CHARACTERIZATION OF A SOYBEAN *BAG* GENE AND ITS POTENTIAL ROLE IN NEMATODE RESISTANCE

A Thesis Presented to the Faculty of the Graduate School at the University of Missouri-Columbia

---

In Partial Fulfillment of the Requirements for the Degree
Master of Science

By

GREGORY J. YECKEL

Dr. Melissa Mitchum, Thesis Advisor

December 2012

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

FUNCTIONAL CHARACTERIZATION OF A SOYBEAN *BAG* GENE AND ITS POTENTIAL ROLE IN NEMATODE RESISTANCE

Presented by Gregory Yeckel, a candidate for the degree master of science, and hereby certify that, in their opinion, it is worthy of acceptance

_____
Dr. Melissa Mitchum

_____
Dr. James Schoelz

_____
Dr. Antje Heese

In appreciation to those whose generous help and knowledge I have relied upon.

Acknowledgments

A major research project like this is never the work of anyone alone. The contributions of many different people, in their different ways, have made this possible. I would like to extend my appreciation especially to the following.

First and foremost I offer my sincerest gratitude and thanks to my advisor, Dr. Melissa Mitchum, who has supported me throughout my schooling with her patience and knowledge, while allowing me the room to work in my own way. I attribute most of my success to her encouragement, patience and effort and without her this thesis, too, would not have been completed. I would not have considered graduate school had it not been for Dr. Mitchum's constant, unwavering support. One simply could not wish for a better or more supportive advisor.

Next I would like to thank Dr. Pramod Kandoth, Dr. Jianying Wang and Dr. Chris Lee for aiding me with their vast technical knowledge and assistance during my project. Without their help this study would not have been successful.

I would also like to thank the many colleagues who I have worked with over the years in the Mitchum lab, thank you for your friendship and technical advice.

Last but not least, I would like to thank my family for their unconditional support, both financially and emotionally throughout my schooling. In particular, the patience and understanding shown by my mom, dad and brothers during these past years is greatly appreciated.

Table of Contents

Chapter 1

Cyst Nematode Introduction

INTRODUCTION

Biology of soybean cyst nematode

The soybean cyst nematode (SCN), *Heterodera glycines*, is an obligate, sedentary endoparasitic nematode that has spread across the soybean producing regions of the United States. Today, SCN is considered the most economically important pathogen of soybean (*Glycine max*). The average bushels of soybean lost to SCN on an annual basis in the United States during 2006 to 2009 was estimated to be 128.6 million bushels, and this loss was valued at $1.286 billion (Koenning & Wrather, 2010).

SCN accomplishes this devastation by establishing an intimate relationship with the host's root system (Hussey & Grundler, 1998). The life cycle of soybean cyst nematode consists of five life stages punctuated by four molts over its 25-30 day life cycle. SCN juveniles penetrate the host's root system using a hollow mouth spear (stylet) and migrate to root vascular cells where a selected cell is reconstructed into a complex feeding sink, called a syncytium. The syncytium is a metabolically, highly active, multinucleate feeding structure formed by partial dissolution of the cell walls surrounding the initial syncytial cell selected by the nematode that leads to the fusion of adjacent cells (Endo, 1986). The syncytium provides a conduit for the nematode to obtain life-sustaining nutrients at the expense of the host plant. Once the nematode initiates syncytium formation, it becomes sedentary due to the loss of somatic musculature and is completely dependent on the successful formation of the syncytium to complete

its life cycle. In a compatible (susceptible) interaction the nematode is able to successfully create a syncytium. As the nematodes feed, the females swell and eventually grow so large that they break through the root tissue and are visible on the surface of the root. Males migrate out of the roots and fertilize the females. When the female dies, her body forms the characteristic lemon-shaped cyst, which encases between 50-300 viable eggs and can persist in the soil for years (Wrather et al., 1984).

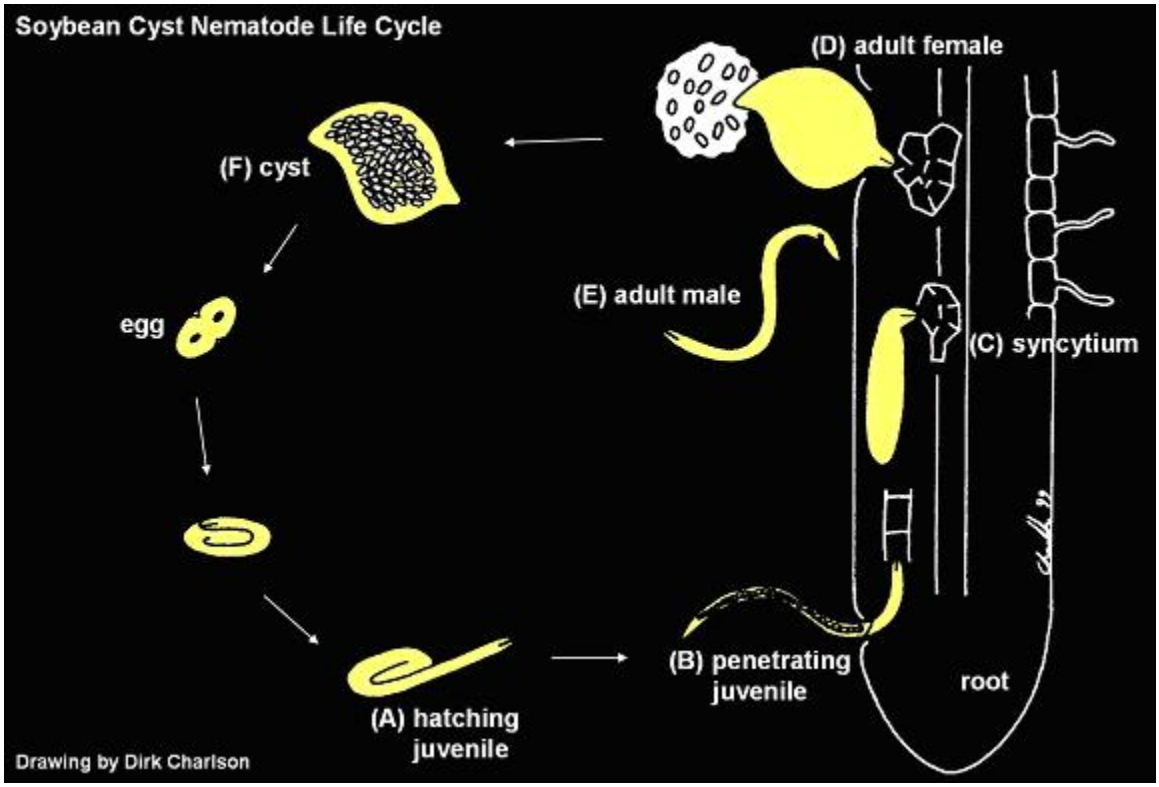


Figure 1.1. Soybean cyst nematode life cycle.

Nematode resistance in soybean

Planting of resistant soybean cultivars is the primary strategy used to manage SCN. Resistant cultivars were developed through identifying SCN-resistant soybean germplasm from numerous plant introductions (PIs) and

incorporating the trait through conventional breeding programs. Although numerous sources of resistance have been identified, only a few PIs have been used in conventional breeding programs to develop resistant cultivars. The effectiveness of resistant cultivars has been reduced due to the repeated planting of resistant cultivars from a narrow genetic base, which has led to the adaptation of more virulent nematodes able to break the resistance (Mitchum et al., 2007). Thus, there is an urgency to identify new types of resistance or develop novel approaches for resistance through the use of biotechnology.

The most common sources of resistance used in commercially available SCN-resistant cultivars are derived from PI 88788, PI 54840 (Peking), and PI 437654, due to its broad nematode resistance (Concibido et al., 2004). In resistant cultivars, the infective juveniles are able to penetrate host roots and begin to form syncytia (Endo, 1965), but the syncytia become necrotic and degenerate. This degeneration has been described as a hypersensitive response (HR)-like cell death response (Mahalingam & Skorupska, 1996). Presumably, the collapse of the syncytium deprives the nematode of nutrition and effectively starves the nematode to death. The timing of the HR and the degeneration of syncytia differ among resistant cultivars (Acedo et al., 1984). In Peking, syncytial collapse is observed as quickly as 2 days post infection (Mahalingam & Skorupska, 1996), whereas in PI 209332, the HR response is delayed, with degeneration of the syncytia not occurring until 8 to 10 days post infection (Acedo et al., 1984). On a cellular level, cell wall apposition formation and

nuclear break down leading to cytoplasmic collapse has been observed within degenerating syncytia (Kim et al., 2010).

Genetic understanding of soybean resistance

Very little is known about the molecular mechanisms underlying resistance to SCN in soybean. Research conducted in the past decades has identified a number of quantitative trait loci (QTLs) associated with SCN resistance in different PIs (Concibido et al., 2004). Among these, two QTLs, *Rhg1* (for resistance to *Heterodera glycines*) on soybean chromosome 18 and *Rhg4* on chromosome 8 have been consistently mapped in a variety of germplasm sources. *Rhg1* exhibits incomplete dominance and is necessary for SCN resistance in most resistant PI's tested, such as PI 88788, PI 90763, PI 209332 and Peking. Thus, soybeans heterozygous for the susceptible and resistant alleles at the *Rhg1* locus allow SCN to reproduce at a level that is intermediate between resistant and susceptible cultivars (Melito et al., 2010). The other major SCN resistance QTL mapped in soybean PIs, *Rhg4*, has been shown to be necessary for full resistance to select races of SCN and is dominant (Matson & Williams, 1965). In certain cultivars such as PI 437654 and Peking, *Rhg4* is necessary for full resistance, but is not required for SCN resistance in PI 88788 or PI 209332 (Brucker et al., 2005). The genes responsible for resistance at the *Rhg1* and *Rhg4* loci were recently reported. *Rhg1*-mediated resistance is conferred by a cluster of three genes, coding for an amino acid transporter, an α-SNAP protein, and a WI12 (wound-inducible domain) protein, which appear to contribute to SCN resistance by copy number variation (Cook et al., 2012). In PI

88788, the three genes are duplicated in tandem repeats ten times leading to an innate overexpression of the genes, whereas, in susceptible cultivars only one copy of the gene cluster is present. For *Rhg4*, a serine hydroxymethyltransferase (*SHMT)* gene was shown to confer resistance due to polymorphisms between the resistant and susceptible lines that alter the regulatory properties of the enzyme (Liu et al., 2012). Although these genes have been identified, the signaling pathways and underlying molecular mechanisms of how these genes contribute to resistance is still unknown.

Unraveling the molecular basis of resistance

In an attempt to identify genes involved in resistance to SCN, microarray studies have been conducted to compare the transcriptional responses of resistant and susceptible lines of soybean to SCN. The first studies used whole soybean roots infected with SCN to assess transcriptional changes of resistant lines to SCN (Alkharouf et al., 2004). More recently, laser capture microdissection (LCM) of syncytial cells coupled with microarray analysis has been useful in expanding our understanding of the genes involved in resistance (Kandoth et al., 2011, Klink et al., 2009, Klink et al., 2007). These studies have provided candidate genes to explore for a role in SCN resistance, which should help reveal the underlying molecular events leading to syncytium degeneration.

