possessed sesquiterpene activities. One gene was determined to be the α-farnesene synthase gene, and the other gene was the β-caryophyllene synthase gene. In addition, semi-quantitative PCR was performed to examine the expression patterns of these genes between undamaged and herbivore-attacked soybean plants.

MATERIALS AND METHODS

Plants, Insects and Plant Treatment. Soybean (*G. max* cv. Williams 82) seeds were sterilized and germinated according to the method described in Chapter 2. Seedlings were planted at 3 plants per pot, and grown 12 h light/12 h dark at 28 °C /25 °C and light condition was 150-200 µ mol $m^{-2} \cdot s^{-1}$ for 10 days. Then they were transferred to greenhouse for exposure to two-spotted spider mites, *Tetranychus urticae*. When the symptom of spider mite damage was observed, the three-week old soybean plants were used for plant volatile analysis.

Plant Volatile Collection. Volatiles emitted from spider mite-damaged soybean plants and control soybean plants were collected in an open headspace sampling system (Analytical Research Systems), respectively. Three plants grown in a pot with root systems wrapped with aluminum foil were placed in a glass chamber of 10 cm in diameter and 30 cm high, with a removable O-ring snap lid with an air outlet port. Charcoal-purified air entered the chamber at a flow rate of 0.8 L min^{-1} from the top through a Teflon® hose. Volatiles were collected overnight by pumping air from the chamber through a SuperQ volatile collection trap (Analytical Research Systems). After collection, volatiles were eluted with 100 µl of methylene chloride containing 1-octanol (0.003%) as an internal standard for quantification.

Volatile Identification by GC-MS Analysis. Plant volatiles and volatile terpenoids from TPS enzyme assays (see below) were analyzed by a Shimadzu 17A gas

chromatograph coupled to a Shimadzu QP5050A quadrupole mass selective detector. Compounds' separation was performed on a Restek SHR5XLB column (30 m * 0.25 mm internal diameter * 0.25 µm thickness, Shimadzu, Columbia, MD). Helium was used as the carrier gas (flow rate of 5 ml min^{-1}), and a splitless injection (injection temperature 250 ºC) was used. A temperature gradient of 5 ºC/min from 60 ºC (6 min hold) to 300 ºC was applied. Products were identified using the National Institute of Standards and Technology (NIST) mass spectral database and by comparison of retention times and mass spectra with authentic reference compounds if available. Quantification was performed as previously reported (Chen et al. 2009). Representative single-ion peaks of each compound were integrated and compared with the equivalent response of the internal standard (single-ion method).

Database Search and Sequence Analysis. The protein sequence of MdAFS1 from apple (*Malus* x *domestica*) (GenBank accession AAO22848.2) was used as a query, BLAST search against the soybean genome database (http://www.phytozome.net/soybean) was performed in order to obtain the complete open reading frames of soybean putative TPS genes. For identifying significant protein matches, the cutoff value was set as ≤e-4. A multiple sequence alignment of selected functionally characterized TPS and putative soybean TPS was performed using ClustalX program. A neighbor-joining rooted phylogenetic tree was constructed and viewed using TreeView software.

Isolation of Soybean TPS Genes. Soybean leaf materials from herbivore-damaged and control plants were harvested immediately at the end of the volatile collection, flash-frozen with liquid nitrogen and stored at 80 ºC until sample preparation. Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) and DNA contamination was removed with an on-column DNase (Qiagen, Valencia, CA, USA) treatment. Then total RNA was reverse transcribed into first strand cDNA in a 15 µl reaction volume using the First-Strand cDNA Synthesis Kit (GE Healthcare, Piscataway, NJ, USA) as previously described (Chen et al. 2003). The full length cDNA sequence of Glyma17g05500 (GmTPS1) and Glyma12g16990 (GmTPS2) were obtained from public soybean database (http://www.phytozome.net/soybean). The following primers were designed for cloning and semi-quantitative RT-PCR as following: *GmTPS1*: 5'-ATGAATCACTCATACGCGAATCAATC-3' (forward) and 5'-CTATCTAAGGGGTTCAACAACCAGTG -3' (reverse); *GmTPS2*: 5'-ATGACTAACGTATCCTCGTCACTTC-3' (forward) and 5'-CTAGCTTATACGGGCAGACATC-3' (reverse). The PCR program used to amplify the target genes was performed as following: 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 56 °C for 30 s and 72 °C for 1 min 50 s, and a final extension at 72 °C for 10 min. PCR products were cloned into vector pEXP5/CT-TOPO. By sequencing, the PCR products of *GmTPS1* and *GmTPS2* were the same as that predicted in the soybean database.

***Escherichia coli*-expression of Soybean TPSs.** To study the biochemical function of

GmTPS1 and GmTPS2, the above mentioned protein expression vector pEXP5/CT-TOPO harboring *GmTPS1* and *GmTPS2* respectively, were transformed into the *E. coli* strain BL21 (DE3) CodonPlus (Stratagene, La Jolla, CA). 50 ml of liquid cultures of the bacteria harboring the expression constructs were grown at 37 °C to an OD600 of 0.6. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 500 μM, and the cultures were incubated for 20 h at 18 °C. Then the cells were collected by centrifugation and disrupted by a 4 x 30 sec sonication treatment in chilled extraction buffer (50 mM Mopso, pH 7.0, with 5 mM $MgCl_2$, 5 mM sodium ascorbate, 0.5 mM PMSF, 5 mM dithiothreitol and 10% (v/v) glycerol). The cell fragments were removed by centrifugation at 14,000 g and the supernatant was desalted into assay buffer (10 mM Mopso, pH 7.0, 1 mM dithiothreitol, 10% (v/v) glycerol) by passage through an Econopac 10DG column (Bio-Rad, Hercules, CA, USA).

GmTPS Activity Assays. To determine the catalytic activity of the putative soybean TPSs, enzyme assays for recombinant proteins of GmTPS1 and GmTPS2 were performed at 30 °C for overnight, using 50 μl of the crude enzymes and 50 μl assay buffer with 10 μM substrate (GPP and FPP, respectively), 10 mM $MgCl_2$, 0.05 mM $MnCl_2$, 0.2 mM $NaWO_4$ and 0.1 mM NaF in a Teflon-sealed, screw-capped 1 ml GC glass vial. A solid phase microextraction (SPME) fiber consisting of 100 μm polydimethylsiloxane (Supelco, Bellefonte, PA, USA) was placed in the headspace of the vial for 10 min for collecting the volatile. For analysis of the absorbed reaction products, the SPME fiber was inserted directly into the injector of the gas

chromatograph. GC-MS analysis and product identification were performed as described above.

Transcript Abundance Analysis of GmTPSs Using Semi-quantitative Reverse Transcription PCR. In performing semi-quantitative RT-PCR analysis for *GmTPS1 and GmTPS2*, soybean ubiqutin3 gene (*GmUBI-3,* GenBank accession D28123) was used as an internal control. Initially, PCR analysis was performed with ubiquitin3 specific primers using 0.1 µl, 0.2 µl, 0.5 µl and 1.0 µl cDNA. The program used to amplify *GmUBI-3* was as follows: 94 °C for 2 min followed by 30 cycles at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 45 s, with a final extension at 72 °C for 10 min. Amplified products were separated on 1.0% agarose gel. Gels were stained with ethidium bromide, visualized under UV-light and quantified using the Bio-Rad Quantity One software (Bio-Rad, Hercules, CA). Analysis showed that the amount of amplified products with the *GmUBI-3* specific primers increased linearly with increasing amounts of template cDNA. Therefore, 0.2 µl cDNA products gained from control and herbivore damaged soybean leaves respectively, were used for the optimal template concentration for PCR analysis with the *GmTPS1* and *GmTPS2* specific primers. The program used to amplify the *GmTPS1* and *GmTPS2* fragment was as follows: 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 56 °C for 30 s and 72 °C for 1 min 50 s, and a final extension at 72 °C for 10 min. Since methyl salicylate (MeSA) was also observed in the herbivore-induced volatile of soybean leaves, gene expression of *GmSAMT1* (described in Chapter 2) was compared between control and herbivore-damaged leaves as well.

The PCR program and primers for *GmSAMT1* were used for gene cloning condition as mentioned in Chapter 2. All PCRs were replicated three times with the first-strand cDNAs made from two independent RNA preparations.

RESULTS

Herbivore-induced Volatiles from Soybean Plants. Headspace collection coupled with GC-MS analysis showed that three-week old soybean plants grown under the normal conditions in the growth chamber emitted only a small amount of α-farnesene and a trace amount of MeSA (**Figure 4.1**). However, for one-week soybean seedling grown in the growth chamber, α-farnesene was not detected under the same volatile collection condition (data not shown). When the three-week soybean plants were severely damaged by spider mites, the volatile profile was significantly different from that of control plants. A total of 7 volatiles were detected only from spider mite-damaged soybean plants (**Figure 4.1**). Similar to the volatiles emitted from other plant species upon herbivory, spider mite-induced volatiles from soybean can be categorized into three major groups: terpenes, shikimic acid derived metabolites and fatty acid-derived metabolites. The terpene group contained two sesqueterpene: α-farnesene and E-β-caryophyllene, and one homoterpne 4,8,12-trimethyl-3,7,11-tridecatrienenitrile (TMTT). α-farnesene was the most abundant sesquiterpene. In addition, (E,E)-farnesol was also detected in trace amounts. There were two compounds derived from the shikimate pathway: methyl salicylate and indole. (Z)-3-hexenyl acetate is a representative of fatty acid-derived volatiles. The major

soybean volatile components of herbivore-induced volatile include (Z)-3-hexenyl acetate (about 17% of the total volatiles), α-farnesene (about 30% of the total volatiles), MeSA (about 45% of the total volatiles), and TMTT (about 6% of the total volatiles).

Identification of TPSs. To investigate the molecular basis of herbivore-induced volatiles of soybean, we sought to identify genes responsible for the formation of the major terpenes of the volatile bouquet of herbivore-damaged soybean plants. The gene, *Glyma02g06070* (*GmSAMT1*), responsible for producing MeSA has been studied in Chapter 2. Considering that the predominant terpene compound of herbivore-induced volatiles from soybean was α-farnesene, the protein sequence of characterized α-farnesene synthase gene (*MdAFS1*) from apple (Pechous and Whitaker, 2004) was used as a query to BLAST search against the soybean genome database (http://www.phytozome.net/soybean). A total of putative TPS genes were identified. One soybean gene with complete open reading-frame, Glyma17g05500 designated as GmTPS1, showed 52.7% amino acid sequence identity to MdAFS1. Another gene with a complete open reading-frame, Glyma12g16990, designated as GmTPS2, showed 34.2% amino acid sequence identity to MdAFS1. Interestingly, *GmTPS2* was predicted to be the β-caryophyllene/α-humulene synthase with 38.4% amino acid sequence identity to one Arabidopsis TPS, AtTPS21 (GenBank accession AA085539).

Phylogenetic Analysis of GmTPSs with Related Proteins. Thirty soybean putative soybean TPS genes that are significantly homologous to known plant terpene synthase

were identified from the soybean genome. These 30 genes are distributed in 11 chromosomes. Since many terpene synthase genes have been characterized in other plant species, the evolutionary relatedness of GmTPS1 and GmTPS2 to some known TPSs from other plants was studied by phylogenetic analysis (**Figure 4.2**). Among 30 putative soybean terpene synthases, GmTPS1 fell in the a cluster with previously described α-farnesene synthases, such as MdAFS1, PcTPS, and PtTPS2 and AtTPS3, from apple (*Malus* × *domestica*), pear (*Pyrus communis*), poplar (*Populus trichocarpa*) and *A. thaliana* (Aubourg et al. 2002; Gapper et al. 2006; Danner et al. 2011; Huang et al. 2010). GmTPS2 are most closely related to known β-caryophyllene synthases, including AtTPS21, ZmTPS10 and ZmTPS23 from *A. thaliana* and maize (*Zea mays*) (Chen et al. 2003; Köllner et al. 2008).

Motif Analysis of GmTPSs. GmTPS1 and GmTPS2 contained several highly conserved domains of known terpene synthases, including the aspartate-rich region DDxxD motif, which is involved in metal cofactor binding, and RxR motif, which is implicated in the complexation of the diphosphate group after ionization of the substrate (Starks et al. 1997). Both NDxxTxxxE or DDxxSxxxE motif and H-α1 loop, which are proposed to be the magnesium and potassium binding regions respectively (Green et al. 2009), also exist in the protein sequences of GmTPS1 and GmTPS2. Interestingly, GmTPS1 contained the $RR(x)_8W$ motif, which was also found in the other α-farnesene synthase MdAFS1 and AtTPS3, whereas GmTPS2 and the known characterized β-caryophyllene synthases, such as ZmTPS23 and AtTPS21, do not have this motif

(**Figure 4.3**).

Cloning and Biochemical Characterization of GmTPSs. To further confirm the terpene synthases' enzyme function of the above two putative soybean terpene synthase genes, *E. coli*-expressed recombinant GmTPS1 and GmTPS2 were then assayed for TPS activity with the potential substrates GPP and FPP. Reaction products were collected from the headspace using SPME and analyzed by GC-MS. Both GmTPS1 and GmTPS2 showed the activity with the substrate FPP, not GPP. GmTPS1 exhibited a specific sesquiterpene synthase activity with (E,E)-α-farnesene as the only product (**Figure 4.4)**. In contrast, GmTPS2 showed a relative broad product specificity with four detectable sesquiterpenes dominated by E-β-caryophyllene, and low amount of α-Cubebene, (Z,E)-α-farnesene and Germacrene D (**Figure 4.5**).

Transcript Abundance of GmTPSs. To further determine whether the production of herbivore-induced volatiles is regulated at the transcription level, the expression of both *GmTPS1* and *GmTPS2* were compared between the control and herbivore-damaged soybean plants by semi-quantitative reverse transcription (RT)-PCR analysis. Meantime, to understand the role of MeSA emission in the herbivore-induced defense, the expression of *GmSAMT1* was also studied. The results showed the expression *GmTPS1* and *GmSAMT1* were highly induced in herbivore-damaged soybean, whereas very low level of expression of these two genes was observed in the control plant. However, there was no obvious change on the expression of *GmTPS2*, which showed similar

levels of expression in control and herbivore-damaged soybean (**Figure 4.6**).

DISCUSSION

Soybean was observed to emit a mixture of volatiles when severely attacked by two-spotted spider mites (**Figure 4.1**). The major components of spider mite-induced volatiles were (Z)-3-hexenyl acetate, α-farnesene, MeSA, and TMTT, all of which have been studied in other plant-herbivore systems. In these plant species, insect-induced volatiles were shown to act as a signal to attract the natural enemies of the attacking insects (De Boer et al. 2004; Himanen et al. 2005; Loughrin et al. 1994). The similar pattern of volatile profiling suggests that these volatiles may play the same role in soybean. However, some plant species were shown to emit a broader spectrum of volatiles upon herbivore feeding (Yuan et al. 2008; Köllner et al. 2008). This is not surprising because studies have demonstrated that insect induced-volatiles exhibit the diverse patterns among different plant species. Meantime, the different growth conditions and different insects may also be the reason leading to the low numbers of volatiles emitted in herbivore-damaged soybean. It is also interesting to note that there was no monoterpene detected from herbivore-damaged soybean plants.

Among the herbivore-induced volatiles, terpenes constitute the main group. Although TPS genes have been reported for many plant species, the study about the legume TPS is very limited. In order to fill this important gap, soybean TPSs were studied in this chapter. Here, two novel soybean TPS genes, *GmTPS1* (*Glyma17g05500*) and *GmTPS2* (*Glyma12g16990*), which were related to the formation of α-farnesene

and β-caryophyllene (**Figure 4.4** and **Figure 4.5**), respectively, were first identified based on their sequenced relatedness to known plant TPSs.

The biochemical assays of the recombinant *GmTPS1* and *GmSAMT1* indicated that they catalyzed the formation of α-farnesene and MeSA. Interestingly the transcription patterns of *GmTPS1* and *GmSAMT1* (**Figure 4.6**) were consistent with the soybean volatile patterns (**Figure 4.1**). This suggests that soybean may respond to herbivore's damage by emitting α-farnesene and MeSA through increasing the expression of *GmTPS1* and *GmSAMT1* genes. In contrast, the emission of E-β-caryophyllene was not strongly induced by spider mite damage. The gene *GmTPS2*, coding β-caryophyllene synthase, also showed no obvious changes in transcript abundance (**Figure 4.6**).

Based on the phylogenetic analysis of soybean TPS genes in this study (**Figure 4.2**) and prior studies from many plant species (Danner et al. 2011), some α-farnesene synthase genes from different angiosperm species are evolutionarily closely related. α-farnesene synthase has also been found in gymnosperms (Danner et al. 2011) and this enzyme catalyzed the formation of non-cyclic product of (E,E)-α-farnesene, by relatively simple reaction mechanism (Sharkey et al. 2005). The gymnosperm α-farnesene synthase is only distantly related to angiosperm α-farnesene synthase, suggesting independent evolution of this enzyme in angiosperms and gymnosperms.

Farnesene is a common compound of herbivore-induced plant volatile blends from some plants such as bean, pear, apple and so on (Du et al. 1998; Boeve et al. 1996; Scutareanu et al. 2003). The laboratory results showed that farnesene was attractive to parasitic wasps including *Aphidius ervi*, *Coleomegilla maculate*, and *Chrysoperla*

carnea (Du et al. 1998; Zhu et al. 1999). James (2005) also reported the parasitic mymarid wasp, *Anagrus daanei*, was attracted to farnesene in the field study too. In this study, spider mite-damaged soybean leaves were analyzed and shown to release (E,E)-α-farnesene as the predominant compound. Soybean α-farnesene synthase gene shared highly conserved motifs (**Figure 4.3**) and was closely related to the other α-farnesene synthase genes from angiosperms. Therefore soybean may share a similar defense mechanism of producing α-farnesene to attract predator insects to activate the indirect defense against herbivores as found in other plant species.

REFERENCES

Aubourg SA, Lecharny AL, Bohlmann JB (2002) Genomic analysis of the terpenoid synthase (*AtTPS*) gene family of *Arabidopsis thaliana*. Mol. Genet. Genomics 267: 730-745

Boeve JL, Lengwiler U, Dorn S, Turlings TCJ, Tollsten L (1996) Volatiles emitted by apple fruitlets infested by larvae of the European apple sawfly. Phytochem. Lett. 42: 373-381

Chen F, Al-Ahmad H, Joyce B, Zhao N, Köllner TG, Degenhardt J, Stewart Jr CN (2009) Within-plant distribution and emission of sesquiterpenes from *Copaifera officinalis*. Plant Physiol. Bioch. 47: 1017-1023

Chen F, Tholl D, Bohlmann J, Pichersky E (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J. 66: 212-229

Chen F, Tholl D, D'Auria JC, Farooq A, Pichersky E, Gershenzon J (2003) Biosynthesis and emission of terpenoid volatiles from Arabidopsis flowers. Plant Cell 15: 481-494

Colazza S, McElfresh JS, Millar J (2004) Identification of volatile synomones, induced by *Nezara viridula* feeding and oviposition on bean spp., that attract the egg parasitoid *Trissolcus basalis*. J. Chem. Ecol. 30: 945-964

Danner H, Boeckler GA, Irmisch S, Yuan JS, Chen F, Gershenzon J, Unsicker SB, Kollner TG (2011) Four terpene synthases produce major compounds of the gypsy moth feeding-induced volatile blend of *Populus trichocarpa*. Phytochemistry 72: 897-908

De Boer JG, Posthumus MA, Dicke M (2004) Identification of volatiles that are used in discrimination between plants infested with prey or nonprey herbivores by a predatory mite. J. Chem. Ecol. 30: 2215-2230

Degenhardt J, Köllner T, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70: 1621-1637

Du Y, Poppy GM, Powell W, Pickett JA, Wadhams LJ, Woodcock CM (1998) Identification of semiochemicals released during aphid feeding that attract parasitoid *Aphidius ervi*. J. Chem. Ecol. 24: 1355-1368

Falara V, Akhtar T, Nguyen TTH, Spyropoulou EA, Bleeker PM, Schauvinhold I, Matsuba Y, Bonini ME, Schilmiller AL, Last RL, Schuurink RC, Pichersky E (2011) The tomato (*Solanum lycopersicum*) terpene synthase gene family. Plant Physiol. 157:

770-789

Gapper NE, Bai J, Whitaker BD (2006) Inhibition of ethylene-induced α-farnesene synthase gene PcAFS1 expression in ‘d’Anjou’ pears with 1-MCP reduces synthesis and oxidation of α-farnesene and delays development of superficial scald. Postharvest Biol. Tec. 41: 225-233

Green S, Squire CJ, Nieuwenhuizen NJ, Baker EN, Laing W (2009) Defining the potassium binding region in an apple terpene synthase. J. Biol. Chem. 284: 8661-8669

Himanen S, Vuorinen T, Tuovinen T, Holopainen JK (2005) Effects of cyclamen mite (*Phytonemus pallidus*) and leaf beetle (*Galerucella tenella*) damage on volatile emission from strawberry (*Fragaria* × *ananassa Duch*.) plants and orientation of predatory mites (*Neoseiulus cucumeris*, *N. californicus*, and *Euseius finlandicus*). J. Agr. Food Chem. 53: 8624-8630

James DG (2005) Further field evaluation of synthetic herbivore-induced plan volatiles as attractants for beneficial insects. J. Chem. Ecol. 31: 481-495

Köllner TG, Held M, Lenk C, Hiltpold I, Turlings TCJ, Gershenzon J, Degenhardt J (2008) A maize (E)-β-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most american maize varieties. Plant Cell 20: 482-494

Loughrin JH, Manukian A, Heath RR, Turlings TC, Tumlinson JH (1994) Diurnal cycle of emission of induced volatile terpenoids by herbivore-injured cotton plant. Proc. Natl. Acad. Sci. USA 91: 11836-11840

Nieuwenhuizen NJ, Wang MY, Matich AJ, Green SA, Chen X, Yauk Y-K, Beuning LL, Nagegowda DA, Dudareva N, Atkinson RG (2009) Two terpene synthases are responsible for the major sesquiterpenes emitted from the flowers of kiwifruit (*Actinidia deliciosa*). J. Exp. Bot. 60: 3203-3219

Pechous S, Whitaker B (2004) Cloning and functional expression of an (*E*,*E*)-a-farnesene synthase cDNA from peel tissue of apple fruit. Planta 219: 84-94

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X-C, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178-183

Scutareanu P, Bruin J, Posthumus MA, Drukker B (2003) Constitutive and herbivore-induced volatiles in pear, alder and hawthorn trees. Chemoecology 13: 63-74

Starks CM, Back K, Chappell J, Noel JP (1997) Structural Basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. Science 277: 1815-1820

Turlings T, Wäckers F (2004) Recruitment of predators and parasitoids by herbivore-injured plants. In: Cardé R MJ (ed) In Advances in Insect Chemical Ecology. Cambridge University Press, Cambridge, pp 21-75

Winter T, Rostás M (2010) Nitrogen deficiency affects bottom-up cascade without disrupting indirect plant defense. J. Chem. Ecol. 36: 642-651

Yuan JS, Köllner TG, Wiggins G, Grant J, Degenhardt J, Chen F (2008) Molecular and genomic basis of volatile-mediated indirect defense against insects in rice. Plant J. 55: 491-503

Zhu J, Cossé AA, Obrycki JJ, Boo KS, Baker TC (1999) Olfactory reactions of the twelve-spotted lady beetle, *Coleomegilla maculata* and the green lacewing, *Chrysoperla carnea* to semiochemicals released from their prey and host plant: electroantennogram and behavioral responses. J. Chem. Ecol. 25: 1163-1177

Zhu JW, Park KC (2005) Methyl salicylate, a soybean aphid-induced plant volatile attractive to the predator *Coccinella septempunctata*. J. Chem. Ecol. 31: 1733-1746

Zhuang X, Köllner TG, Zhao N, Li G, Jiang Y, Zhu L, Ma J, Degenhardt J, Chen F (2011) Dynamic evolution of herbivore-induced sesquiterpene biosynthesis in sorghum and related grass crops. Plant J. doi: 10.1111/j.1365-1313X.2011.04771.x.