A prior study in our lab employed LCM and microarray analysis for a comparative analysis of syncytia gene expression profiles using near-isogenic lines (NILs) to identify transcripts regulated by specific soybean resistance genes

(Kandoth et al., 2011). The advantage of using NILs is that the effects of specific gene loci can be analyzed with reduced genetic background effects. Soybean NILs, derived from a cross between the susceptible cv. Evans and the resistant PI 209332, were used for comparative analysis. These NILs are predicted to share 98% of their genome, differing only at the major SCN resistance locus, *Rhg1* (Mudge, 1999). Syncytial cells from infected soybean roots were captured by LCM, the RNA was then isolated from syncytial cells, labeled, and hybridized onto the Affymetrix GeneChip Soybean Genome Arrays. Out of the 37,000 genes printed on the chip, 1,447 genes were differentially expressed within syncytia of NIL-R and NIL-S lines. From this analysis, 828 genes were found to be upregulated in syncytia of the resistant line compared to the susceptible line (Kandoth et al., 2011).

Out of the 828-upregulated genes in the NIL-R line, 241 of them were related to stress and defense. Defense-related genes up-regulated within syncytia of the resistant line included those predominantly involved in apoptotic cell death, the plant hypersensitive response, and salicylic acid-mediated defense signaling (Kandoth et al., 2011). A soybean gene coding for a BCL2-associated athanogene (BAG) domain protein with highest homology to the *Arabidopsis* BAG6 protein (AtBAG6) was the most highly up-regulated gene in syncytia of the resistant line (87-fold; Kandoth et al., 2011). Quantitative real-time RT-PCR analysis using RNA isolated from excised SCN-infected whole root pieces at different time points post inoculation (2, 4, 6, and 8 dpi) confirmed the specific induction of this soybean *BAG* gene in the resistant soybean line

whereas expression was barely detectable in the susceptible variety of soybean in response to SCN (Kandoth et al., 2011). In addition, the level of up-regulation was much lower in response to infection by a nematode population compatible on the resistant soybean line providing further evidence of a role for this gene in SCN resistance.

What are BAG proteins?

BAG proteins were first described in mammalian systems (Takayama et al., 1995). BAG proteins belong to a family of genes found to modulate a diverse number of cellular processes such as the activity of molecular chaperone heat shock protein 70 (Hsp70) and function as regulators of apoptosis (Yan et al., 2003) with a range of activities from inhibition to promotion of cell death (Reed, 1994). The members of the BAG family are distinguished by the presence of a conserved C-terminal BAG domain spanning 110–130 amino acids. The BAG domain is comprised of three α helices of 30–40 amino acids each (Doong et al., 2002). A total of six *BAG* genes, *BAG-1* to *BAG-6*, were identified in humans that encode proteins with at least one BAG domain (Takayama et al., 1999). BAG proteins have also been documented in multiple organisms such as mice, *Xenopus*, *Drosophila*, *Bombyx mori* (silk worm), *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Arabidopsis* (Briknarova et al., 2001, Moribe et al., 2001, Takayama & Reed, 2001, Takayama et al., 1999, Thress et al., 1998). BAG-1 was first identified as an interacting protein of the anti-apoptosis protein BCL-2 (Takayama et al., 1995). BCL-2 protein blocks programmed cell death, and BAG-1 enhances the

anti-apoptotic effects of BCL-2 (Sawitzki et al., 2002). BAG-1 was also found to interact with the ATPase binding domain of HSC70, HSC70 and BCL-2 to modulate their activity either positively or negatively (Pascale et al., 2010). BAG-1 contains a ubiquitin domain thought to act as a bridge between the proteasome and HSP's (Luders et al., 2000). BAG-1 achieves multiple functions by existing as three functionally distinct and differentially localized isoforms that originate from a single mRNA and interact with a broad range of cellular targets, including heat shock proteins, nuclear hormone receptors, signaling molecules, the anti-apoptotic BCL-2 protein and components of the ubiquitination and proteasome machinery (Townsend et al., 2005). The BCL-2 protein family can be broken down into three functional classes. The first class of proteins inhibits apoptosis. The second class promotes apoptosis, whereas the third class is thought to inhibit or regulate the anti-apoptotic proteins in class one (Youle & Strasser, 2008). BAG-1 and BAG-6 contain an ubiquitin domain. Although the specific function of this domain is unknown, its conservation in *C. elegans*, *S. pombe* and *Arabidopsis* homologues of BAG-1 suggests that the ubiquitin domain plays an important role in some aspect of either the function of this protein or its regulation (Takayama & Reed, 2001). BAG-3 lacks an ubiquitin domain, but contains two highly conserved tryptophans that binds proline-rich peptide motifs (WW domain) that has been implicated in protein complex assembly (Ingham et al., 2005, Takayama et al., 1999). These unique upstream sequence elements likely target individual BAG family members to their unique partners and reflect their specific

roles in different processes such as protein folding and degradation, signal transduction, and apoptosis. (Sondermann et al., 2001).

It wasn't until recently that *BAG* genes were identified in plants (Doukhanina et al., 2006, Rana et al., 2012). In the model plant, *Arabidopsis*, a family of seven genes, *BAG-1* through *BAG-7* were identified (Doukhanina et al., 2006). The proteins encoded by these *AtBAG* genes contain a conserved BAG domain, which is partially conserved between plants and mammals. The domain organization of AtBAG1-4 was found to be similar to animal homologs and contain a conserved BAG domain and an upstream ubiquitin domain. AtBAG5-7 lack the ubiquitin domain, but contain a calmodulin-binding domain (IQ) (Rhoads & Friedberg, 1997) upstream of a conserved BAG domain. The IQ domain has been shown as the site of interaction of these proteins with calcium and calmodulins (Black et al., 2005). An *Arabidopsis* gene, *IQD1* (IQ-DOMAIN 1) was identified, which encodes a novel calmodulin-binding nuclear protein. IQD1 is proposed to integrate intracellular Ca^{2+} signals towards stimulation of plant defenses, including accumulation of secondary metabolites (Levy et al., 2005). IQ domains are protein segments that contain an 11-amino acid consensus sequence defined as (I/L/V) QXXXRXXXX(R/K) (where X is any residue) (Van Petegem et al., 2005). The presence of an IQ domain is considered a unique feature found within plant BAG proteins reflecting possible divergent mechanisms within the BAG family as compared to their mammalian counterparts (Doukhanina et al., 2006).

The plant BAG members, like their mammalian homologs, have been shown to regulate apoptosis-like processes (Doukhanina et al., 2006). AtBAG4, has been shown to enhance cold tolerance and prevents features of programmed cell death when overexpressed in tobacco (Doukhanina et al., 2006). AtBAG7 was also shown to promote cell survival in *Arabidopsis* plants undergoing heat or cold stress (Williams et al., 2010). In contrast, *AtBAG6* was shown to induce cell death in both yeast cells and plant cells (Kang et al., 2006). These data indicate that, like their mammalian counterparts, the plant BAG proteins perform a plethora of different functions.

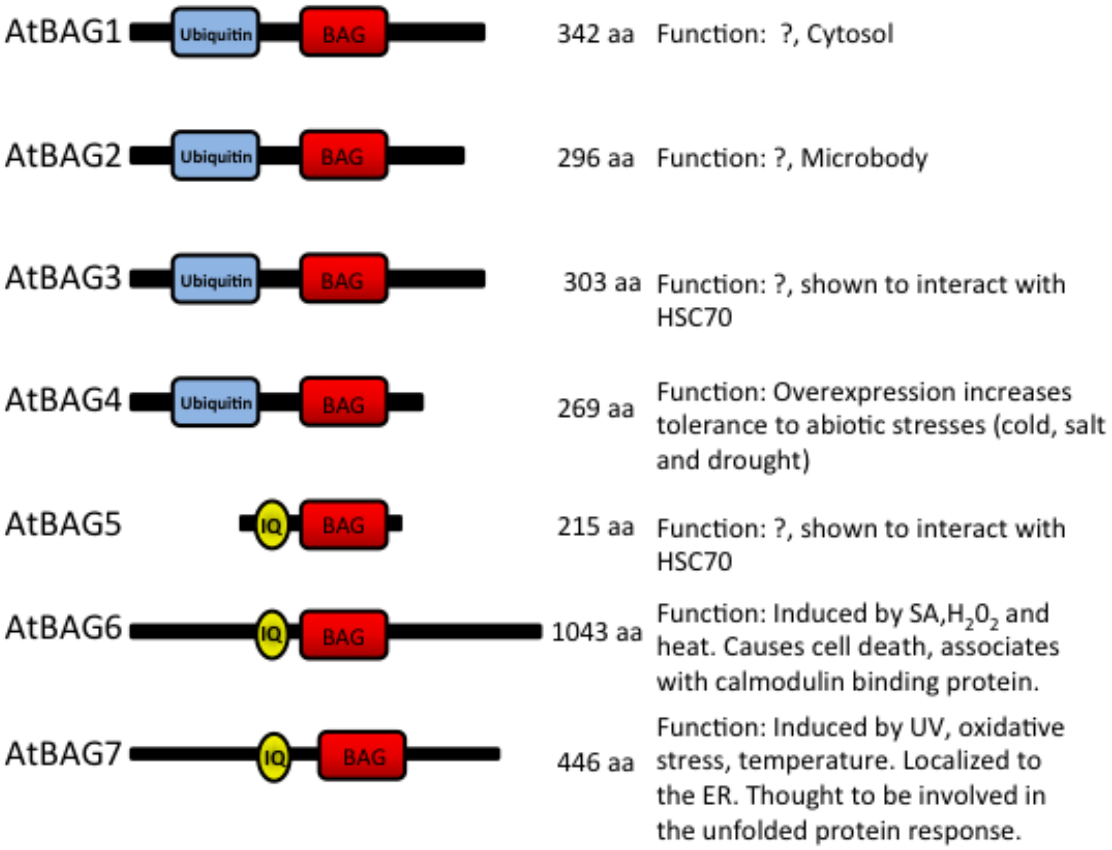


Figure 1.2. Gene models of *Arabidopsis* BAG family and their putative functions.

Bioengineering plant resistance to nematodes

Two of the primary control strategies used today for the management of plant-parasitic nematodes are nematicides and planting of resistant cultivars. However, there are limitations associated with each approach. Nematicides are toxic to the environment and not always cost-effective for agricultural producers. Repeated annual planting of resistant varieties has the potential to select for populations of nematodes that can overcome resistance. Moreover, natural resistance typically targets select species of plant-parasitic nematodes. For these reasons, researchers have begun to explore bioengineering-based control strategies to combat plant-parasitic nematodes. Bioengineering approaches can be classified into three broad categories. The first category is focused on targeting the nematode and includes expressing proteinase inhibitors (Urwin et al., 1997a, Urwin et al., 1997b, Urwin et al., 1998), BT toxins (Wei et al., 2003, Li et al., 2007) or dsRNA for RNA interference (RNAi) of endogenous nematode targets (Yadav et al., 2006). The second category is focused on targeting the plant-nematode interface, such as disrupting nematode virulence (Huang et al., 2006). The third category is focused on approaches acting through the plant response, for instance causing a hypersensitive response in the developing feeding cells. Most efforts have focused on the sedentary endoparasitic plant-parasitic nematodes because these nematodes establish intimate associations with host plants by forming feeding cells. Sedentary cyst and root-knot nematodes form syncytia and giant-cells, respectively, within host roots. The nematodes are completely dependent on the formation of these feeding cells to

complete their life cycle. Thus, targeted degradation of the feeding cells, such as is observed in naturally resistant cultivars, effectively stops the nematode from reproducing. Consequently, focus has been placed on creating transgenic plants that specifically target this interface for an effective means of resistance.

There have been limited studies exploring the use of transgenic approaches at the plant-interface to induce targeted cell death by disrupting plant physiological processes to block nematode feeding site formation. This approach has many challenges such as restricting expression to only the feeding site to prevent phytotoxicity from expression elsewhere in the plant. One study utilized a previously characterized tobacco water channel protein promoter, *TobRB7*, because it was found to be expressed in roots and induced in root-knot nematode giant-cells (Opperman et al., 1994). It was shown that fusing this promoter to a previously characterized cytotoxic protein, barnase (Prior et al., 1996), hindered giant-cell formation and negatively affected the nematode life-cycle (Conkling et al., 1998). Although this approach showed promise, it was later found that the promoter was not exclusively feeding site specific, as there was some expression of barnase in other organs, resulting in severe growth defects in the plants (Vrain, 1999).

To date, no studies have been published demonstrating targeted degeneration of feeding cells using an endogenous plant gene as a viable approach to engineer resistance. Depending on the transgene being expressed *in planta*, its expression outside the feeding site may be detrimental to yield or cause phytotoxic effects in the plant. Therefore, discovery of nematode-inducible

promoters specifically induced at the plant-nematode interface may aid in limiting any potential off-target effects. If we can successfully engineer plants that target and terminate the forming feeding cells, then we may be able to effectively engineer plants resistant to sedentary endoparasitic nematodes.

Rationale

From a prior microarray analysis, we identified a soybean *BAG* gene upregulated 87-fold higher in laser-microdissected feeding cells induced by SCN in resistant soybean compared to susceptible soybean. The consensus sequence of the corresponding probeset shared highest homology to the *BAG6* gene of *Arabidopsis*. The *Arabidopsis BAG6* gene was previously shown to induce programmed cell death (PCD) in both yeast and plants. Soybean resistance to SCN is manifested by the targeted degeneration of feeding cells in the form of a hypersensitive response (HR)-like PCD. Therefore, this soybean *BAG6* gene became a strong candidate for further characterization for a role in SCN resistance.

The work described in this thesis focused on i) cloning of the *BAG6*-like gene from soybean, ii) testing the function of this gene for a role in promoting cell death in yeast and plants, and iii) conducting proof-of-concept experiments to test its potential use in engineering a novel form of nematode resistance in plants.

The results of this study contribute new knowledge to our understanding of the soybean genes involved in SCN resistance and open up novel avenues for

bioengineering control for plant-parasitic nematodes. Orthologs of the *BAG* genes identified in this study can be identified from other plant species, which might also be exploited and translated to other major crop plants that cyst nematodes infect, such as tomato and potato, to develop biotech approaches utilizing the plants own genes for nematode control. Moreover, although it remains to be tested, this approach may have application for other types of plant-parasitic nematodes.

Chapter 2

Characterization of a soybean *AtBag6-like* gene in resistance to soybean cyst nematode, *Heterodera glycines*

ABSTRACT

Plants resistant to the soybean cyst nematode (SCN) mount a hypersensitive cell death-like response upon nematode feeding, but the genes regulating this process are not known. Laser-assisted microdissection of nematode feeding cells coupled with microarray analysis identified a soybean gene upregulated 87-fold in plants resistant to SCN that shared sequence similarity with the *Arabidopsis BAG6* (Bcl-2 associated athanogene 6) gene. *BAG* genes encode an evolutionarily conserved family of proteins in animals, yeast and plants. These proteins contain a conserved BAG domain which mediates interaction with the molecular chaperone HSP70/HSC70. Members of the BAG protein family in animals and yeast function in apoptosis to regulate a range of activities from inhibition to promotion of cell death. However, much less is known about the role of BAG proteins in plants. A family of seven *BAG* genes (*AtBAG1-7*) has been identified in *Arabidopsis*. *AtBAG6* was shown to induce programmed cell death in both yeast and *Arabidopsis*. Expression of a truncated version of the protein, spanning a calmodulin-binding IQ domain and the BAG domain, enhanced the cell death phenotype. Here, we demonstrate that similar to AtBAG6, overexpression of the full length GmBAG6A protein or a truncated version spanning the IQ and BAG domains, induced cell death in yeast and plants, with the truncated form showing an enhanced cell death phenotype. Expression of the truncated form in *Arabidopsis* and soybean under the control of a nematode-inducible promoter significantly reduced nematode development demonstrating

its potential use in engineering a novel form of nematode resistance in crop plants.

INTRODUCTION

The soybean cyst nematode (SCN; *Heterodera glycines*) continues to be a problem across soybean producing regions of the United States shows no signs of slowing. SCN is a microscopic endoparasitic roundworm of soybean that causes an estimated billion dollars in soybean yield loss annually in the United States alone (Wrather & Koenning, 2006). Major control strategies used today consist of deploying naturally resistant lines of soybean. Naturally resistant soybean lines have been used as the primary strategy to manage SCN because they have evolved a natural mechanism for resisting SCN infection. However, soybean resistance against SCN is derived from a small genetic base and repeated annual plantings of resistant soybean has selected for populations of SCN that can break the resistance. Therefore, understanding the underlying molecular mechanisms to resistance is crucial in order to improve soybean resistance to SCN and develop new biotechnological approached to nematode control.

In an attempt to identify additional genes involved in resistance to SCN, microarray studies have been conducted to compare the transcriptional responses of resistant and susceptible lines of soybean to SCN. The first studies used whole soybean roots infected with SCN to assess transcriptional changes of resistant lines to SCN (Alkharouf et al., 2004). More recently, laser capture microdissection (LCM) of syncytial cells coupled with microarray analysis has been useful in expanding our understanding of the genes involved in resistance (Kandoth et al., 2011, Klink et al., 2009, Klink et al., 2007).

Previously, we employed LCM and microarray analysis for a comparative analysis of syncytia gene expression profiles using near-isogenic lines (NILs) to identify transcripts regulated by specific soybean resistance genes (Kandoth et al., 2011). The advantage of using NILs is that the effects of specific gene loci can be analyzed with reduced genetic background effects. Syncytial cells from infected soybean roots were captured by LCM, the RNA was then isolated from syncytial cells and hybridized onto the Affymetrix GeneChip Soybean Genome Arrays. Out of the 37,000 genes printed on the chip, 1,447 genes were differentially expressed within syncytia of NIL-R and NIL-S lines. A gene coding for a BCL2-associated athanogene (BAG) domain protein with highest homology to the *Arabidopsis* BAG6 protein (AtBAG6) was the most highly up-regulated gene in syncytia of the resistant line (87-fold; Kandoth et al., 2011). *AtBAG6* was shown to induce cell death in both yeast cells and plant cells (Kang et al., 2006). Quantitative real-time RT-PCR analysis using RNA isolated from excised SCN-infected whole root pieces at different time points post inoculation (2, 4, 6, and 8 dpi) confirmed the specific induction of this soybean *BAG* gene in the resistant soybean line whereas expression was barely detectable in the susceptible variety of soybean in response to SCN (Kandoth et al., 2011). In addition, the level of up-regulation was much lower in response to infection by a nematode population compatible on the resistant soybean line providing further evidence of a role for this gene in SCN resistance.

BAG proteins were first described in mammalian systems (Takayama et al., 1995). BAG proteins belong to a family of genes found to modulate a diverse

number of cellular processes such as the activity of molecular chaperone heat shock protein 70 (Hsp70) and function as regulators of apoptosis (Yan et al., 2003) with a range of activities from inhibition to promotion of cell death (Reed, 1994). The members of the BAG family are distinguished by the presence of a conserved C-terminal BAG domain spanning 110–130 amino acids. The BAG domain is comprised of three α helices of 30–40 amino acids each (Doong et al., 2002).

Recently *BAG* genes were identified in plants (Doukhanina et al., 2006, Rana et al., 2012). In the model plant, *Arabidopsis*, a family of seven genes, *BAG-1* through *BAG-7* were identified (Doukhanina et al., 2006). The proteins encoded by these *AtBAG* genes contain a conserved BAG domain, which is partially conserved between plants and mammals. The domain organization of AtBAG1-4 was found to be similar to their animal homologs and contain a conserved BAG domain and an upstream ubiquitin domain. AtBAG5-7 lack the ubiquitin domain, but contain a calmodulin-binding domain (IQ) (Rhoads & Friedberg, 1997) upstream of a conserved BAG domain. The IQ domain has been shown as the site of interaction of these proteins with calcium and calmodulins (Black et al., 2005). IQ domains are protein segments that contain an 11-amino acid consensus sequence defined as (I/L/V) QXXXRXXXX(R/K) (where X is any residue) (Van Petegem et al., 2005). The presence of an IQ domain is considered a unique feature found within plant BAG proteins reflecting possible divergent mechanisms within the BAG family as compared to their mammalian counterparts (Doukhanina et al., 2006).

Our finding of a soybean *BAG6*-like gene upregulated 87-fold in feeding cells of the resistant line prior to their degeneration along with aforementioned studies indicating that *AtBAG6* can cause programmed cell death in yeast, led us to select this candidate for further characterization of a potential role in SCN resistance. In this study, a soybean *BAG6*-like gene was cloned and shown to induce cell death in yeast and plants, similar to *AtBAG6*. In addition, its potential use in engineering a novel form of nematode resistance in crop plants was demonstrated.

MATERIALS AND METHODS

Plant and nematode material

Soybean (*Glycine max* (L.) Merr) cv. Williams 82 (W82) and soybean near-isogenic lines (NILS) 7923R and 7923S (Li et al., 2004) were used in this study. *Arabidopsis* thaliana, ecotype Col-0, was used in this study. Soybean cyst nematode (SCN), *Heterodera glycines* inbred line PA3 (HG-type 0), was maintained on W82 according to standard procedures (Niblack et al., 1993). The beet cyst nematode (BCN), *Heterodera schachtii,* was propagated on greenhouse-grown sugar beets (*Beta vulgaris* cv. Monohi). SALK_047959C (*bag6-1*) and SALK_073331C (*bag6-2*) lines were obtained from the *Arabidopsis* Biological Resource Center (ABRC).

AtBAG6 mutant genotyping and phenotyping

Plant genomic DNA was extracted using the CTAB method from rosette leaves of 14-day-old seedlings. Plants were determined by PCR to be homozygous for the T-DNA insertion with allele-specific primers for both SALK_047959C and SALK_073331C and a left border of T-DNA primer (LBSKA). Primers used to genotype for SALK_047959C and SALK_073331C were as follows: SALK_047959C-LP 5’-GCCTGTGTACATGGATCCATC-3’, SALK_047959C-RP 5’-CTCAAAATTCCCTTCTTTGCC-3’ SALK_009534C-LP 5’-GCGATTGATTTAGTTTGGCAG-3’ and SALK_009534C-RP 5’-

TTCCTTCACATCTTGACCCTG-3’ for the gene specific primer set, and LBSKA 5’-GATTTCGGAACCACCATCAAAC-3’ to bind to the left border of the T-DNA insertion.

BAG6 cloning

The full-length *AtBAG6* (At2g46240) and *GmBAG6A* sequences were PCR amplified from cDNA generated from RNA extracted from *Arabidopsis* Col-0 leaf tissue and NIL-7923R root tissue 4-days post SCN infection, respectively. Primers used were as follows, *AtBAG6*-FP 5’-ATGATGCCTGTGTACATGGATC-3’, *AtBAG6*-RP 5’- TCATAATACGGCATCGGTTGG-3’, *GmBAG6A*-FP 5’-ATGAAGCTTGATCCATCCAAACC-3’ and *GmBAG6A*-RP 5’-TTAGGAGGATTTCATTTCATGGG-3’. Sequences amplified were then cloned into pGEM-T Easy plasmid (Promega) and verified by sequencing. The soybean NIL-7923R *GmBAG6A* cDNA sequence was deposited in Genebank under accession No. JX665043.