Appendix

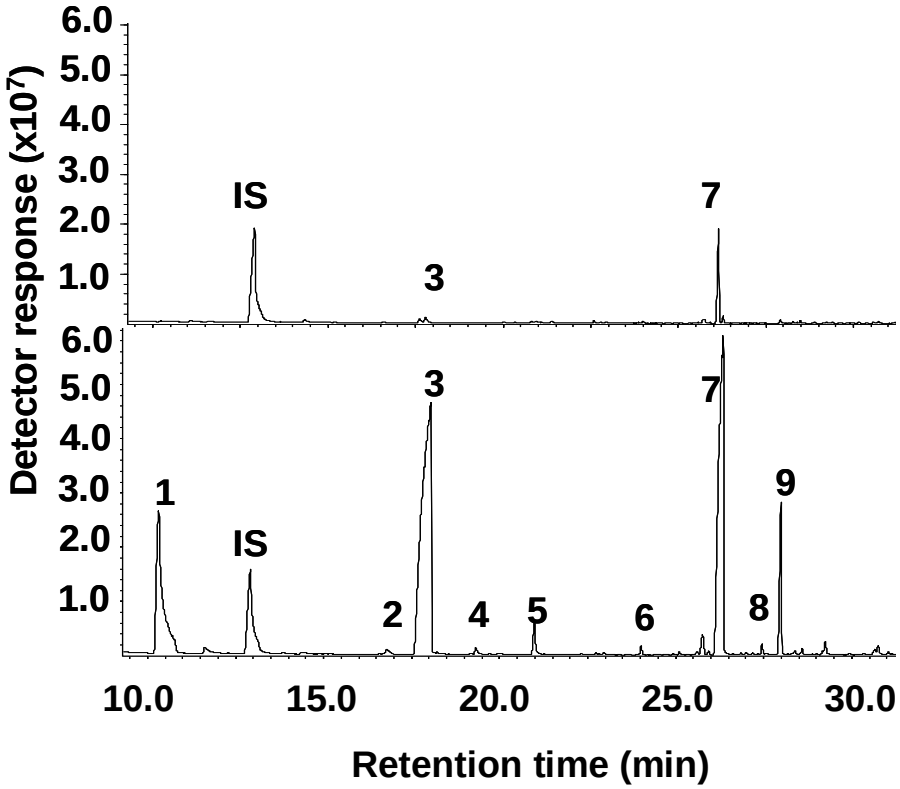

Figure 4.1. Volatiles emitted from spider mite-damaged Williams 82 soybean plants and untreated control plants. The upper panel shows a GC chromatogram of the volatiles from control plants and the lower panel shows a GC chromatogram from spider mite-treated plants. IS represents the internal standard. Peaks 1-9 are compounds: 1: 3-Hexen-1-ol, acetate; 2: Butanoic acid, 3-hexenyl ester; 3: Methyl salicylate; 4: 1,2,3-Propanetriol, monoacetate; 5: Indole; 6: E-β-Caryophyllene; 7: (E,E)-α-Farnesene; 8: (E,E)-Farnesol; 9: 4,8,12-trimethyl-3,7,11-Tridecatrienenitrile (TMTT).

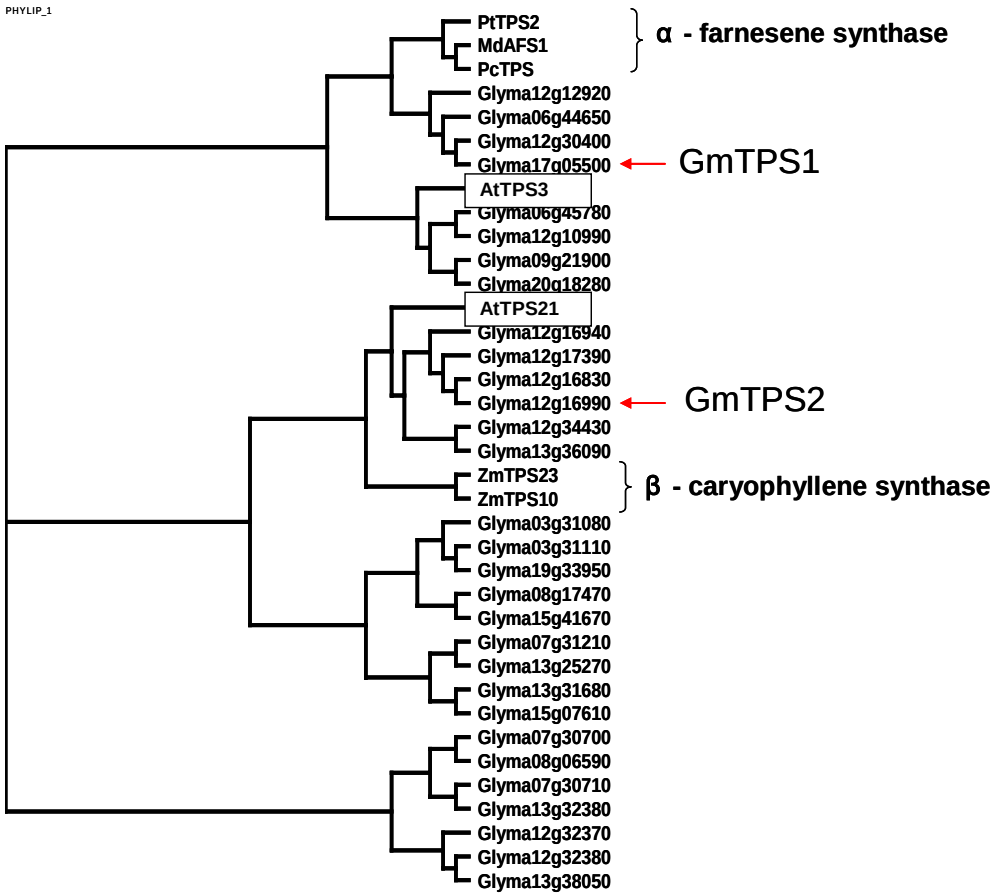

Figure 4.2. Phylogenetic tree of selected terpene synthase proteins from other plant species together with 30 putative soybean terpene synthase genes. The rooted phylogenetic tree was accomplished with the ClustalX algorithm. All sequences that start with 'Glyma' are soybean terpene synthase genes. The two terpene synthases studied in this chapter, GmTPS1 and GmTPS2, are indicated. MdAFS1, PtTPS2 (GenBank accession AEI52902), PcTPS (GenBank accession ABC25002) and AtTPS3 (At4g16740) are α-farnesene synthases. ZmTPS10 (GenBank accession AAX99146), ZmTPS23 (GenBank accession ABY79207), and AtTPS21 (At5g23960) are β-caryophyllene synthases.

Figure 4.3. Comparison of the deduced amino acid sequence of GmTPS1 and GmTPS2 with some known terpene synthases. The functionally characterized terpene synthases include AtTPS3, AtTPS21, ZmTPS23 and MdAFS1. Amino acids identical in all four proteins are marked by black boxes and amino acids with similar side chains are marked by gray boxes. The highly conserved metal cofactor binding regions are labeled DDxxD, NDxxTxxxE/DDxxSxxxE, respectively. RR$(x)_8$W motif which was shown in GmTPS1 not in GmTPS2. H α1-loop was also shown in both GmTPS1 and GmTPS2.

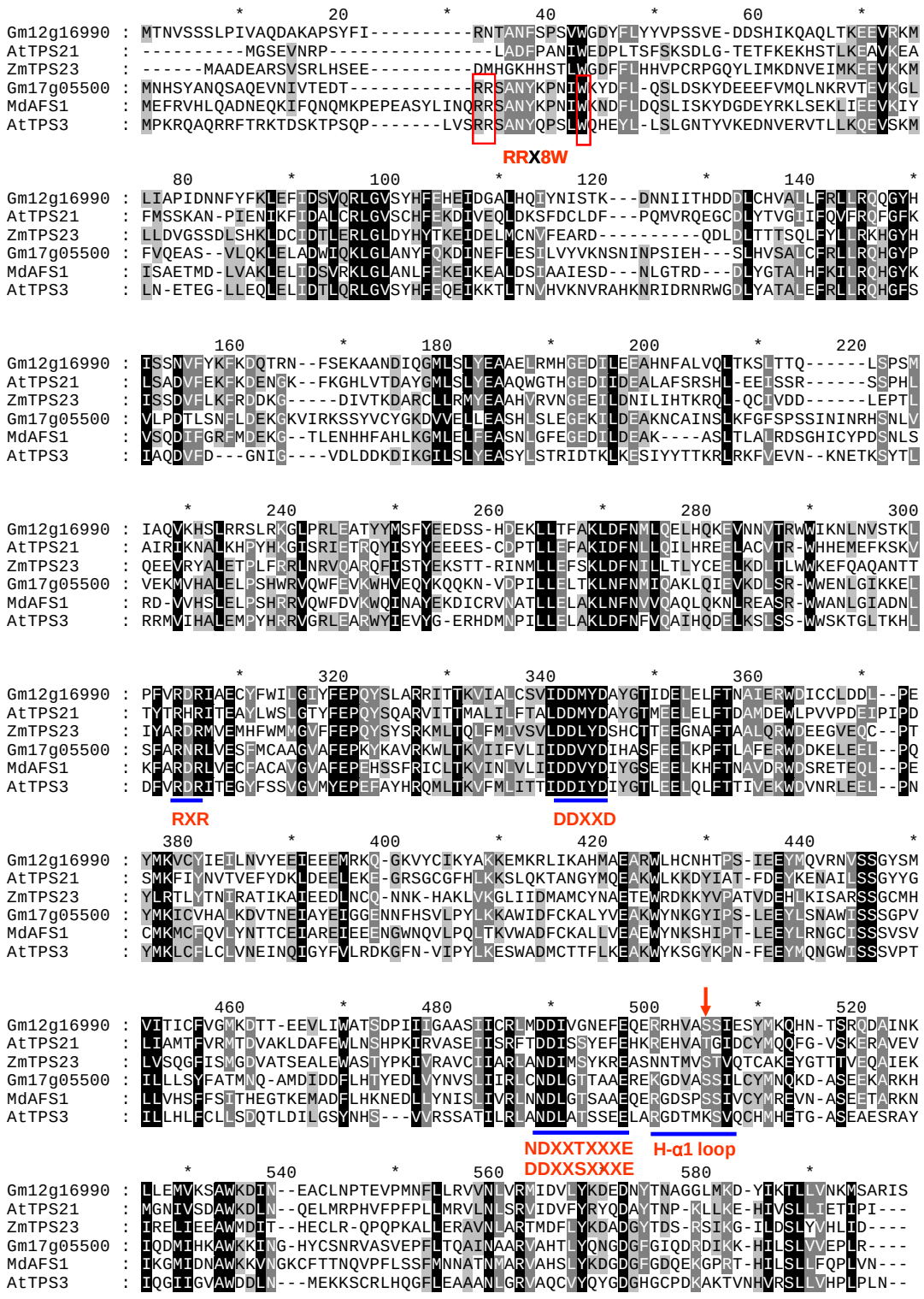

Figure 4.3. Comparison of the deduced amino acid sequence of GmTPS1 and GmTPS2 with some known terpene synthases.

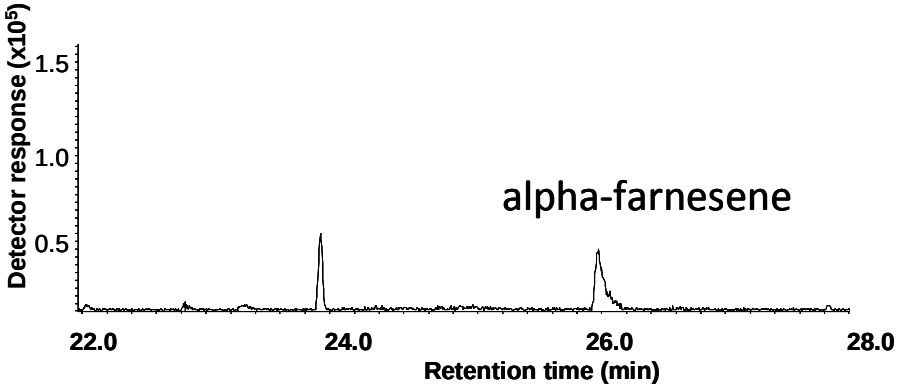

Figure 4.4. GC–MS total ion chromatograms of reaction products from a terpene synthase assay of *E. coli* expressed GmTPS1. The peak is (E,E)-α-farnesene with retention time (about 26 min), and other peaks are non-terpene bacterial contaminants that were present in all TPS assay extracts (data not shown).