Soybean BAC pool screen

Prior to the release of the soybean (cv. Williams 82) genome sequence (Schmutz et al., 2010), a soybean (*Glycine max* cv. W82) BAC library developed in Dr. Henry Nguyen’s lab was screened by post-doctoral research associate Xiaolei Wu using PCR primers specific to *GmNHL1* and *GmNHL2*. Primers RS_FX88182.1F 5’-TCTGCAAACATGGCCGATAA-3’ and RS_FX88182.1R 5’-GCTCCGCCGGAGGAAT-3’ were used to screen for BACs containing *GmNHL1*.

Primers RS-MA11004.1F 5'-TTAAACGGGTGAAGGAAAAACC-3' and RS-MA11004.1R 5'-ATATCCCACGGGCAAAAGG-3' were used to screen for *GmNHL2*. BAC M0116A09 was positive for both *GmNHL1* and *GmNHL2* and fully sequenced by the Department of Energy Joint Genome Institute as part of the soybean genome sequencing project (Schmutz et al., 2010).

NHL-GUS **gene constructs**

Based on sequences obtained from sequencing of BAC M0116A09, primers incorporating restriction digestion sites into the 5' UTR for both *GmNHL1* and *GmNHL2* were designed. Primers GmNHL1-SalI-F 5' ACGCGTCGACTTCAGAATTCCCTCTCTATAAACAAA- 3' (underlined letters are non-*GmNHL1* sequences designed to incorporate a *Sal*I site into amplified fragment) and GmNHL1-BamHI-R 5'CGGATCCGGCGTAATATGAATTGAATAACC-3' (underlined letters are non-*GmNHL1* sequences designed to incorporate a *Bam*HI site into amplified fragment) were used to amplify 1kb of *GmNHL1* promoter region from BAC M0116A09 by PCR. The *NHL1* promoter fragment was column purified (Qiagen, Valencia, CA), digested with *Sal*I and *Bam*HI, and cloned into *Sal*I and *Bam*HI digested binary vector pBI121 (Jefferson, 1987) to replace the 35S promoter sequence to generate the *GmNHL1-GUS* construct. Primers GmNHL2-HindIII-F 5'-CCCAAGCTTTTCTATCTTTTGGCCCACGTA-3'(underlined letters are non-*GmNHL2* sequences designed to incorporate a *Hind*III site into amplified fragment) and GmNHL2-BamHI-R

5’CGGGATCCGCTTTAGGATTGGTGATATAATAGG-3’(underlined letters are non-*GmNHL2* sequences designed to incorporate a *Hind*III site into amplified fragment) were used to amplify 1.98 kb *GmNHL2* promoter region from BAC M0116A09 by PCR. The amplified fragment was column purified (Qiagen, Valencia, CA), digested with *Hind*III and *Bam*HI, and cloned into a *Hind*III and *Bam*HI digested binary vector pBI121 (Jefferson, 1987) to replace the 35S promoter sequence to generate the *GmNHL2-GUS* construct.

Overexpression and nematode-inducible gene constructs

The full-length *AtBAG6* and *GmBag6A* sequences, and the truncated *AtIQBAG* and *GmIQBAG* sequences were PCR amplified using sequence verified pGEM-T Easy plasmid (Promega) carrying either full-length *AtBAG6* or *GmBAG6A* cDNA as a template and cloned into the Gateway cloning vector pDONR-Zeo (Invitrogen). The clones were sequenced verified using vector-specific primers and internal sequencing primers. The sequences were then cloned from the pDONR-Zeo donor plasmid into gateway compatible vectors pAKK-2X35S and pAKK-NHL1p to generate *2X35S-GmBAG6A, 2X35S-GmIQBAG, 2X35S-AtBAG6, 2X35S-AtIQBAG, NHL1-GmIQBAG and NHL1-AtIQBAG. pAKK-NHL1* was constructed by replacing the 35S promoter from the 35S-CGT vector (Wang et al., 2010) with the 1kb *NHL1* promoter. A gateway compatible RFP cassette containing ATTB1 and ATTB2 recombination sites was amplified from CGT11017A described in (Libault et al., 2009) and inserted downstream of the *NHL1* promoter in the CGT backbone using *KpnI* and *XbaI* restriction enzymes.

The cassette containing both the promoter and RFP cassette was excised using *SBF1* restriction enzyme and subcloned into binary vector pAKK1467B, which has both green fluorescent protein (GFP) and BASTA selectable markers. Primers used for amplification for gateway compatibility were as follows: *AtBAG6*, 5′-AAAAAAGCAGGCTTAGAAGGAGATAGAACCATGATGCCTGTGTACATGGATC-3′ and 5′-AAGAAAGCTGGGTCTCATAATACGGCATCGGTTGG-3′ (underlined letters are non-*AtBAG6* sequences designed to incorporate a AttB sites into amplified fragment for gateway cloning); *AtIQBAG*, 5′-AAAAAAGCAGGCTTAGAAGGAGATAGAACCATGCCTGCAAAGAAGAGCTTTACA-3′ and 5′-AAGAAAGCTGGGTCTTAAGGCTGAGATTTAATTTCCAC-3′ (underlined letters are non-*AtIQBAG* sequences designed to incorporate a AttB sites, start and stop codons into amplified fragment for gateway cloning); *GmBAG6A*, 5′-AAAAAAGCAGGCTTAGAAGGAGATAGAACCATGAAGCTTGATCCATCCAAACC-3′ and 5′-AAGAAAGCTGGGTCTTAGGAGGATTTCATTTCATGGG-3′ (underlined letters are non-*GmBAG6A* sequences designed to incorporate a AttB sites into amplified fragment for gateway cloning); *GmIQBAG* 5′-AAAAAAGCAGGCTTAGAAGGAGATAGAACCATGGCAGGTCGAAAAGATGGAAGAG-3′ and 5′-AAGAAAGCTGGGTCTTAGTTCATTGGAGTGATTTCAACG-3′ (underlined letters are non-*AtIQBAG* sequences designed to incorporate a AttB sites, start and stop codons into amplified fragment for gateway cloning).

Soybean hairy roots

Agrobacterium rhizogenes K599 was used to generate soybean hairy roots. 1µl of plasmid DNA was added to 50µl K599 competent cells and transformed by heat shock for 5 min at 37°C, 100µl of LB was next added to the cells, then incubated with shaking at 28 °C for 1–2.5 h. Transformed cells were then plated on LB agar containing150 µg/mL kanamycin and incubated at 28 °C for 36 h. Colonies were picked and DNA was extracted from randomly selected clones and PCR was conducted to confirm presence of correct sequence. Transgenic soybean hairy roots were generated according to (Cho et al., 2000) with slight modifications. Cotyledons from 7- to 9-day old W82 seedlings were harvested and cut so that ¾ of the cotyledon remained intact. Cut cotyledons were placed in a solution of *A. rhizogenes* K599 resuspended in ¼ Gamborg's B-5 liquid medium to an OD_{600} of 0.2 to 0.3 and vacuum infiltrated for 20 minutes. The infiltrated cotyledons were incubated abaxial side up on moistened filter paper for 3 days under 16 hours light at 25 °C. Cotyledons were transferred and cultured with the cut surface facing up on MXB agar medium [MS basal nutrient salts (Caisson Laboratories, North Logan, UT), 1× Gamborg's B-5 vitamins (Sigma), 3% sucrose, 0.8% Type A agar (Sigma), pH 5.7]. The medium also contained 238 µg/mL Timentin (GlaxoSmithKline, Research Triangle Park, NC) to inhibit *A. rhizogenes* growth and GFP fluorescence was used to select for transformed roots. Approximately 10-14 days after root emergence, 1-2 cm-long root tips were transferred and freed from bacteria by 2 additional passages on the same medium at 25°C in the dark. After selection, 1-2 cm-long root tips were

transferred to MXB agar medium and allowed to grow for 1 week under the same growth conditions. The established cultures were then transferred every 2-½ weeks on MXB medium with 238 μg/mL timentin and used for nematode inoculations.

Arabidopsis transformation

Agrobacterium tumefaciens strain GV3101 was transformed with *2X35S-GmBAG6A, 2X35S-AtBAG6, 2X35S-AtIQBAG, 2X35S-GmIQBAG, NHL1-GUS, NHL2-GUS, and NHL1-AtIQBAG* by heat shocking for 5 min at 37°C, then incubated with shaking at 28 °C for 1–2.5 h. Transformed cells were then plated on LB agar containing 50 μg/mL kanamycin and incubated at 28 °C for 48 h. Colonies were picked and DNA was extracted from randomly selected clones and PCR was conducted to confirm presence of correct construct. Constructs were then transformed into *Arabidopsis thaliana* (ecotype Col-0) by the floral dipping method (Clough & Bent, 1998). Transgenic plants (T_1) were selected on 0.5X MS agar plates containing 20 μg/mL BASTA. T_1 transformants were then transferred to soil, observed for phenotypes and grown to maturity at 22 °C under long-day growth conditions. Seed was collected from individual T1 transformants and checked for a segregation ratio of 3:1 on 0.5X MS agar plates containing 20 μg/mL BASTA. Plants segregating at a 3:1 ratio were then advanced and T_3 seed was harvested from plants. T_3 seed was then plated on 0.5X MS agar plates containing 20 μg/mL BASTA to identify homozygous lines. *NHL1-GUS* T_3 lines 2-4-6, 2-4-1 and 2-3-2 were used for GUS developmental expression assays.

NHL2-GUS T_3 lines 1-2-1, 1-2-3 and 3-1-3 were used for GUS developmental expression assays.

Nematode infection assays

Isolation and hatching of nematode eggs were performed as described previously in (Mitchum et al., 2004). Nematodes were surface sterilized with sterilizing solution (0.004% [w/v] mercuric chloride, 0.004% [w/v] sodium azide, and 0.002% [v/v] Triton X-100) for 8 min followed by five washes with sterile water and resuspended in 0.1% (w/v) agarose. *Arabidopsis* seeds were sterilized using the chlorine gas method (Wang et al., 2011) for 6 h and stratified by incubation for 2 days at 4 °C. For BCN inoculations, seedlings were grown as described in (Replogle et al., 2012). Plants were seeded one per well in12-well plates, containing modified Knop's medium (Sijmons et al., 1991). Ten day old seedlings were inoculated with ~175 surface-sterilized BCN J2s. Third-stage juveniles (J3) were counted at 9 days post infection (DPI) and fourth-stage juveniles (J4) were counted at 14 DPI. Soybean hairy roots (3–4 cm) were root tip propagated and grown on MXB medium. Soybean hairy roots were inoculated with approximately 225 J2 / root in a 25-µL volume for *NHL-GUS* expression assays and 175 J2 / root in a 25-µL volume for *NHL1-GmIQBAG* infection assays as described in (Wang et al., 2007), except that 238 µg/ml ml timentin (GlaxoSmithKline, Research Triangle Park, NC, USA) was used instead of carbenicillin. Roots were inoculated approximately 1 cm above the root tip. The plates were incubated in the dark at room temperature for 30 days. After 30 days,

cysts were counted under a stereomicroscope. The experiment was conducted independently three times with a minimum of 16 independent hairy root lines per treatment. All infection assay results were plotted and analyzed for statistical significance by students *t*-test using GraphPad PRISM® software.

Histochemical GUS assays

β-Glucuronidase (GUS) assays were performed on soybean hairy roots by infiltration with GUS substrate (1.0 mM 5-bromo-4chloro-3-indoyl glucuronide, 100 mM Tris, pH 7.0, 50 mM NaCl, .006% Triton X-100, 1.0 mM potassium ferrocyanide) and incubated overnight at 37°C (Jefferson, 1987). GUS-stained root pieces were placed in 70% ethanol for long-term storage. For sectioning, GUS-stained root pieces were fixed with 4% paraformaldehyde in phosphate buffered saline (PBS), pH 7.2 for 3 hours under vacuum. Root sections were washed in PBS, pH 7.2 six times for ten min each, dehydrated with increasing concentrations of ethanol, and infiltrated with an automated Tissue-Tek VIP5 vacuum infiltration processor (Sakura Finetek) and embedded in paraffin wax. Sections 10 μm in thickness were cut from the blocks with an HM325 rotary microtome (Richard Allen) and floated in water at 50°C. Sections were mounted and cover-slipped on positively charged Superfrost Plus glass slides (Fischer Scientific, PA) and dried at 42°C for 24 hours on a slide warmer.

Microscopy

GUS-stained root photographs were taken using a Leica MZ FIII fluorescence stereo-microscope (Leica Microsystems Inc. Bannockburn, Ill). Root section photographs were taken using an AHBT3 compound microscope (Olympus, Tokyo, Japan) and captured with a Leica DFC295 (Leica Microsystems Inc. Bannockburn, Ill) digital camera using bright field microscopy.

Virus-induced overexpression in soybean

Post-doctoral research associate, Pramod Kandoth, conducted overexpression analysis in soybean. The *Bean pod mottle virus* (BPMV) VIGS vector, pBPMV and pBPMV-IA-V4, were used in this study (Zhang et al., 2010). pBPMV-IA-V4 is a derivative of pBPMV-IA-R2 containing *XhoI* and *SmaI* restriction sites between the cistrons encoding movement protein and the large coat protein subunit. A 312-bp fragment (spanning bps 1597-1908) of the *GmBag6A* cDNA sequence was amplified from soybean NIL 7923R root cDNA by RT-PCR. Primers for *GmIQBAGOE* amplification are as follows, 5′-CTCGAGATGGCAGATGCTGCTGTTTTGATACAA -3′ and 5′-GCAAGAAAGGCTTGATTCTATAATGCCCGGG -3′. (underlined letters are non-*GmIQBAG* sequences designed to incorporate restriction sites *XhoI, SmaI*, and start codon into amplified fragment for restriction site based cloning). PCR products were digested with *XhoI* and *SmaI* and ligated into pBPMV-IA-V4 digested with the same enzymes to generate pBPMV-IA-IQBAGOE. Gold particles coated with plasmid DNA corresponding to pBPMV-IA-R1M and

pBPMV-IA-IQBAGOE were co-bombarded into soybean leaf tissue as described in (Zhang et al., 2010). At 3-4 weeks post-inoculation, BPMV-infected leaves were collected, lyophilized, and stored at -20 °C for future experiments. Infected soybean leaf tissues were ground in a mortar and pestle with 0.05 M potassium phosphate buffer (pH 7.0) and used as virus inoculum for assays. The SCN-resistant RIL 7923R was inoculated with pBPMV-IA-IQBAGOE. Control plants were infected with BPMV only. Each treatment consisted of at least 12 plants. Unifoliate leaves of 9-day-old plants were rub inoculated with virus using carborundum according to (Zhang et al., 2010). Plants were grown in a growth chamber set to the following conditions: 20-21°C, 16 h light/8 h dark, and 100 µmol $m^{-2}s^{-1}$ light intensity.

RNA isolation and quantitative real-time PCR

RNA was extracted from 14-day old *Arabidopsis* seedlings using the RNeasy Plant mini kit (Qiagen, Valencia, CA) and treated with RNase-free DNase (Qiagen). cDNA was made from 1 µg of RNA using the First Strand cDNA Synthesis Kit for RT-PCR (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's instructions. Real-time quantitative RT-PCR (qPCR) was carried out using the Applied Biosystems 7500 Real-Time PCR system. qPCR primers were designed for *AtBAG6*, *GmBAG6A* and an *Arabidopsis* polyubiquitin endogenous control gene, *AtUBC10* (At4g05320) (Czechowski et al., 2005), using the Primer Express software (Applied Biosystems, Foster City, CA). The primer sets were as follows: For *AtBAG6,* 5′-TGCGTTCACGGCAACTTC -3′ and

5′-CAGCATTGATGGTACGGTACTTG -3′; *AtIQBAG,* 5′-CTCTCAAGCATTCAAGACAAGCTT -3′ and 5′-CACTTGTTCCTTTACAGCCTCTTTC -3′; *GmBAG6A,* 5′-AAGGATTAAATTCAGAGAGCTATGCAA -3′ and 5′-ACCGTGTTTCTCCATTAGAAATATTCA -3′; *GmIQBAG*, 5′-GAGACTCCTGCTGAAGTTGGATACT-3′ and 5′-TCCCTAGCCAAGGATTTTCTGA-3′; *AtUBC10*, 5′-GGCCTTGTATAATCCCTGATGAATAAG -3′ and 5′-AAAGAGATAACAGGAACGGAAACATAGT-3′. Triplicate qPCRs were set up and analyzed as described in (Wang et al., 2007).

Yeast cell death assay

Plasmids containing full-length *AtBAG6*, *GmBAG6A* and the truncated forms spanning the IQBAG domain for both genes were cloned into the pYES2-DEST52 Gateway vector system (Invitrogen). The plasmids were transformed into the wild-type *S. cerevisiae* strain, *W303-1a* (Petrezselyova et al., 2010) (MATa *ade2-1 ura3-1 his3-11 trp1-1 leu2-3 leu2-112 can1-100*). Strains were pregrown in SD medium lacking uracil in the presence of 2% glucose (SD-U/Glu) at 30°C to an optical density of .17. After washing three times in sterile water to remove any excess glucose, aliquots (10 µl) of 10-fold serial dilutions were pipetted on plates of SD medium lacking uracil, in the presence of 2% glucose (SD-U/Glu) or 2% galactose (SD-U/Gal). Plates were incubated at 30°C, and photographed after 4 days.

Tobacco agroinfiltration assay

Constructs *2X35S-GmBAG6A, 2X35S-AtBAG6, 2X35S-AtIQBAG, 2X35S-GmIQBAG,* and *2X35S-Empty Vector* (negative control) were transiently expressed in *Nicotiana benthamiana* tobacco leaves. *RX,* a *NB-LRR* protein of potato and the coat protein (CP) from potato virus X (PVX) were co-infiltrated as a positive control (Sacco et al., 2007). Constructs were introduced into *Agrobacterium tumefaciens* strain C58C1 competent cells by heat shocking for 5 min at 37°C, then incubated with shaking at 28 °C for 1–2.5 h. Transformed cells were then plated on LB agar containing 50 µg/mL kanamycin and incubated at 28 °C for 36 h. Colonies were picked and DNA was extracted from randomly selected clones and PCR was conducted to confirm presence of correct construct. *A. tumefaciens* strain C58C1 was grown at 28°C in LB broth with 50 µg/ml kanamycin. Bacteria were centrifuged for 5 min at room temperature and re-suspended in 10 mM $MgCl_2$. Cells were brought to an optical density (OD_{600}) of ~0.5-0.6. Cells were then infiltrated into the abaxial air spaces of 3–4-week-old *N. benthamiana* plants using a syringe and monitored for phenotypes for 8 days post infiltration.

Phylogenetic analysis

Phylogenetic tree construction and bootstrap analysis was performed using RAxML (Randomized Axelerated Maximum Likelihood) software comparing *Arabidopsis* (AtBAG1: At5g52060, AtBAG2: At5g62100, AtBAG3: At5g07220, AtBAG4: At3g51780, AtBAG5: At1g12060, AtBAG6: At2g46240 and AtBAG7: At5g62390), rice (OsBAG1: LOC_Os09g35630, OsBAG2: LOC_Os08g43270, OsBAG3: LOC_Os06g03640, OsBAG4: LOC_Os01g61500, OsBAG5: LOC_Os11g31060 and OsBAG6: LOC_Os02g48780) and soybean (GmBAG6A: JX665043 and GmBAG6B Glyma16g03320) BAG protein sequences.

RESULTS

Cloning and sequence analysis of *GmBAG6*

In a previous study, laser capture microdissection was coupled with microarray profiling to compare gene expression profiles of SCN feeding cells induced in resistant and susceptible near-isogenic lines (NILs) differing only for the *Rhg1* locus on chromosome 18 of soybean. The most highly upregulated gene was found to be 87-fold higher in the resistant line. A nucleotide blast search to the soybean genome (cv. Williams 82) database (http://phytozome.net; (Goodstein et al., 2012) using the consensus sequences for the corresponding Affymetrix probeset Gma.7623.1.A1 (Kandoth et al., 2011) identified two best hits, Glyma07g06750 on chromosome 7 and Glyma16g03320 on chromosome 16 (Figure 2.1). A subsequent blast search to the TAIR protein database using these two soybean protein sequences identified AtBAG6 (At2g46240) as the most closely related *Arabidopsis* protein (50% sequence similarity). Correspondingly, these two soybean genes were named *GmBAG6A* (Glyma07g06750) and *GmBAG6B* (Glyma16g03320). The predicted protein sequences for GmBAG6A and GmBAG6B shared a similar domain arrangement with AtBAG6, including conserved IQ and BAG domains (Figure 2.2, Figure A1). A phylogenetic analysis including all *Arabidopsis* and rice BAG protein family members clustered the GmBAG6 proteins with AtBAG6 (Figure 2.3).

To clone the full-length *GmBAG6A* and *GmBAG6B* cDNA sequences, primers were designed based on the Williams 82 genome sequence and used in

PCR to amplify the corresponding sequences from cDNA generated from total RNA isolated from nematode-infected root tissue of Williams 82, NIL-7923R and NIL-7923S at 6 days post-inoculation. A fragment of approximately 3.3 kb was amplified and cloned into a gateway-cloning vector, pDONR-Zeo. All clones sequenced matched to the *GmBAG6A* sequence. No *GmBAG6B* clones could be recovered from nematode-infected root tissues in repeated attempts. These data suggest that *GmBAG6A*, but not *GmBAG6B* is upregulated in response to nematode infection. A soybean whole transcriptome analysis shows that *GmBAG6A* is predominately expressed in roots, shoot apical meristem, and green pods, whereas *GmBAG6B* was only detectable in leaves (Libault et al., 2010) (Figure 2.4). The sequences obtained from NIL-7923R and NIL-7923S with Williams 82 show a discrepancy between the predicted numbers of introns within the genomic sequence of *GmBAG6A.* The predicted gene model found on Phytozome.net predicts 6 intron splice sites, where as sequencing of the cDNA for *GmBAG6A* showed the presence of only one, 471 nucleotide intron, spliced out starting at nucleotide 1824 and ending at nucleotide 2295 (Figure A2). This splicing pattern was consistent between all three lines sequenced. The full-length 3,387-bp *GmBAG6A* cDNA sequence from Williams 82, NIL-7923R and NIL-7923S contained an open reading frame encoding a 1129 amino acid protein with a predicted mass of 126.47 kD. As expected, the *GmBAG6A* sequences from NIL-7923R and NIL-7923S shared 100% sequence conservation. An alignment of *GmBAG6A* sequences from NIL-7923R and NIL-7923S with that of

Williams 82 identified 14 single nucleotide polymorphisms resulting in 11 amino acid changes. (Figure A3, Figure A4).