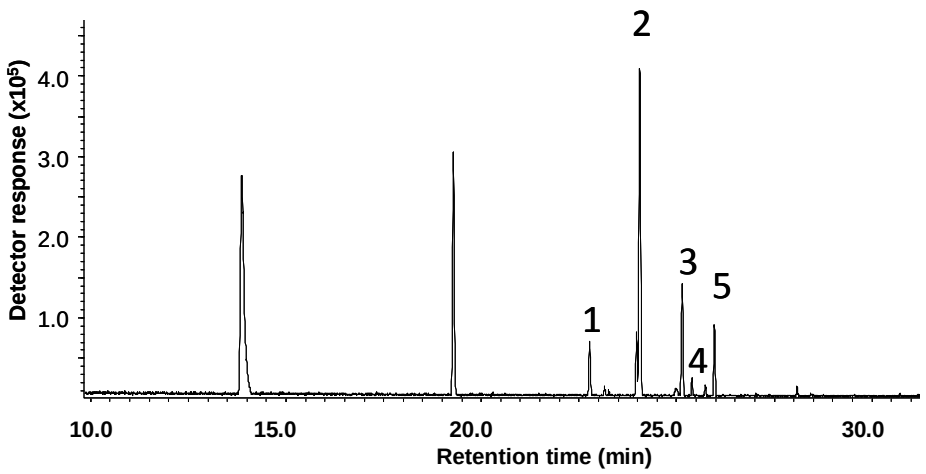

Figure 4.5. GC–MS total ion chromatograms of reaction products from a terpene synthase assay of *E. coli* expressed GmTPS2. Five peaks are identified as following 1: α-Cubebene; 2: E-β-Caryophyllene; 3: α-Caryophyllene; 4: (Z,E)-α-Farnesene; 5: Germacrene D.

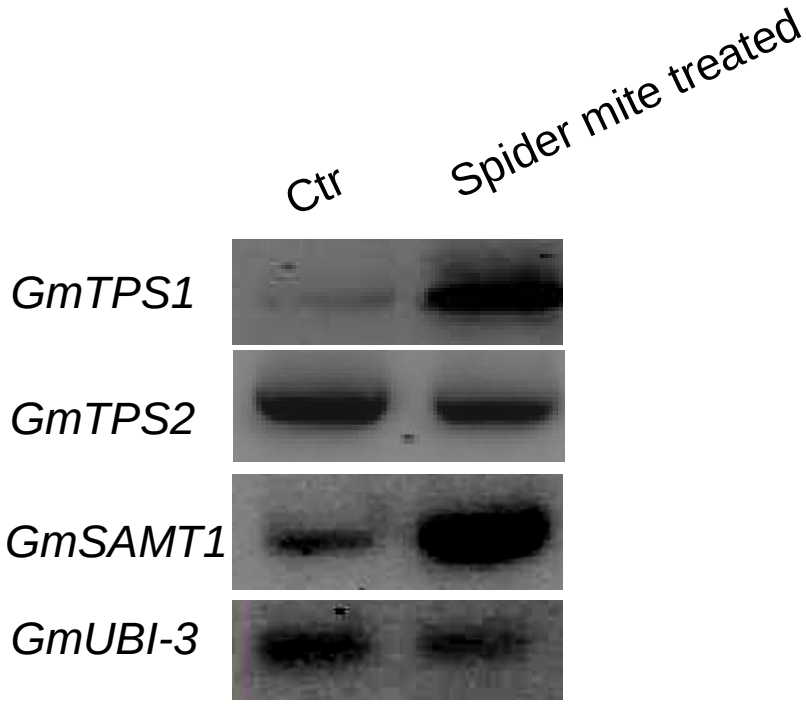

Figure 4.6. Expression analysis of *GmTPS* genes and *GmSAMT1* in spider mite-damaged (herbivory) and untreated (control) soybean leaves using semi-quantitative RT-PCR. The PCR products for soybean ubiquitin3 (*GmUBI-3*) were used to judge equality of concentration of cDNA templates in different samples.

Chapter V. Molecular Cloning and Biochemical Characterization of an Endo-β-Mannanase Gene from Soybean for Soybean Meal Improvement

Adapted from:

Jingyu Lin, Vincent R. Pantalone, Guanglin Li, Feng Chen (2011) Molecular cloning and biochemical characterization of an endo-β-mannanase gene from soybean for soybean meal improvement. J. Agric. Food Chem. 59:4622-4628

ABSTRACT

Soybean meal is the most commonly used protein source in animal feeds. Among the undesirable attributes of soybean meal is the high level of β-mannan, which was determined to be detrimental to the growth performance of animals. β-mannan is a type of hemicellulose in the plant cell wall and can be hydrolyzed by endo-β-mannanase. The goal of this study is to isolate and characterize an endo-β-mannanase gene from soybean that can be used for genetic improvement of soybean meal. From the sequenced soybean genome, 21 putative endo-β-mannanase genes were identified. Based on their relatedness to known functional plant endo-β-mannanases, four soybean endo-β-mannanase genes (*GmMAN1* to *GmMAN4*) were chosen for experimental analysis. *GmMAN1* and *GmMAN4* showed expression in the soybean tissue examined, and their cDNAs without the sequences for signal peptide were cloned and expressed in *E. coli* to produce recombinant enzymes. Only GmMAN1 showed endo-β-mannanase hydrolase activity. Further gene expression analysis showed that *GmMAN1* is specifically expressed in cotyledons of seedlings, suggesting a role of *GmMAN1* in degrading mannan-rich food reserves during soybean seedling establishment. Purified recombinant GmMAN1 exhibited an apparent Km value of 34.9 mg/mL. The catalytic efficiency (kcat/Km) of GmMAN1 was determined to be 0.7 ml/mg · s. GmMAN1 was also shown to be active in hydrolyzing the β-mannan-rich cell wall of soybean seeds.

KEYWORDS: endo-β-mannanase, soybean meal, kinetics, cotyledon

INTRODUCTION

Soybean meal, the product remaining after removal of most of the oil from soybean seeds, is a major source of protein in livestock and poultry feeds in most countries (Hsiao et al. 2006). The quality of soybean meal is determined by a number of parameters, including protein content, amino acid composition, carbohydrate composition and digestibility, among which, a high content of β-mannan was determined to be an undesirable character of soybean meal (Duncan et al. 2002). β-mannan is one type of hemicellulosic polysaccharide found in the plant cell wall. It is composed of repeating units of mannose, with galactose or glucose, or both, often found attached to the β-mannan backbone (Ebringerova et al. 2005). The β-mannan content of soybean meal ranges from 1.0% to 1.5% for dehulled samples and from 1.3% to 2.1% for nondehulled samples (Hsiao et al. 2006). The relatively high content of β-mannan in animal feeds appears to have a deleterious effect on animal performance. β-mannan was shown to stimulate the innate immune system of animals, which causes an increase in proliferation of macrophages and monocytes, as well as increased cytokine production (van Kooyk and Geijtenbeek 2003). Although the stimulation of the immune system may provide benefits under certain conditions (Staykov et al. 2007), these physiological changes often lead to an increased severity of disease symptoms and a decrease in the efficiency of nutrient utilization (Anderson and Warnick 1964).

As a type of polysaccharide, β-mannan can be hydrolyzed by the action of a number of hydrolytic enzymes, with endo-β-mannanase (EC 3.2.1.78) being the most important. Endo-β-mannanase catalyzes the random hydrolysis of the β-mannosidic linkage in the

mannan backbone. A number of studies have shown that the addition of exogenous endo-β-mannanase derived from microorganisms to diets containing soybean meal can improve the performance of poultry and swine (Pettey et al. 2002; Wu et al. 2005). Previous studies showed that the quality of soybean meal, such as the content of phytate, can be genetically improved (Yang et al. 2011). These results suggest that the desirable soybean meal with a low content of β-mannan may be produced by overexpression of an endo-β-mannanase gene in soybean seeds through genetic engineering.

A successful genetic engineering program of soybean meal improvement for reducing β-mannan requires an endo-β-mannanase gene. No endo-β-mannanase gene has been reported from soybean, whereas genes encoding endo-β-mannanase have been isolated from a wide range of organisms including plants, animals and microorganisms (Yuan et al. 2007). For plants, endo-β-mannanase genes have been isolated from tomato (*Solanum lycopersicum*), coffee (*Coffea arabica*) and lettuce (*Lactuca sativa*) (Bewley et al. 1997; Marraccini et al. 2001; Wang et al. 2004). Sequence comparison and phylogenetic analysis imply that plant endo-β-mannanase genes have a common evolutionary origin (Yuan et al. 2007). The biological function of plant endo-β-mannanase genes has been best characterized for their involvement in seed biology. For example, *LeMAN1*, the first endo-β-mannanase gene to be isolated from a plant, plays a role in post-germination reserve mobilization in tomato seeds (Bewley et al. 1997). In contrast, a germination-specific endo-β-mannanase gene from tomato, *LeMAN2*, was suggested to play a role in regulating germination by modulating the process of endosperm weakening (Nonogaki et al. 2000). In addition,

endo-β-mannanase genes have been shown to be involved in fruit ripening (Carrington et al. 2002) and pollen development (Filichkin et al. 2004). While most of the known endo-β-mannanases were shown to possess hydrolase activity, a recent study suggested that some endo-β-mannanases may have transglycosylase activity (Schröder et al. 2009).

Although any endo-β-mannanase gene can be used as a molecular tool for the genetic improvement of soybean meal quality, a gene from soybean will have an advantage for the creation of cisgenic plants, which are better perceived by the public (Schouten et al. 2006). The genome of soybean has recently been fully sequenced (Schmutz et al. 2010), which provides a unique opportunity for the identification and isolation of endo-β-mannanase genes from soybean. In this study, an endo-β-mannanase gene (*GmMAN1*) that showed hydrolase activity with specific expression in cotyledons of germinated soybean seeds was selected from a family of 21 putative endo-β-mannanase genes and subjected to detailed biochemical characterization. Recombinant GmMAN1 is active in hydrolyzing the β-mannan-rich cell wall of soybean seeds, suggesting that it can be used as a candidate gene for genetic improvement of soybean meal quality.

MATERIALS AND METHODS

Plant Material and Chemicals. ‘Williams 82’ soybeans (Bernard and Cremeens 1988), the cultivar utilized for whole genome sequencing (Schmutz et al. 2010), were used for gene cloning and expression analysis of soybean endo-β-mannanase genes. Germinating seed samples were collected from soybean seeds imbibed in water on germination paper

at 22 °C at different time points. In a growth chamber, soybean seeds were planted in soil and seedlings were grown with a photoperiod set for 18 h at 25 °C in the light and 6 h at 23 °C in the dark. Soybean tissues for gene expression analysis, including root, epicotyls, hypocotyls, cotyledons, first foliage leaves and leaf buds, were harvested from 14-day old seedlings. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

Database Search and Sequence Analysis. The soybean genome sequence (Schmutz et al. 2010) was analyzed to identify endo-β-mannanase genes using the BLASTP program (Altschul et al. 1997) with the protein sequence of the tomato endo-β-mannanase protein LeMAN1 (GenBank accession AAB87859) as a query. The E-value of ≤e-5 was set as the cutoff value for identifying significant protein matches. In phylogeny reconstruction, a multiple sequence alignment of selected endo-β-mannanase proteins was performed using ClustalW2 (Larkin et al. 2007) with default parameters. The neighbor-joining phylogenetic tree was constructed using the MEGA software (ver.4.0) (Tamura et al. 2007).

Gene Expression Analysis Using Semi-quantitative RT-PCR. Total RNA was extracted from germinating seeds and other tissues using an RNeasy Plant Mini Kit (Qiagen, Valencia, CA) with DNA contamination removed using an on-column DNase treatment (Qiagen). After purification, total RNA (1.5 μg) was reverse-transcribed into first-strand cDNA in a 15 μl reaction volume using the first-strand cDNA Synthesis Kit

(Amersham Biosciences, Piscataway, NJ) as previously described (Zhao et al. 2010; Zhao et al. 2009). With the first strand cDNA as the template, RT-PCR was conducted using the gene specific primers for analyzing the expression of *GmMAN1* to *GmMAN4*. The mixed cDNAs (1.0 µl) from various tissues was used as the template for the PCR reactions. In performing semi-quantitative RT-PCR analysis for *GmMAN1*, soybean ubiqutin3 gene (*GmUBI-3*, GenBank accession D28123) was used as an internal control. The primers for semi-quantitative RT-PCR were designed (**Table 5.1**).