Yeast cell death assay

The BAG domain of the human BAG-1 protein was shown to bind to BCL2 and promote cell survival (Takayama et al., 1995). In contrast, AtBAG6 was shown to have the opposite effect and induce apoptosis in yeast. Unlike human BAGs, a subset of plant BAGs, including AtBAG6, contain an IQ domain immediately upstream of a BAG domain. The AtBAG6 cell death phenotype was found to be more pronounced when a truncated version of the protein spanning the IQ and BAG domains was expressed in yeast (Kang et al., 2006). To investigate whether *GmBAG6A* functions similarly to *AtBAG6* in causing cell death in yeast, the full-length *GmBAG6A* and a truncated version of *GmBAG6A* spanning the IQBAG domain were expressed in a *S. cerevisiae* W303-1A yeast strain (Veal et al., 2003). For these experiments, sequences corresponding to the full-length *GmBAG6A* and the *GmBAG6A IQBAG* domain were cloned into the gateway compatible pYES2-Dest52 vector under control of the *GAL1* promoter, which allows for the conditional expression of this protein when cells are plated on a media containing galactose (Kang et al., 2006). The full-length *AtBAG6* and *AtIQBAG* sequences were cloned into pYES2-Dest52 and included as positive controls. Similar to *AtBAG6* and *AtIQBAG*, both *GmBAG6A* and *GmIQBAG* induced cell death in yeast. The cell death phenotype was more pronounced with the IQBAG domains (Figure 2.5). Unfortunately, the activity of the *BAG6* and

IQBAG proteins was partially disrupted when tagged with a 6XHIS tag on either the C and N terminus, with the C-terminal tag causing the most significant inhibition of cell death (Figure 2.5). Consequently, tagged versions of these proteins were not used in subsequent experiments.

Identification of *AtBAG6* mutants

To further test whether *GmBAG6A* had a similar function to *AtBAG6*, we wanted to test *GmBAG6A* for complementation of *AtBAG6* mutant phenotypes. However, two prior studies reported contradictory results for *atbag6* mutant phenotypes. One study reported that *atbag6* T-DNA insertion mutants displayed reduced vegetative growth, earlier flowering, and more branched roots and inflorescences (Doukhanina et al., 2006). SALK_009534 and SALK_058290 were used in the study by Doukhanina et al (2006). A second study reported no phenotypes in two different *atbag6* T-DNA insertion mutants exposed to a variety of different growth conditions including high and low temperature, high salt and UV light (Kang et al., 2006), although, the line names were not mentioned. We requested two *atbag6* T-DNA insertion lines, SALK_047959C (*bag6-1)* and SALK_073331C (*bag6-2)* (Figure 2.6A) from ABRC. Individual plants were genotyped to identify plants homozygous for the insertion. Both mutants were determined to not produce full-length transcript by qRT-PCR (Figure 2.6B). No growth phenotypes were observed in either of the mutant plants. These results were consistent with that reported by Kang et al (2006). Consequently, the

atbag6 mutants could not be used in complementation analyses to test the function of *GmBAG6A*.

The *atbag6* mutant plants were also tested for a nematode infection phenotype. No significant differences in nematode development were observed between wild-type Col-0 controls and *bag6-1* mutant plants (Figure 2.7).

Expression of *AtBAG6* and *GmBAG6* in Arabidopsis

To further characterize the function of *GmBAG6A* in plants, *GmBAG6A* and *GmIQBAG* were cloned into the pAKK binary vector under the control of a constitutive cauliflower mosaic virus double 35S promoter (2X35S) and transformed into *Arabidopsis*. *AtBAG6* and *AtIQBAG* were cloned into the pAKK binary vector under control of a 2X35S promoter and transformed into *Arabidopsis* for comparison. Out of 53 T_1 primary transformants for *AtBAG6*, with plants derived from 12 independent pools of seed, 77% of the plants were indistinguishable from wild-type control plants (Figure 2.8, A). Only 23% of the plants screened showed any observable phenotype and were classified as having either an intermediate phenotype (15%) represented by plants with smaller rosettes, leaf lesions, and malformed leaves or a severe phenotype (8%) represented by dwarf plants showing anthocyanin accumulation (Table 1; Figure 2.8). Out of 105 T_1 primary transformants for *GmBAG6A*, with plants derived from 9 independent pools of seed, 95% of the plants were indistinguishable from wild-type control plants (Figure 2.4, A). Only 5% of the plants screened showed any observable phenotype and were classified as having either an intermediate

phenotype (5%) represented by plants with smaller rosettes, leaf lesions, and malformed leaves or a severe phenotype (0%) represented by dwarf plants showing anthocyanin accumulation (Table 1: Figure 2.9). In contrast to the full-length *AtBAG6*, 67% of plants expressing *AtIQBAG* had an observable phenotype. Of the 133 primary T_1 transformants observed, with plants derived from 9 independent pools of seed, 21% were classified as an intermediate phenotype and 46% of the plants displayed a severe phenotype exhibiting plant death, necrotic lesions, misshapen leaves, anthocyanin accumulation, shoot apical meristem termination, delayed flowering and aborted siliques (Table 2; Figure 2.8). In contrast, only 10% of the plants expressing *GmIQBAG* had any observable phenotype. Of the 128 primary T_1 transformants observed, with plants derived from 9 independent pools of seed, 8% were classified as intermediate and only 2% as severe (Table 2, Figure 2.9).

Overexpression of the transgenes was confirmed via qRT-PCR using sequence-specific primers and RNA extracted from rosette leaves of 14 day-old T_1 plants. All lines tested showed an increase in transcript abundance compared to wild-type control plants and the level of expression correlated with the severity of symptoms observed in the plants tested (Figure 2.10). Intriguingly, the expression of the full-length *AtBAG6* and *GmBAG6A* in the overexpression lines was much lower (Figure 2.11, A-C) than the expression level observed for the *AtIQBAG* and *GmIQBAG* (Figure 2.11, B-D).

Expression of *GmIQBAG* in soybean using *Bean pod mottle virus*

A 312-bp fragment (bps 1597-1908) of the *GmBAG6A* cDNA sequence (Genbank accession number JX665043) spanning the IQBAG domain was amplified by post-doctoral research associate, Dr. Pramod Kandoth, from SCN-resistant soybean NIL-7923R root cDNA and cloned into *pBPMV-IA-V4* (Figure 2.11, A). The full-length *GmBAG6A* of 3,387-bp exceeded the size limit for recombinant virus expression and therefore was unable to be tested using this approach. The SCN-resistant NIL-7923R was inoculated with *pBPMV-IA-IQBAGOE* while control plants were infected with BPMV only and observed for phenotypes. Plants inoculated with *pBPMV-IA-IQBAGOE* showed severe stunting in growth and developed sporatic necrotic lesions on developing leaves (Figure 2.11, B-C) when compared with controls. These data indicated that *GmIQBAG* can induce cell death in soybean.

Transient expression of *BAG6* in tobacco by agroinfiltration

The *Arabidopsis* and soybean BAG6 proteins were also tested for their ability to induce cell death in tobacco. Four constructs, *GmBAG6A, AtBAG6, AtIQBAG* and *GmIQ*BAG were cloned into the pAKK binary vector under the control of a constitutive cauliflower mosaic virus double 35S promoter and transformed into *Agrobacterium tumefaciens* strain C58C1 for transient expression in *Nicotiana benthamiana* via syringe infiltration. C58C1 transformed

with the empty vector was used a negative control. Co-infiltration of *RX,* a *NB-LRR* protein of potato and the coat protein (CP) from potato virus X (PVX) (Bendahmane et al., 1999) was used as a positive control for a hypersensitive response (HR). Agroinfiltrated plants were monitored for the formation of an HR in leaf sections infiltrated with the constructs for 8 days. By day 8, there was no visible HR in any construct tested except for the positive control (Figure 2.12).

Cloning of nematode-inducible promoters to drive expression of *IQBAG*

The IQBAG domain of AtBAG6 and GmBAG6 was sufficient to cause cell death in yeast and in plants. Therefore, we wanted to test whether this protein could be used to engineer a novel form of resistance to cyst nematodes. However, in order to test this, a nematode-inducible promoter would need to be employed to avoid the severe plant growth defects associated with constitutive overexpression of IQBAG in plants. For this, our previously published LCM-microarray dataset was used to identify genes upregulated in response to nematode infection within feeding cells. Two probesets, GmaAffx.88182.1.S1_at and Gma.11004.1.S1_at showing an 8-fold upregulation in syncytia of the resistant soybean line (NIL-R) vs. the susceptible soybean line (NIL-S**)** (Kandoth et al., 2011) and corresponding to consensus sequences annotated as *NDR1/HIN1-like* (*NHL*) genes were selected for further analysis. These two genes were named *NHL1* (Glyma03g35930) and *NHL2* (Glyma03g35920). Primers were designed to the consensus sequences and used by post-doctoral research associate Xiaolei Wu to screen a soybean BAC library. BAC M0116A09

was identified to contain both genes and was fully sequenced. The *NHL1* and *NHL2* genes were found to reside within a cluster of 8 other *NHL* genes residing within a 30 KB region of the BAC. A PCR-based approach was used to amplify 1 kb of *GmNHL1* promoter region and 1.98 kb of *GmNHL2* promoter region from BAC M0116A09. These sequences were cloned upstream of the β-glucuronidase (*GUS*) reporter gene to generate *NHL1-GUS* and *NHL2-GUS* constructs for further analysis (Figure 2.13). The *GmNHL1* and *GmNHL2* promoter sequences from NIL-7923R and NIL-7923S were also cloned and sequenced and found to be identical to each other and to the Williams 82 sequences (data not shown).

Analysis of soybean *NHL* gene promoters during plant development and nematode infection

To determine where the *GmNHL1* and *GmNHL2* genes were expressed during plant growth and development, we analyzed the tissue-specific and cell-type expression pattern of the *NHL1-GUS* and *NHL2-GUS* genes in *Arabidopsis* and soybean roots. Transgenic *Arabidopsis* and soybean hairy roots that contained *NHL1-GUS* and *NHL2-GUS* fusions were generated. Single-insertion, homozygous *Arabidopsis* lines were generated for each construct. Organ-specific regulation of *NHL1-GUS* and *NHL2-GUS* was characterized in transgenic *Arabidopsis* during plant development. In *NHL1-GUS* lines, expression was undetectable in 7-day old seedlings and barely detected in the vasculature of developing leaves at 14 days post germination (Figure 2.14, A-C). *NHL1-GUS*

expression increased in 21-day-old rosettes in a sporadic pattern (Figure 2.14, D). Expression of *NHL1-GUS* was also observed in cauline leaves in a punctate pattern (Figure 2.14, E). *NHL1-GUS* expression was undetected in flowers, siliques, and roots (Figure 2.14, F-H). In 7-day-old seedlings, *NHL2-GUS* expression was observed in the root tips, and a sporadic expression pattern was detected in the leaves, emerging laterals and root vasculature (Figure 2.15, A-C). In 14 day-old seedlings, expression of *NHL2-GUS* was observed in the leaf vasculature (Figure 2.15, D-E). In 21 day-old plants, strong expression of *NHL2-GUS* was seen within select rosette leaves and appeared to have a 'leaky' expression profile (Figure 2.15, F). *NHL2-GUS* was also strongly expressed in the vasculature of cauline leaves (Figure 2.15, G). During *Arabidopsis* reproductive development, *NHL2-GUS* was strongly expressed in nodes and flowers (Figure 2.15, H-I), Expression was seen in the base and tips of siliques (Figure 2.15, J). Multiple independent soybean hairy root lines were generated and analyzed for each construct. *NHL1-GUS* expression was undetectable in transgenic soybean hairy roots (Figure 2.16, A). *NHL2-GUS* expression in soybean hairy roots was only detected in root tips and lateral root initials (Figure 2.16, B-C).