Initially, PCR analysis was performed with ubiquitin3 specific primers using 0.1 µl, 0.2 µl, 0.5 µl and 1.0 µl cDNA. The program used to amplify ubiquitin3 was as follows: 94 °C for 2 min followed by 30 cycles at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 45 s, with a final extension at 72 °C for 10 min. Amplified products were separated on 1.0% agarose gel. Gels were stained with ethidium bromide, visualized under UV-light and quantified using the Bio-Rad Quantity One software (Bio-Rad, Hercules, CA). Analysis showed that the amount of amplified products with the ubiquitin3 specific primers increased linearly with increasing amounts of template cDNA. Therefore, 0.2 µl cDNA was chosen as the optimal template concentration for PCR analysis with the *GmMAN1*-specific primers. The program used to amplify the *GmMAN1* fragment was as follows: 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 1 min 30 s, and a final extension at 72 °C for 10 min. All PCRs were replicated three times each with the first-strand cDNAs made from two independent RNA preparations.

cDNA Cloning for *GmMAN1* and *GmMAN4* and Protein Expression. Plant endo-β-mannanases are cell wall proteins and contain signal peptides. Previous studies showed that the removal of signal peptides is important for the activity of mature proteins (Nonogaki et al. 2000). We first analyzed the protein sequences of GmMAN1 and GmMAN4 using the SignalP program (http://www.cbs.dtu.dk/services/SignalP/) and their corresponding signal peptides were identified (first 28 and 18 amino acids for GmMAN1 and GmMAN4 respectively). To clone the truncated *GmMAN1* and *GmMAN4* without the signal peptide, specific primers for cloning were used (**Table 5.1**). PCR products were cloned into the pEXP5/CT-TOPO vector and fully sequenced. To express recombinant GmMAN1 and GmMAN4, the protein expression construct was transformed into the *E. coli* strain BL21 (DE3) CodonPlus (Stratagene, La Jolla, CA). Protein expression was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) at a concentration of 500 μM for 18 h at 25 °C.

GmMAN1 was also expressed with a His-tag. GmMAN1 was first subcloned into the pET100/D-TOPO vector (Invitrogen, Carlsband, CA). *E. coli*-expressed GmMAN1 with a His-tag was purified from the *E. coli* cell lysate using Ni-NTA agarose following the manufacturer instructions (Invitrogen). Protein purity was verified by SDS–PAGE and protein concentrations were determined by the Bradford assay (Bradford 1976).

Endo-β-mannanase Hydrolase Activity Assay. Qualitative analysis of hydrolase activity of GmMAN1 and GmMAN4 was performed using a modified gel diffusion method (Still et al. 1997). The agarose 0.8% (w/v) plates containing 0.05% (w/v) locust

bean galactomannan (Sigma) were solidified, and wells were formed on the plates by scoring with a 3-mm cork borer and removing the plug by suction. The extracts (2 µL) from purified recombinant protein solution were applied to the wells, and the plates were incubated at 28 °C for 20 h. After incubation, the agarose gel plates were stained by 0.5% (w/v) congo red dye (Sigma). The hydrolyzed areas were visible as clear circles on a dark background. The diameter of the hydrolyzed area is logarithmically related to the enzyme activity. The 1000-fold diluted commercial endo-β-mannanase from *Aspergillus niger* (Megazyme, Bray, Ireland) was used as a positive control.

The kinetic parameters of GmMAN1 were determined by measuring reducing sugars released after incubation with locust bean gum (LBG) in 0.1 M pH 5.0 McIlvaine buffer at 28 ^{o}C for 30 min according to a previously published method (Miller 1959). Absorbency was measured at 540 nm. The standard curve was obtained with D-mannose. One unit of endo-β-mannanase activity was defined as the amount of enzyme required to liberate 1 µmol of mannose per minute at a given assay temperature. Increases in reaction rate by increasing concentrations of LBG were evaluated through the endo-β-mannanase activity assay described above and were found to obey Michaelis-Menten kinetics. Appropriate enzyme concentration and incubation time were determined in time course assays. For determination of the catalytic parameters Km and Vmax, LBG served as the substrate and was used at concentrations of 1-10 mg/mL. The optimal pH of endo-β-mannanase activity was examined at pH 2.2-8.0 using 0.1 M McIlvaine buffer.

The effect of temperature on endo-β-mannanase activity was determined by

incubating the purified enzyme with the substrate at temperatures ranging from 0 to 60 °C in 0.1 M McIlvaine buffer, pH 5.0. Thermal stability of the enzyme was determined by assaying for residual enzyme activity after incubation at various temperatures for 30 min in 0.1 M McIlvaine buffer, pH 5.0. The effect of various metal ions on endo-β-mannanase activity was determined by assaying for residual activity after adding metal ions to reactions in the form of chloride salts at 5 mM final concentration. Final values represent the average of three independent measurements.

Isolation of Cell Walls from Soybean Seeds. Williams 82 soybean seeds were ground and homogenized in 1 ml distilled water and centrifuged at 10,000 g for 5 min. After decanting the supernatant, the pellet was washed three times, each time with 1 M NaCl, 70% (v/v) ethanol, and chloroform-methanol (2:1) and then dried at room temperature. The dried cell wall material (20 mg) was suspended in 0.1 M McIlvaine buffer, pH5, and subjected to enzyme digestion with the purified recombinant GmMAN1 protein (about 10 μg) at 28 °C for 20 h. Reducing sugars released into the supernatant were assayed by the dinitrosalicylic acid (DNS) method using mannose as the standard (Miller 1959).

RESULTS

Identification of Endo-β-mannanase Genes in the Soybean Genome and Phylogenetic Analysis. Twenty-one genes encoding proteins that are significantly homologous to known plant endo-β-mannanases were identified from the soybean

genome. These 21 genes are distributed in 14 chromosomes. The average length for these 21 mannanases is 422 amino acids, with the shortest being 365 amino acids and the longest being 471 amino acids.

To investigate the evolutionary relationships between the soybean mannanase genes and endo-β-mannanases of other origins, a phylogenetic tree of endo-β-mannanases was constructed (**Figure 5.1**). The 21 endo-β-mannanase genes in soybean are clustered into three major clades. Based on their gene structure and relatedness to functional known plant endo-β-mannanases, four genes were selected for further experimental study. These include *Glyma19g41090.1* (*GmMAN1*), *Glyma03g38490.1* (*GmMAN2*), *Glyma03g38840.1* (*GmMAN3*) and *Glyma01g37720.1* (*GmMAN4*).

Gene Expression Analysis of Four Soybean Endo-β-mannanase Genes. Soybean RNA was obtained from a pool of soybean tissues, including roots, epicotyls, hypocotyls, cotyledons, leaf buds, first foliage leaves and imbibed soybean seeds. RNA was isolated and reverse-transcribed into first-strand cDNA, which was used as the template for PCR. Using specific primers, the amplified PCR products of soybean mannanase candidate genes were gained for *GmMAN1* and *GmMAN4* (**Figure 5.2**). Both of these amplified PCR products were cloned into a pEXP5/CT-TOPO vector and fully sequenced.

GmMAN1 and *GmMAN4* contain the opening reading frame sequences of 1233 bps and 1296 bps, which encode proteins of 410 and 431 amino acids, respectively.

GmMAN1 and GmMAN4 show moderate levels of sequence similarity to known plant endo-β-mannanases. For example, they are about 57% similar to LeMAN4a. Compared with LeMAN4a, whose three-dimensional (3-D) crystal structure has been solved (Bourgault et al. 2005), the active sites of GmMAN1, including W97, N212, W295, W371 and I409, are strictly conserved.

Molecular Cloning of GmMAN1 and GmMAN4 and their Biochemical Activities. The presence of a signal peptide in an endo-β-mannanase can interfere with its hydrolase activity (Nonogaki et al. 2000). Similarly, proteins expressed using full-length cDNAs of GmMAN1 and GmMAN4 did not show endo-β-mannanase activity. When the truncated forms of GmMAN1 and GmMAN4 without signal peptide were assayed, GmMAN1 showed endo-β-mannanase activity but no endo-β-mannanase activity was detected for GmMAN4 (**Figure 5.3**).

Expression Patterns of *GmMAN1*. To identify the specific tissues of soybean seedlings in which *GmMAN1* was expressed, total RNAs were isolated from 14-day old soybean seedling tissues, including roots, epicotyls, hypocotyls, cotyledons, first foliage leaves and leaf buds. In addition, soybean seeds imbibed in water for 36 and 48 h were also collected. The cDNAs synthesized from individual RNAs were used for semi-quantitative RT-PCR analysis. The expression of *GmMAN1* was only observed in the cotyledons of seedlings but not in other tissues examined (**Figure 5.4**).

Biochemical Properties of GmMAN1. Recombinant GmMAN1 with a His-tag expressed in *E. coli* was purified to electrophoretic homogeneity (**Figure 5.5**) and subjected to detailed biochemical characterization. When enzyme assays were performed in buffers of different pH values, GmMAN1 showed the highest level of catalytic activity at pH 5.0 (**Figure 5.6A**). The optimal temperature for GmMAN1 activity was 50 ^{o}C (**Figure** 5.**6B**). The enzyme was stable up to 37 ^{o}C but about 80% of its activity was lost at 50 ^{o}C after 0.5 h of incubation (**Figure 5.6C**).

The effects of various metal ions on GmMAN1 activity were measured. While Zn^{2+}, Fe^{2+}, NH_4^+, K^+ and Na^+ had minimal effect on GmMAN1 activity, GmMAN1 activity was strongly inhibited by Mn^{2+}, and moderately inhibited by Cu^{2+} and Ca^{2+} (**Figure 5.6D**).

The kinetic parameters of GmMAN1 were also determined. All assays were conducted at 28 ^{o}C in McIlvaine buffer, pH 5.0. The Km of GmMAN1 was determined to be 34.86±4.78 mg/mL, Vmax was 33.80±3.99 μmol/min·mg and Kcat was 25.2 s^{-1}. The catalytic efficiency of GmMAN1 (kcat/Km) was determined to be 0.72 ml/mg·s.

Activity of Recombinant GmMAN1 on Isolated Cell Wall from Soybean Seeds. To test whether the recombinant GmMAN1 can hydrolyze the soybean cell wall, cell walls were isolated from soybean seeds. After 20 h of reaction, 10 μg of GmMAN1 recombinant enzyme hydrolyzed approximately 0.2 mg of the isolated soybean cell walls.

DISCUSSION

Hemicellulose β-mannan is rich in many seeds and has important biological functions. In tomato seeds, β-mannan is an important component of endosperm and plays a role in regulating seed germination (Nonogaki et al. 2000). In seeds of other plants, including soybean, β-mannan mainly serves as a food reserve to support seedling growth after germination (Reid and Bewley 1979). Because there is essentially no endosperm in the mature soybean seeds, the high content of mannan is very likely localized in the cell walls of cotyledons. Despite its biological importance, the high content of β-mannan in soybean meal is undesirable (Duncan et al. 2002). The properties of soybean meal, including carbohydrate content and composition, can be modified genetically. Our long term goal of this study is to improve soybean meal quality, attributed to β-mannan, through genetic engineering. To this end, our first objective was to identify a useful soybean endo-β-mannanase gene.

Endo-β-mannanase genes belong to an ancient gene family, which is present in plants (Bewley et al. 1997; Marraccini et al. 2001; Wang et al. 2004), microorganisms (Dhawan and Kaur 2007) and animals (Yamamura et al. 1996). Phylogenetic analysis suggested that plant endo-β-mannanase genes share a common evolutionary origin (Yuan et al. 2007). Our research presents the first report, to our knowledge, on functional characterization of endo-β-mannanase genes from soybean. Despite the well-characterized functions of endo-β-mannanase in seed germination (Nonogaki et al. 2000), fruit ripening (Carrington et al. 2002) and pollen development (Filichkin et al. 2004), the knowledge of endo-β-mannanases in plant biology is still limited. Because

β-mannan is a component of the primary cell wall of many plants (Carpita et al. 2001; Hosoo et al. 2006), the modification of β-mannan via the action of endo-β-mannanase is likely involved in many biological processes. It is interesting that the soybean genome contains a relatively large family of endo-β-mannanase genes with 21 members (**Figure 5.1**), in contrast to Arabidopsis, rice and poplar, which contain 8, 9 and 11 genes, respectively (Yuan et al. 2007). Therefore, soybean represents a useful model for studying the biochemical and biological functions of endo-β-mannanase gene. For instance, GmMAN4, which is 55% identical to GmMAN1 at the protein sequence level, did not show hydrolase activity (**Figure 5.3**). It will be interesting to determine the specific biological role of GmMAN4.

A number of endo-β-mannanse genes have been cloned from various plant species, including tomato, coffee and lettuce (Bewley et al. 1997; Marraccini et al. 2001; Wang et al. 2004). However, at the biochemical level, most of these genes were only characterized to have hydrolase activity without detailed knowledge of their kinetics. Due to the many commercial uses of endo-β-mannanases, many microbial endo-β-mannanases have been fully characterized biochemically (Hogg et al. 2003; Xu et al. 2002). To our knowledge, GmMAN1 is among the first plant endo-β-mannanases to be fully characterized for its kinetics. The Km of GmMAN1 is about 35 mg/mL on locust bean gum, which is significantly higher than those of microbial endo-β-mannanases. For instance, the Kms for endo-β-mannanases from *Aspergillus tamarill* and *Streptomyces lividans* were reported to be 0.02 mg/mL and 0.77 mg/mL, respectively (Arcand 1993; Civas et al. 1984). While these differences may be due to

the different mannan substrates used in the assays, they also suggest that GmMAN1 is less efficient than microbial endo-β-mannanases. Nonetheless, the detailed kinetic parameters of GmMAN1 presented in this study will be useful for a rational design of a genetic engineering strategy for reducing β-mannan in the soybean meal.

With an endo-β-mannanase gene at hand, our next objective will be to test the effect of overexpression of *GmMAN1* on the β-mannan content of soybean. This could be conducted using different strategies. One is to overexpress *GmMAN1* in developing seeds. Assuming that mannan is synthesized during seed development and maturation, the overexpression of *GmMAN1* in cotyledons will lead to immediate degradation of this hemicellulose. However, if this would cause problems for seed development, an alternative strategy is to accumulate the proteins in a different compartment, such as endoplasmic reticulum, which will physically separate the enzyme from its substrate in intact seeds. Overproduced endo-β-mannanase can function in hydrolyzing β-mannan only after the soybean seeds have been milled, a strategy mimicking the addition of microbe-produced endo-β-mannanase to soybean meal. In this way, the potential negative impact of endo-β-mannanase overproduction on seed vigor and seedling establishment can be minimized.

REFERENCES

Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman D (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic. Acids Res. 25: 3389 – 3402

Anderson JO, Warnick RE (1964) Value of enzyme supplements in rations containing certain legume seed meals or gums. Poult Sci. 43: 1091-1097

Arcand NK, D; Paradis, F.W.; Morosoli, R.; Shareck, F. (1993) β-mannanase of *Streptomyces lividans* 66: cloning and DNA sequence of the manA gene and characterization of the enzyme. Biochem. J. 290: 857-863

Bernard RL, Cremeens CR (1988) Registration of 'Williams 82' soybean. Crop Sci. 28: 1072

Bewley JD, Burton RA, Morohashi Y, Fincher GB (1997) Molecular cloning of a cDNA encoding a (1→4)-β-mannan endohydrolase from the seeds of germinated tomato (*Lycopersicon esculentum*). Planta 203: 454-459

Bourgault R, Oakley AJ, Bewley JD, Wilce MCJ (2005) Three-dimensional structure of (1,4)-β-d-mannan mannanohydrolase from tomato fruit. Protein Sci. 14: 1233-1241

Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 7: 248-254

Carpita NC, Defernez M, Findlay K, Wells B, Shoue DA, Catchpole G, Wilson RH, McCann MC (2001) Cell wall architecture of the elongating maize coleoptile. Plant Physiol. 127: 551-565

Carrington CMS, Vendrell M, Domínguez-Puigjaner E (2002) Characterisation of an endo-(1,4)-β-mannanase (LeMAN4) expressed in ripening tomato fruit. Plant Sci. 163: 599-606

Civas A, Eberhard R, Le Dizet P, Petek F (1984) Glycosidases induced in *Aspergillus tamarii*. Secreted alpha-D-galactosidase and beta-D-mannanase. Biochem. J. 219:857–863.

Dhawan S, Kaur J (2007) Microbial mannanases: an overview of production and applications. Crit. Rev. Biotechnol. 27: 197-216

Duncan CJG, Pugh N, Pasco DS, Ross SA (2002) Isolation of a galactomannan that

enhances macrophage activation from the edible fungus *morchella esculenta*. J. Agr. Food Chem. 50: 5683-5685

Ebringerová A, Hromádková Z, Heinze T (2005) Hemicellulose
Polysaccharides I. In: Heinze T (ed). Springer Berlin / Heidelberg, pp 1-67

Filichkin SA, Leonard JM, Monteros A, Liu P-P, Nonogaki H (2004) A novel endo-β-mannanase gene in tomato LeMAN5 is associated with anther and pollen development. Plant Physiol. 134: 1080-1087

Hogg D, Pell G, Dupree P, Goubet F, Martín-Orúe SM, Armand S, Gilbert HJ (2003) The modular architecture of *Cellvibrio japonicus* mannanases in glycoside hydrolase families 5 and 26 points to differences in their role in mannan degradation. Biochem. J. 371: 1027-1043

Hosoo Y, Imai T, Yoshida M (2006) Diurnal differences in the supply of glucomannans and xylans to innermost surface of cell walls at various developmental stages from cambium to mature xylem in *Cryptomeria japonica*. Protoplasma 229: 11-19

Hsiao H-Y, Anderson DM, Dale NM (2006) Levels of beta-mannan in soybean meal. Poult Sci. 85: 1430-1432

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948

Marraccini P, Rogers WJ, Allard C, André M-L, Caillet V, Lacoste N, Lausanne F, Michaux S (2001) Molecular and biochemical characterization of endo-β-mannanases from germinating coffee (*Coffea arabica*) grains. Planta 213: 296-308

Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31: 426-428

Nonogaki H, Gee OH, Bradford KJ (2000) A germination-specific endo-β-mannanase gene is expressed in the micropylar endosperm cap of tomato seeds. Plant Physiol. 123: 1235-1245

Pettey LA, Carter SD, Senne BW, Shriver JA (2002) Effects of beta-mannanase addition to corn-soybean meal diets on growth performance, carcass traits, and nutrient digestibility of weanling and growing-finishing pigs. J. Anim. Sci. 80: 1012-1019

Reid JSG, Bewley JD (1979) A dual role for the endosperm and its galactomannan reserve in the germinative physiology of fenugreek (*Trigonella foenun-graecum* L.) an endospermic leguminous seed. Planta 147: 145-150

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X-C, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178-183

Schouten HJ, Krens FA, Jacobsen E (2006) Cisgenic plants are similar to traditionally bred plants. EMBO Rep. 7: 750-753

Schröder R, Atkinson RG, Redgwell RJ (2009) Re-interpreting the role of endo-β-mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. Ann. Bot. 104: 197-204

Staykov Y, Spring P, Denev S, Sweetman J (2007) Effect of a mannan oligosaccharide on the growth performance and immune status of rainbow trout (*Oncorhynchus mykiss*). Aquacult. Int. 15: 153-161

Still DW, Dahal P, Bradford KJ (1997) A single-seed assay for endo-beta-mannanase activity from tomato endosperm and radicle tissues. Plant Physiol. 113: 13-20

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596-1599

van Kooyk Y, Geijtenbeek TBH (2003) DC-SIGN: escape mechanism for pathogens. Nat. Rev. Immunol. 3: 697-709

Wang A, Li J, Bewley J (2004) Molecular cloning and characterization of an endo-β-mannanase gene expressed in the lettuce endosperm following radicle emergence. Seed Sci. Res. 14: 267-276

Wu G, Bryant M, Voitle R, Roland D, Sr (2005) Effects of beta-mannanase in corn-soy diets on commercial leghorns in second-cycle hens. Poult Sci. 84: 894-897

Xu B, Sellos D, Janson J-C (2002) Cloning and expression in *Pichia pastoris* of a blue mussel (*Mytilus edulis*) β-mannanase gene. Eur. J. Biochem. 269: 1753-1760

Yamamura I, Nozaki Y, Matsumoto T, Kato T (1996) Purification and some properties of endo-β-1,4 mannanase from a marine mollusk *Littorina brevicula*. Biosci. Biotechnol. Biochem. 60: 674-676

Yang S, Li G, Li M, Wang J (2011) Transgenic soybean with low phytate content

constructed by *Agrobacterium* transformation and pollen-tube pathway. Euphytica 177: 375-382

Yuan J, Yang X, Lai J, Lin H, Cheng Z-M, Nonogaki H, Chen F (2007) The *endo-β-mannanase* gene families in *Arabidopsis*, rice, and poplar. Funct. Integr. Genomics 7: 1-16

Zhao N, Guan J, Ferrer J-L, Engle N, Chern M, Ronald P, Tschaplinski TJ, Chen F (2010) Biosynthesis and emission of insect-induced methyl salicylate and methyl benzoate from rice. Plant Physiol. Bioch. 48: 279-287

Zhao N, Guan J, Forouhar F, Tschaplinski TJ, Cheng Z-M, Tong L, Chen F (2009) Two poplar methyl salicylate esterases display comparable biochemical properties but divergent expression patterns. Phytochemistry 70: 32-39

Appendix

Table 5.1. Primers for RT-PCR analysis of soybean mannanase gene expression and cloning target gene into the expression vectors.

Goal	Gene Name	Orientation	Sequence (5'-3')
RT-PCR	*GmMAN1*	Forward	ATGGGTTGGAATGGACGTCTCA
		Reverse	TTAATGCTTGTTCACTAAGCAATACTTG
RT-PCR	*GmMAN2*	Forward	ATGGGGGACTCAGCTCGCATGTTAC
		Reverse	CTAAGCAATACTTGACATTTTGTGG
RT-PCR	*GmMAN3*	Forward	ATGAAGGATATAGTCAATCCTCATCATCAC
		Reverse	TTATGCAAGACTTGACATCTTCTTAG
RT-PCR	*GmMAN4*	Forward	ATGTTCCATGTCAGTGTGGAAGCAAGG
		Reverse	TCAGTTGTCAATGTATCTGCCTCCGTTTC
RT-PCR	*GmUBI-3*	Forward	GGGTTTTAAGCTCGTTGTGT
		Reverse	TCCCCTCTAGCCAATTCAGA
Cloning	*GmMAN1*	Forward	ATGGGGGACTCAGCTCGCATATTG
		Reverse	TTAATGCTTGTTCACTAAGCAATACTTG
Cloning	*GmMAN4*	Forward	ATGTTCCATGTCAGTGTGGAAGCAAGG
		Reverse	TCAGTTGTCAATGTATCTGCCTCCGTTTC

Figure 5.1. The phylogeny of 21 soybean endo-β-mannanase proteins, known endo-β-mannanases from other plants and representative microbial endo-β-mannanases. All sequences that start with 'Glyma' are soybean endo-β-mannanases. Known plant endo-β-mannanases include LeMAN1 (GenBank accession AAB87859), LeMAN3 (GenBank accession AAG14352), LeMAN4a (GenBank accession AAK97760), LeMAN5 (GenBank accession AAG00315) from tomato; CoffeeMANA (GenBank accession CAC08208), CoffeeMANB (GenBank accession CAC08442) from coffee (*Coffea arabica*); AaMAN1 (GenBank accession ACN78662) from *Actinidia arguta*; DcMAN1 (GenBank accession AAN34823) from carrot (*Daucus carota*); MdMAN1 (GenBank accession ACN78663), MdMAN2 (GenBank accession ACN78664), MdMAN3 (GenBank accession ACN78665) from apple (*Malus domestica*); MaMAN1 (GenBank accession ABF69949) from *Musa acuminate* and VcoMAN1 (GenBank accession ACN78666) from *Vaccinium corymbosum*. The microbial endo-β-mannanases include BliMAN1 (AAU23418.1) from *Bacillus licheniformis*; TneMAN1 (CAB56856.1) from *Thermotoga neapolitana*, TmaMAN1 (AAD36302.1) from *Thermotoga maritima*, AbiMAN1 (CAB76904.1) from *Agaricus bisporus* and AacMAN1 (AAA67426.1) from *Aspergillus aculeatus*. Bootstrap values larger than 50% were shown in the phylogenetic tree. The four endo-β-mannanases studied in this chapter, GmMAN1, GmMAN2, GmMAN3 and GmMAN4, are indicated.