We then examined the spatial and temporal expression pattern of *GmNHL1* and *GmNHL2* in response to nematode infection. Multiple independent transgenic soybean hairy root lines were generated for each construct and infected with SCN. Roots were stained at different timepoints 2, 5, 10 and 17 days following inoculation. *NHL1-GUS* expression was specifically expressed in

the area of nematode feeding (Figure 2.17). This expression pattern was observed in roots fed upon by parasitic second-stage juveniles (Figure 2.17, A-C) and was still observed in roots fed upon by fourth-stage females (Figure 2.17, D). Longitudinal sections through transgenic soybean hairy roots infected with SCN indicated that *NHL1-GUS* was specifically expressed within developing syncytia (Figure 2.17, C). Similarly, *NHL2-GUS* expression was also observed in and around nematode feeding sites (Figure 2.18). Longitudinal sections through soybean roots infected with SCN confirmed expression within syncytia and surrounding cells (Figure 2.18, C). Transgenic *Arabidopsis* that contained *NHL1-GUS* and *NHL2-GUS* fusions were also infected with *Heterodera schachtii*, the beet cyst nematode (BCN). Similar to that observed in soybean, *NHL1-GUS* and *NHL2-GUS* were induced in syncytia formed by BCN upon infection of *Arabidopsis*. (Figure 2.19).

The *NHL1-GUS* and *NHL2-GUS* expression patterns are consistent with the LCM microarray data (Kandoth et al., 2011) indicating that *GmNHL1* and *GmNHL2* are expressed in nematode-induced syncytia.

Nematode infection assays of *Arabidopsis* and soybean expressing IQBAG under control of the *GmNHL1* promoter

Based on the observation that the *NHL1* gene promoter had very low, if any, background expression in root tissues of *Arabidopsis* and soybean it was selected to drive the specific expression of IQBAG in nematode feeding sites to

assess for potential effects on nematode infection. For these experiments, *NHL1-AtIQBAG* and *NHL1-GmIQBAG* constructs were generated and transformed into *Arabidopsis* and soybean hairy roots, respectively. Multiple single-insertion, segregating T_2 *Arabidopsis* lines derived from independent transgenic events were generated for the *NHL1-AtIQBAG* construct. Primary T_1 transformants, as well as segregating T_2 transgenic seedlings grown on soil, began exhibiting sporadic lesion formation on rosette leaves by 14 days (Figure 2.20). The pattern of lesion formation was consistent with the sporadic *NHL1-GUS* expression observed in leaves during *Arabidopsis* development (Figure 2.14, C-D). However, root growth of the *NHL1-AtIQBAG* lines remained unaffected (Figure 2.21) consistent with the lack of observable *NHL1-GUS* expression in *Arabidopsis* roots (Figure 2.14, H). Multiple, independent, transgenic soybean hairy root lines were also generated for the *NHL1-GmIQBAG* construct. No observable root growth phenotypes were detected, which is consistent with the low level of *NHL1-GUS* expression observed in soybean hairy roots (Figure 2.16).

We next examined nematode development on the *NHL1-AtIQBAG* *Arabidopsis* lines and *NHL1-GmIQBAG* soybean hairy root lines. *Arabidopsis* was infected with BCN and the soybean hairy roots were infected with SCN. For the *Arabidopsis* infection assays, two segregating T_2 *NHL1-AtIQBAG* lines (3-1 and 6-4) and wild-type Col-0 were randomized in 12-well plates and grown on modified Knop's medium. Ten days after germination, seedlings were inoculated with infective J2s. At this stage of development, no sporadic lesions were

observed on the foliar portion of the plants. However, by day 14, sporadic lesions began to appear on rosette leaves. Thus, we counted nematodes at both 9 and 14 days post-inoculation (dpi) to minimize potential effects of compromised plant growth and development on nematode development. J3 nematodes were counted at 9 dpi and J4 females were counted at 14 dpi. Both *NHL1-AtIQBAG* lines showed a statistically significant reduction in nematode infection compared to wild-type controls at both timepoints analyzed in three biological replicates of the experiment (Figure 2.22).

For nematode infection of soybean hairy root lines, at least 15 independent hairy roots lines were generated for the *NHL1-GmIQBAG* and the empty vector constructs and inoculated with infective J2s. Cysts were counted at 30 dpi. The *NHL1-GmIQBAG* lines showed a statistically significant reduction in nematode infection compared to empty vector control roots in three biological replicates of the experiment. At 30 dpi, nematode infection was reduced by approximately 45-70% in *NHL1-GmIQBAG* hairy roots (Figure 2.23; $P<0.001$). Taken together, these data indicate that expression of the IQBAG domain specifically within feeding cells in response to nematode infection is detrimental to nematode growth and development.

DISCUSSION

In a prior study, we identified a soybean gene sequence that was expressed 87-fold higher in SCN-induced feeding cells of resistant soybean compared to susceptible soybean (Kandoth et al., 2011). This sequence shared homology with the *Arabidopsis* gene, *AtBAG6* (At2g46240), which was shown to induce programmed cell death (PCD) in yeast and plants (Kang et al, 2006). Further analysis of the completed soybean genome sequence identified two soybean genes sharing similarity with *AtBAG6* and were named *GmBAG6A* (Glyma07g06750) and *GmBAG6B* (Glyma16g03320). We were unable to clone *GmBAG6B* from nematode-infected root tissues, indicating that this gene is unlikely induced upon nematode infection, and therefore we focused our studies on *GmBAG6A*. The predicted GmBAG6A protein was similar in size and domain structure to that of AtBAG6. Like AtBAG6, GmBAG6A contained an IQ motif immediately upstream of a conserved BAG domain, with the highest sequence similarity between the proteins spanning the region containing the IQ and BAG domains.

To date, no soybean BAG proteins have been characterized. A previous study demonstrated that AtBAG6 is a CaM-binding protein involved in plant PCD that is upregulated upon various stresses (Kang et al., 2006). In yeast and *Arabidopsis*, overexpression of *AtBAG6* induced PCD. Moreover, a 134-amino-acid sequence, including both the IQ motif and BAG domain, was sufficient to induce cell death. The IQ motif was shown to be required for Ca^{2+} independent binding of BAG6 to CAM and is required for cell death. As a first test of

GmBAG6A function, the cDNA encoding full-length *GmBAG6A* was subcloned under control of the *GAL1* promoter in the pYES2 vector to allow for the inducible expression of this protein when cells are grown on medium containing galactose. Similar to *AtBAG6*, we found that expression of *GmBAG6A* in yeast cells induced cell death. In addition, expression of the *GmIQBAG* domain was sufficient for cell death and an even more robust cell death phenotype was observed with the truncated form of the protein. These data indicated that *GmBAG6A* functions similarly to *AtBAG6* in causing PCD in yeast cells.

To test *GmBAG6A* for functional complementation of *atbag6* loss-of-function mutants, we obtained two T-DNA insertions lines (SALK_047959C; SALK_073331C) in *AtBAG6* from *Arabidopsis* Biological Resource Center and advanced these to homozygosity. Neither mutant produced full-length transcript. Contradictory reports of *atbag6* mutant phenotypes have been reported previously. Doukhanina et al (2006) reported that *atbag6* mutants exhibited earlier flowering and reduced vegetative growth. In addition, the plants produced more branched roots and shoots. However, Kang et al (2006) did not observe any phenotypes in *atbag6* mutants under a variety of different stress conditions. We also did not observe any phenotypes in the *atbag6* mutants, including response to nematode infection, consistent with the report by Kang et al (2006) so complementation analysis using *GmBAG6A* was not possible.

To further analyze the function of *GmBAG6A*, overexpression studies were conducted in *Arabidopsis*, soybean, and tobacco. For a direct comparison, *AtBAG6* was included in these studies. Sequences encoding full-length or

IQBAG domains of *GmBAG6A* and *AtBAG6* were subcloned under control of the double 35S promoter in the pAKK binary vector. Transgenic plants were screened for dwarfism and induction of disease-like necrotic lesions, which were previously observed in the *AtBAG6* overexpression plants (Kang et al., 2006). Interestingly, lines harboring the *AtBAG6* overexpression construct showed a low frequency (23%) of only minor phenotypes such as, smaller rosettes, misshapen and serrated leaves and delayed flowering, which is contrary to the severe dwarf phenotypes and observed necrotic leaf lesions reported by Kang et al (2006). Out of 53 T_1 primary transformants screened, 77% were indistinguishable from wild-type controls. However, the phenotypes of plants harboring the *AtIQBAG* overexpression construct were consistent with the report by Kang et al (2006). Of the 133 T_1 primary transformants screened, 67% displayed severe symptoms including plant death, necrotic leaf lesions, misshapen leaves, anthocyanin accumulation, shoot apical meristem termination, delayed flowering and aborted siliques. The difference in the severity of the cell death phenotypes in *Arabidopsis* expressing the full-length vs. the IQBAG domain, is consistent with the different severity of cell death phenotypes observed with the full-length and IQBAG domain constructs in yeast.

In contrast to *AtBAG6*, the frequency of plants displaying observable phenotypes in *Arabidopsis* overexpressing *GmBAG6A* or *GmIQBAG* was relatively low (5-10%). Out of the 233 T_1 primary transformants screened for both the *GmBAG6A* full-length and *GmIQBAG* constructs, 90-95% of the plants were indistinguishable from wild-type controls. Transgene expression was

confirmed in all overexpression lines via qRT-PCR. All lines tested showed an increase in transcript abundance that correlated with symptom severity; however, much higher levels of transgene expression were observed in the *IQBAG* overexpression lines compared to the full-length *BAG6* overexpression lines. It has been reported that low constitutive levels and intermediate constitutive levels of *AtBAG4* expression in tobacco plants correlated with tolerance to abiotic stresses (Doukhanina et al., 2006). It was also recently reported that low concentrations of the human BAG1M activate the refolding activity of Hsp70/Hsc70 under physiological conditions, whereas high concentrations of BAG1M inhibited the refolding activity of their chaperones. This suggests that the concentration of a given BAG protein in the cell relative to the concentration of its binding partner may be critical for optimal activity (Gassler et al., 2001). Thus, the lower expression of the full-length *BAG6* constructs, suggests that elements regulating *BAG6* expression are present on the C and N-terminus of the full-length protein, which may function in negative feedback to regulate *BAG6* expression level. In contrast, the penetrating phenotype of the AtIQBAG domain may be due to the lack of the C and N-terminus regulatory activity. Consequently, the high levels of IQBAG protein in the cell may result in non-specific binding with a number of other calmodulins causing a detrimental disruption in the cells calmodulin signaling pathways, leading to apoptosis. Interestingly, despite the similar cell death phenotypes of *AtIQBAG* and *GmIQBAG* in yeast cells, only *AtIQBAG* caused noticeable cell death phenotypes when overexpressed in *Arabidopsis*. The cause for this difference might be due to species-specific

interactions of plant BAG6 proteins with binding partners. This idea is supported by the results of overexpression of *GmIQBAG* in soybean using BPMV. Soybean plants overexpressing *GmIQBAG* were severely stunted and developed sporadic necrotic lesions on newly formed leaves. These phenotypes are consistent with the phenotypes observed in *Arabidopsis* lines overexpressing *AtIQBAG*. In addition, transient expression of both full-length *BAG6* and *IQBAG* proteins from *Arabidopsis* and soybean failed to elicit a hypersensitive response in tobacco. Taken together, these results support the idea that soybean and *Arabidopsis* BAG6 proteins function most efficiently in their host systems.