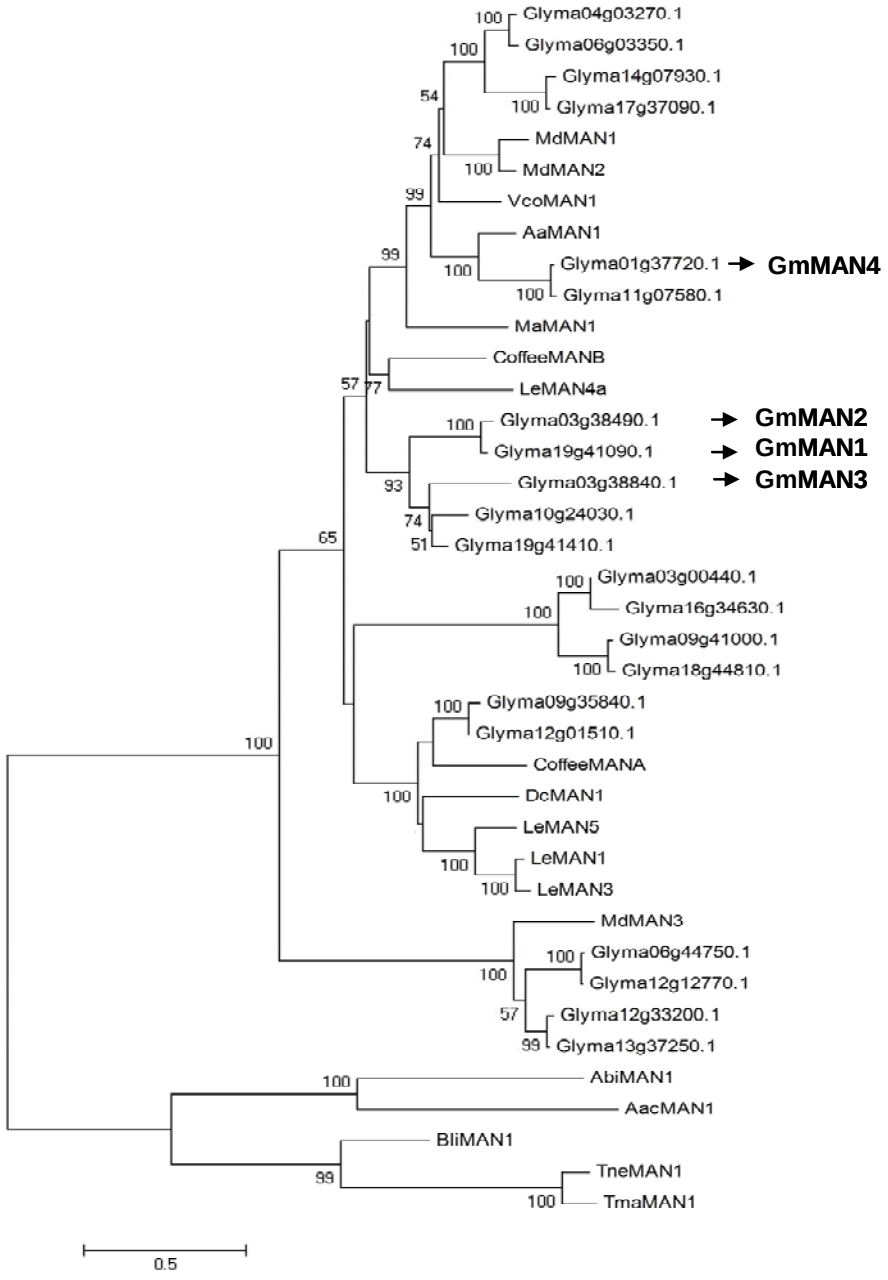

Figure 5.1. The phylogeny of 21 soybean endo-β-mannanase proteins, known endo-β-mannanases from other plants and representative microbial endo-β-mannanases.

Figure 5.2. Expression analysis of four soybean endo-β-mannanase genes *GmMAN1*, *GmMAN2*, *GmMAN3* and *GmMAN4* using semi-quantitative RT-PCR. Soybean RNAs pooled from soybean tissues, including roots, epicotyls, hypocotyls, cotyledons, leaf buds, first foliage leaves and imbibed soybean seeds, were reverse-transcribed to first-strand cDNA as the template for gene expression analysis.

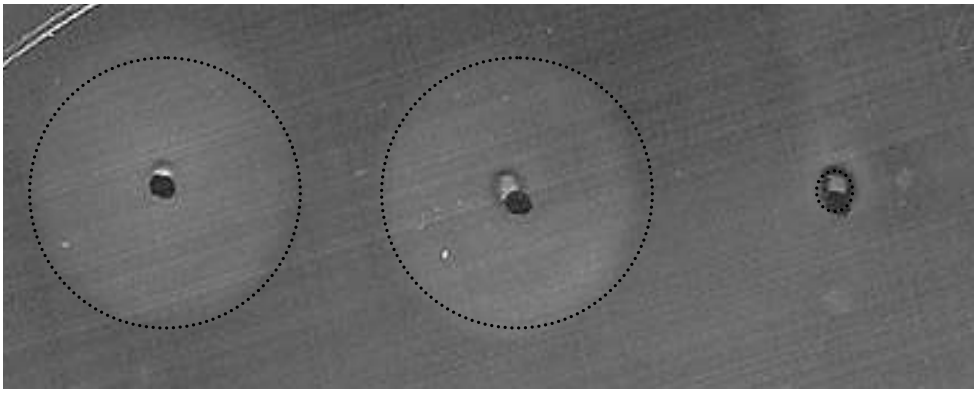

Figure 5.3. The hydrolase activity analysis of the crude protein extract of truncated *GmMAN1* and *GmMAN4* without signal peptide. Endo-β-mannanase from *Aspergillus niger* was used as positive control (Ctr), the crude protein extract of truncated *GmMAN1* and *GmMAN4* were analyzed by gel diffusion method. The size of the hydrolyzed area is indicated by a circular of broken line.

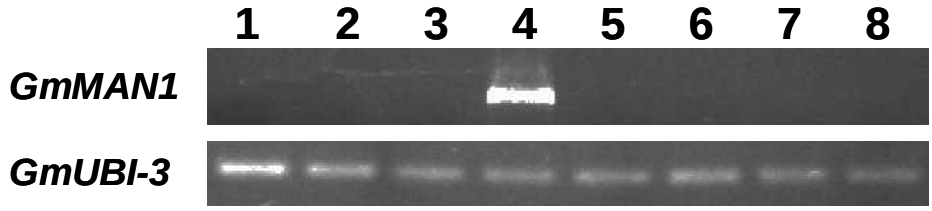

Figure 5.4. Expression of *GmMAN1* in various tissues of soybean. Semi-quantitative RT-PCR was used to analyze *GmMAN1* expression. Seedling tissues, including roots (lane 1), epicotyls (lane 2), hypocotyls (lane 3), cotyledons (lane 4), first foliage leaves (lane 5), and leaf buds (lane 6), were collected from 14 days old soybean plants. Seed samples were collected at 36 h (lane 7) and 48 h (lane 8) after imbibition. Total RNA was extracted and used for RT-PCR analysis. PCR using primers for soybean ubiquitin3 was used to judge equality of concentration of cDNA in different samples.

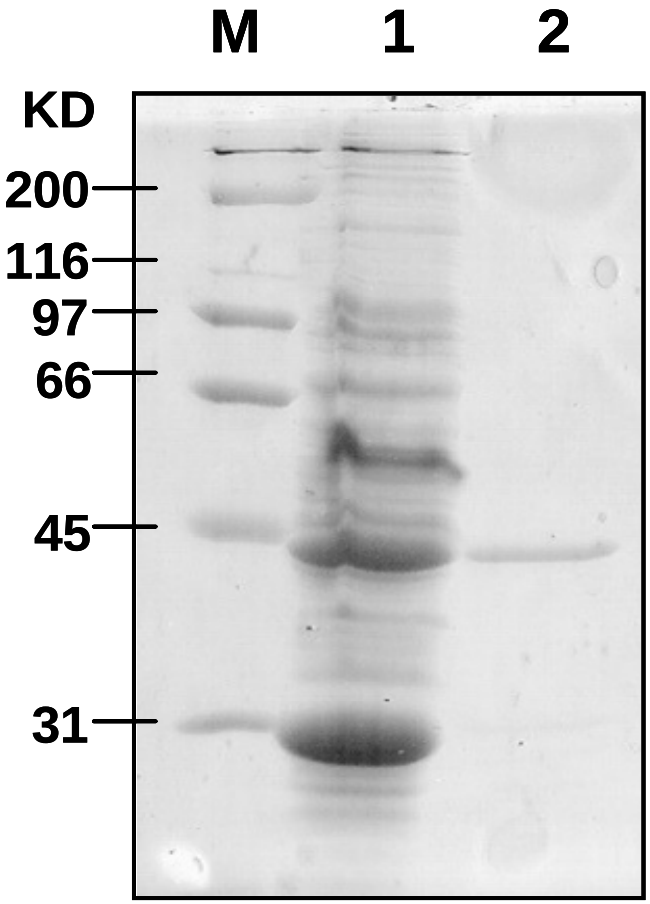

Figure 5.5. SDS-PAGE of purified recombinant GmMAN1 protein. His-tagged GmMAN1 expressed in *E. coli* was purified as described in the materials and methods. Lane M contained protein molecular weight markers. Lane 1 contained crude extract and lane 2 contained about 2 μg of purified GmMAN1 protein. The gel was stained with Coomassie blue.

Figure 5.6. Biochemical properties of GmMAN1. **A,** pH effect on GmMAN1 activity. Activity of the purified GmMAN1 was assayed from pH 2.2-8.0. Level of GmMAN1 activity in 0.1M McIlvaine buffer, pH 5.0, was arbitrarily set at 1.0. **B,** Optimal Temperature of GmMAN1. Activity of purified GmMAN1 was assayed at temperatures ranging from 0 °C to 60 °C. The level of GmMAN1 activity at 50 °C was arbitrarily set at 1.0. **C,** Temperature effect on GmMAN1 stability. Thermostability of the purified GmMAN1 was determined by pre-incubating samples at temperatures ranging from 0°C to 60°C for 30 min and subsequently assaying the remaining activity. **D,** Effects of metal ions on activity of GmMAN1. Metal ions were added to reactions in the form of chloride salts at 5 mM final concentrations. The level of GmMAN1 activity without any metal ion added as control (Ctr) was arbitrarily set at 1.0.

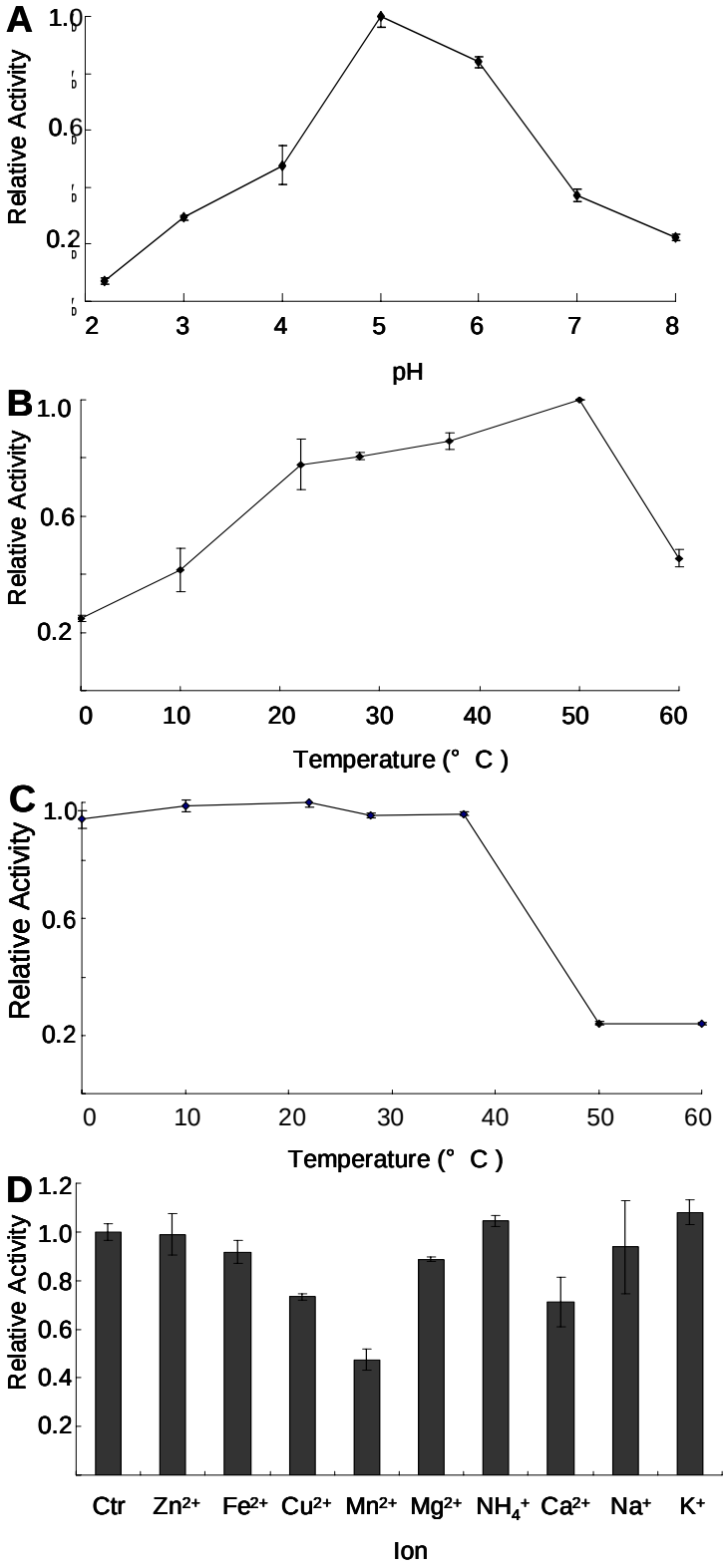

Figure 5.6. Biochemical properties of GmMAN1.

Chapter VI. Conclusions and Perspectives

1. Conclusions

In this dissertation, comparative functional genomics approaches were employed to study four soybean gene families, which included *SABATH* gene family, methyl ester esterase gene family, terpene synthase gene family, and endo-β-mannanase gene family. Six soybean genes from these above gene families have been functionally characterized. The biological functions of the first three gene families were studied in order to understand the soybean defense mechanisms in response to attacks from both above- and below-ground pests. The biological function of the endo-β-mannanase gene family was explored for soybean meal improvement. With respect to different study objectives, my research has led to several major conclusions.

A soybean hairy root system with over-expressed target genes and a reporter gene was optimized and verified to be an efficient tool with which to evaluate the function of target genes between the interaction of soybean and its root parasite, soybean cyst nematode (*Heterodera glycines*, SCN). In Chapters 2 and 3, *Agrobacterium rhizogenes*-mediated transformed soybean hairy roots were generated to study the function of *GmSAMT1* and *GmSABP2-1* genes. Aided by the reporter gene, transgenic soybean hairy roots were easily identified. It took only one month to generate transgenic hairy roots after inoculation by the *A. rhizogenes* K599 harboring the target binary vector. Compared to the traditional stable transformation method, this method is very efficient for studying the function of genes potentially involved in soybean resistance against SCN and some other root pathogens.

The soybean gene *Glyma02g06070* was identified and found to be up-regulated

only in the resistant soybean line but not in the susceptible soybean line in response to SCN infection. A similar gene expression pattern was also observed in soybean leaves under soybean mosaic virus infection. Based on the alignment and phylogenetic analysis, this gene was indicated to be a member of the *SABATH* family.

Using the biochemical assay, *Glyma02g06070* was demonstrated to encode a salicylic acid (SA) methytransferase and subsequently was designated as *GmSAMT1*. *Escherichia coli*-expressed recombinant GmSAMT1 exhibited the activity of converting SA to form methyl salicylate (MeSA) with an apparent Km value of 32.9±1.7 μM for SA.

Another soybean gene, *Glyma16g26060,* was characterized through the biochemical assay to encode a methyl salicylate esterase. It was then designated as *GmSABP2-1* since this gene is the homolog gene of a characterized tobacco SABP2 gene. *E. coli*-expressed recombinant GmSABP2-1 exhibited an apparent Km value of 46.2±2.2 μM for MeSA with the optimal pH at pH 7.0.

The biological roles of GmSAMT1 and GmSABP2-1 were assayed using the transgenic soybean hairy roots system with over-expression of these two target genes. The results suggested that they functioned in producing and accepting, respectively, the defense signal, MeSA, to increase the soybean resistance against SCN attack. The result of qPCR for soybean hairy roots with over-expression of *GmSAMT1* also revealed that the SA pathway may interact with the jasmonic acid (JA) pathway to integrate both the local and systemic acquired resistance.

Soybean plants were found to release volatiles when they were infested by

two-spotted spider mites. Such herbivore-induced volatiles have been demonstrated or implicated to function as an indirect defense signal against herbivorous insects feeding on other plant species. In my study, the components of herbivore-induced soybean volatiles included methyl salicylate, α-farnesene, (Z)-3-hexenyl acetate, 4,8,12-trimethyl-3,7,11-tridecatrienenitrile (TMTT) and a trace amount of E-β-caryophyllene.

The soybean terpene synthase gene family was investigated. Based on phylogenetic analysis with other known terpene synthases and the analysis of conserved motifs, two soybean terpene synthase genes were chosen for further characterization. One gene, *Glyma17g05500* (*GmTPS1*) was characterized to encode the specific (E,E)- α-farnesene synthase. The other gene, *Glyma12g16990* (*GmTPS2*) was characterized to encode a terpene synthase, which catalyzes the formation of four detectable sesquiterpenes dominated by the compound E-β-caryophyllene, and trace amount of (Z,E)-α-farnesene. The transcript abundance of these two genes between herbivore-damaged and undamaged soybean revealed that the two genes have different expression patterns, which was consistent with the volatile emission pattern. *GmTPS1* was up-regulated, while *GmTPS2* did not significantly change gene expression level in response to herbivory.

Among 21 soybean putative endo-β-mannanase genes, four soybean endo-β-mannanase genes with open reading frames (*GmMAN1* to *GmMAN4*) were chosen for experimental analysis. GmMAN1 showed endo-β-mannanase hydrolase activity and was specifically expressed in cotyledons of seedlings, suggesting a role of

degrading mannan-rich food reserves during soybean seedling establishment.

2. Future Perspectives

Although substantial progress has been made in this research, there are many questions that remain to be answered about soybean defenses and soybean meal improvement. Future research can be extended from the following perspectives.

2.1. The SA Signaling Pathway for Soybean Defense

In Chapters 2 and 3, two soybean genes *GmSAMT1* and *GmSABP2-1* were cloned and characterized. Their biochemical functions were demonstrated to catalyze the conversion between SA and MeSA. Since the expression of these two genes were induced in the resistant soybean recombinant inbred line in response to the infection of SCN, and not in the susceptible recombinant inbred line, over-expression of these two genes in the susceptible soybean background was utilized to examine the their biological functions. The results showed the complement of the expression of *GmSAMT1* and *GmSABP2-1* in the susceptible soybean background dramatically benefited the resistance against SCN. However, how these genes are regulated to defend soybean against SCN remains unanswered. A study to investigate the details of the SA signaling pathway will be helpful for understanding the defense mechanism. How the SA signaling pathway crosstalks with the JA signaling pathway during SCN infection is a critical question for future improvement of soybean resistance against SCN, since JA pathway was also suggested to be induced when soybean was infected (Klink et al.

2010).

2.2. Biochemical Function of Soybean TPS Genes and the Evolution of Soybean TPS Genes

In Chapter 4, two soybean TPS genes were cloned and characterized and were suggested to be associated with the formation of α-farnesene and E-β-caryophyllene. However, it is still not clear what the other terpenoids are in response to herbivory and the functions they may have. Does soybean release different volatile profiles to attract the predatory insects to respond to different herbivore insects? In Chapter 4, there are some other putative soybean TPS genes with unknown functions. Köllner et al. (2008) reported maize (*Zea mays*) emitted E-β-caryophyllene from both the leaf tissues in response to attack by lepidopteran larvae and the root tissues after damage by larvae of the coleopteran. The expression of maize terpene synthase, Terpene synthase 23 (TPS23) for producing E-β-caryophyllene, was indicated to be controlled at the transcript level and induced independently in roots and leaves. Interestingly, ZmTPS23 was observed to be maintained by positive selection in maize and its closest wild relatives. The author also argued that β-caryophyllene as a defense signal was lost during breeding of the North American lines. Therefore, to know the functions of soybean TPS genes and explore the evolution of these genes would provide the insight into how soybean plants adapt to the attack of herbivores. Continued research might help develop soybean cultivars with increased resistance against agronomically important pests.

2.3. The Relationship between MeSA and α-farnesene in Soybean Defense

In Chapter 4, both MeSA and α-farnesene were observed in the volatile profile of soybean leaves upon damage by spider mites. It was also reported that aphid-infested soybean also emitted these two compounds (Zhu and Park, 2005). In my dissertation, the genes responsible for producing these two herbivore-induced volatiles were cloned and characterized. Also their expressions were proven to be controlled at the transcript level. Therefore intriguing questions need to be answered in the future. Why do these two compounds occur at the same time and do they work together for soybean defense? Are these two soybean genes co-expressed in response to herbivory? Are they regulated by the same transcription factors and do they contain the same cis-regulatory elements? When we find out the answers to the above questions, we will be able to draw a more detailed picture about soybean defense against pathogens.

2.4. Biological Function of Soybean Mannanase Genes

In Chapter 5, one soybean endo-β-mannanase gene was cloned and characterized through biochemical assay. This gene was found to be expressed specifically in cotyledons, suggesting a role in seedling establishment. To further verify its biological function, β-mannan content of soybean may be tested on stable transgenic soybeans with over-expression or knock-out of *GmMAN1*. Since there are some other putative soybean mannanase genes, it is worthy to figure out their biochemical and biological functions. Some mannanase genes from other plant species have been characterized and shown functions in seed germination (Nonogaki et al. 2000), fruit ripening (Carrington

et al. 2002), and pollen development (Filichkin et al. 2004). Therefore, we may ask which soybean genes function at seed germination, pollen development and other important growth stages. Another interesting question is about the biochemical activity of mannanase. Despite hydrolase activity, a recent study suggested that some endo-β-mannanases may have transglycosylase activity (Schröder et al. 2009). We may also examine whether soybean mannanases contain these dual activities in the future.

Overall, this research has led to further understanding of gene functions in soybean defense against pests and soybean meal improvement.

REFERENCES

Carrington CMS, Vendrell M, Domínguez-Puigjaner E (2002) Characterisation of an endo-(1,4)-β-mannanase (LeMAN4) expressed in ripening tomato fruit. Plant Sci. 163: 599-606

Filichkin SA, Leonard JM, Monteros A, Liu P-P, Nonogaki H (2004) A novel endo-β-mannanase gene in tomato LeMAN5 is associated with anther and pollen development. Plant Physiol. 134: 1080-1087

Klink VP, Hosseini P, Matsye PD, Alkharouf NW, Matthews BF (2010) Syncytium gene expression in *Glycine max*([PI 88788]) roots undergoing a resistant reaction to the parasitic nematode *Heterodera glycines*. Plant Physiol. Bioch. 48: 176-193

Köllner TG, Held M, Lenk C, Hiltpold I, Turlings TCJ, Gershenzon J, Degenhardt J (2008) A maize (E)-β-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most american maize varieties. Plant Cell 20: 482-494

Nonogaki H, Gee OH, Bradford KJ (2000) A germination-specific endo-β-mannanase gene is expressed in the micropylar endosperm cap of tomato seeds. Plant Physiol. 123: 1235-1245

Schröder R, Atkinson RG, Redgwell RJ (2009) Re-interpreting the role of endo-β-mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. Ann. Bot-London. 104: 197-204

Zhu JW, Park KC (2005) Methyl salicylate, a soybean aphid-induced plant volatile attractive to the predator *Coccinella septempunctata*. J. Chem. Ecol. 31: 1733-1746

Vita

Jingyu Lin was born in 1979 in Changchun, Jilin Province, China, where she finished her early education. She obtained her bachelor and master degrees in 2002 and 2005 from Nankai University, Tianjin, China. After graduation, she married Nan Zhao in July, 2005. Soon after, she became a lecturer in Tianjin Modern Vocational Technology College. She came to U.S. to be with her husband in 2006 and started her Ph.D. study in University of Tennessee, Knoxville in January, 2008. Her study is focused on soybean defense against pests and soybean meal improvement using functional genomics. She received her Ph.D. in 2011.