GmBAG6A is highly expressed in feeding cells induced in resistant soybean cultivars upon SCN infection. Although the role of *GmBAG6A* in SCN resistance is still not fully understood, these data indicate that it may be contributing to the HR-like cell death response elicited in response to nematode infection in resistant cultivars. One possibility is that *GmBAG6A* expression is misregulated in feeding cells of the resistant cultivar, thereby exceeding the expression threshold, and this leads to a hypersensitive response. Further research will be needed to fully establish a role for *GmBAG6A* in nematode resistance.

The robust cell death phenotype elicited by overexpression of either the *AtIQBAG* domain in *Arabidopsis* or the *GmIQBAG* domain in soybean led us to test its potential use in engineering a novel form of resistance to nematodes. This is based on the fact that cyst nematodes requires the correct formation of a syncytium in order to survive. Therefore, if we can successfully engineer plants

that specifically target and terminate syncytium formation, then we may be able to effectively engineer plants resistant to cyst nematodes. For this work, the nematode-inducible promoter *GmNHL1*, which was found to be induced in syncytia during nematode infection, and had very low, if any, background expression in root tissues in the absence of nematode infection, was selected for targeted *IQBAG* expression studies. Sequences encoding IQBAG domains of *AtBAG6* and *GmBAG6A* were subcloned under control of the *GmNHL1* promoter in the pAKK binary vector and transformed into *Arabidopsis* and soybean hairy roots, respectively. *Arabidopsis* plants expressing *NHL1-AtIQBAG* grew similar to wild-type plants. However, sporadic necrotic leaf lesions were observed on rosette leaves consistent with the sporadic GUS expression pattern observed in leaves of the transgenic *Arabidopsis NHL1-GUS* lines. Root growth of the *Arabidopsis NHL1-AtIQBAG* lines and soybean *NHL1-GmIQBAG* hairy root lines was unaffected, consistent with the lack of *NHL1-GUS* expression in root tissues. Transgenic lines were infected with either BCN (*Arabidopsis*) or SCN (soybean hairy roots) and nematode development was measured and compared with that of nematodes developing on the control lines. In nematode infection assays of the transgenic *Arabidopsis* T_2 lines expressing *GmNHL1-AtIQBAG,* nematode development was measured at 9 and 14 days post inoculation. At both timepoints, nematode development was reduced by 35-50%. Similar results were observed in soybean hairy roots generated from the SCN-susceptible soybean cv. Williams 82. On soybean expressing the *NHL1-GmIQBAG* construct, a 45-70% reduction in nematode development was measured compared with control

lines harboring an empty *NHL1* cassette. Taken together, these data indicate that the *NHL1-IQBAG* expression constructs are having a deleterious effect on normal nematode development most likely by disrupting the forming syncytia.

The results from this study indicate that the GmBAG6A protein can induce cell death in yeast and plants similar to AtBAG6. Moreover, we have demonstrated that specific expression of the IQBAG domain in nematode feeding sites could be a viable means of nematode control. Even though the promoter is not ideal for use in commercial lines due to its sporadic expression in the vegetative tissue, it nonetheless shows the potential benefits and efficacy in using *IQBAG* under the control of a more stringent nematode-inducible promoter. This type of transgenic approach contrasts with the specificity of resistance genes that are normally protective for only one or select nematode species. Combining natural and transgenic resistance shows promise in bestowing an additive effect in protecting soybeans against a broad range of plant-parasitic nematodes. The potential to use an endogenous soybean gene for nematode control substantiates further research. The benefits to agriculture include reduced crop loss and less dependence on environmentally toxic nematicides

ACKNOWLEDGMENTS

We thank Bob Heinz for maintaining nematode cultures, Joe Mercurio from the molecular cytology core for imaging assistance and Pat Edger for assistance with phylogenetic tree construction. Funding for this work was provided by Missouri Soybean Merchandising Council and the National Science Foundation Plant Genome Program.

TABLES

Table 1. Summary of *BAG6* overexpression phenotypes in *Arabidopsis*

	T_1 Shoot Phenotypes[a]				
Construct	WT-like (%)	Intermediate (%)	Severe (%)	T_1 lines (no.)	Total T_1 (no.)
2X35S:AtBAG6	77	15	8	12	53
2X35S:GmBAG6A	95	5	0	9	105

[a]3 week-old seedlings; WT-like = indistinguishable from wildtype; intermediate = smaller rosettes, leaf lesions, malformed leaves; severe = dwarf, anthocyanin accumulation

Table 2. Summary of *IQBAG6* overexpression phenotypes in *Arabidopsis*

	T_1 Shoot Phenotypes[a]				
Construct	WT-like (%)	Intermediate (%)	Severe (%)	T_1 lines (no.)	Total T_1 (no.)
2X35S:AtIQBAG	32	21	46	9	133
2X35S:GmIQBAG	90	8	2	9	128

[a]3-week-old seedlings; WT-like = indistinguishable from wildtype; intermediate = smaller rosettes, leaf lesions, malformed leaves; severe = dwarf, anthocyanin accumulation, meristem termination

FIGURES

Affymetix Chip ID	Probeset	Glyma Number	e-value
Gma.7623.1.A1	gaagaggcaattatccctgatgataaagacacagaaaatttggccaaag agaaaactgaagtatctgcagaaccaccacctgcattgcaagaccgag ggttaaacggtgactcgaagttattagaagagaatgagaagttaaggga gatgatgaagaagttgcttgaagccgggaatgaacagttaagcgtgatat cagatttgactgtcagagtgaaggacttggagaagaaattagccaggag aaggagtaagagagtgaagacaaaacagtatagacccgcagcttcca aaatgtctacccatgaaatgaaatcctcctaatattccactccatgtgagag cctttgatgttgcaatgtaa	Glyma07g06750 Glyma16g03320	0 4.1e-110

Figure 2.1. Soybean Affymetrix Genechip ID corresponding to the upregulated *BAG* gene. Consensus probe sequence, best matches to the consensus sequence and associated e-values.

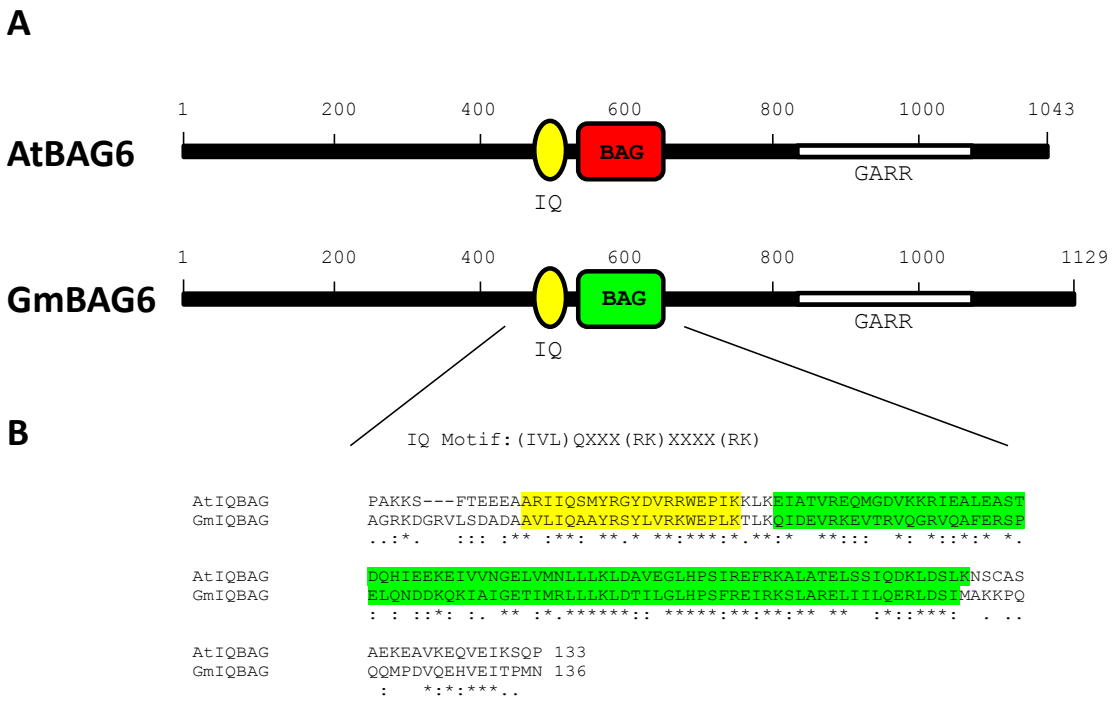


Figure 2.2. BAG6 protein domain structures. (A) Schematic representation of full-length *AtBAG6* (At2g46240) and *GmBAG6* (Glyma07g06750; Glyma16g03220). The predicted IQ domain (light grey), BAG domain (dark grey) and glutamic acid-rich regions (GARR) (white) are indicated. (B) Alignment of IQ and BAG domain sequences from AtBAG6 with GmBAG6A. Amino-acid residues representing the IQ domain (light grey) and BAG domain (dark grey).

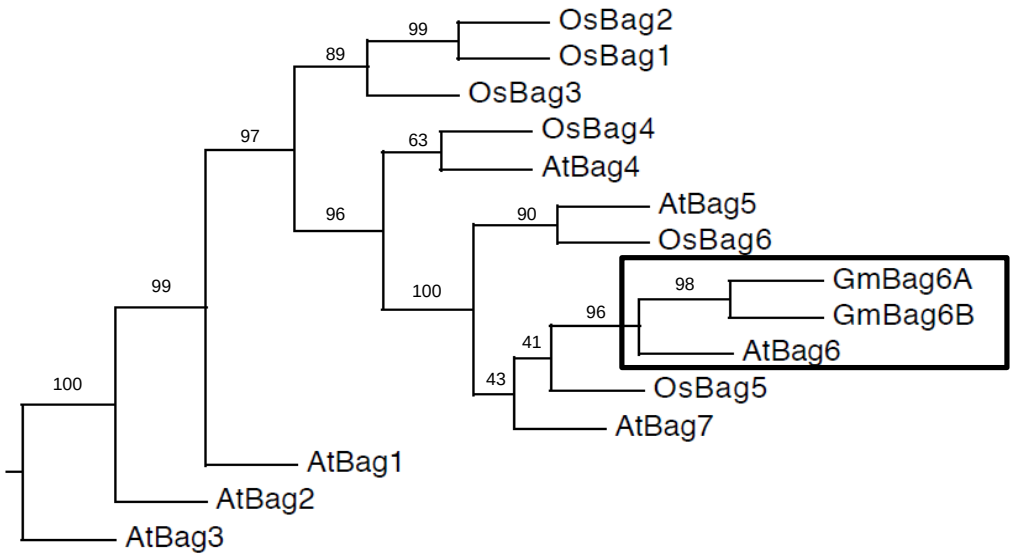


Figure 2.3. Phylogenetic tree construction and bootstrap analysis using RAxML (Randomized Axelerated Maximum Likelihood) software comparing *Arabidopsis*, rice and soybean BAG protein sequences. The numbers beside nodes are the percentages of bootstrap values calculated for 1000 replicates. The numbers represent the probability that each branch point is correct.

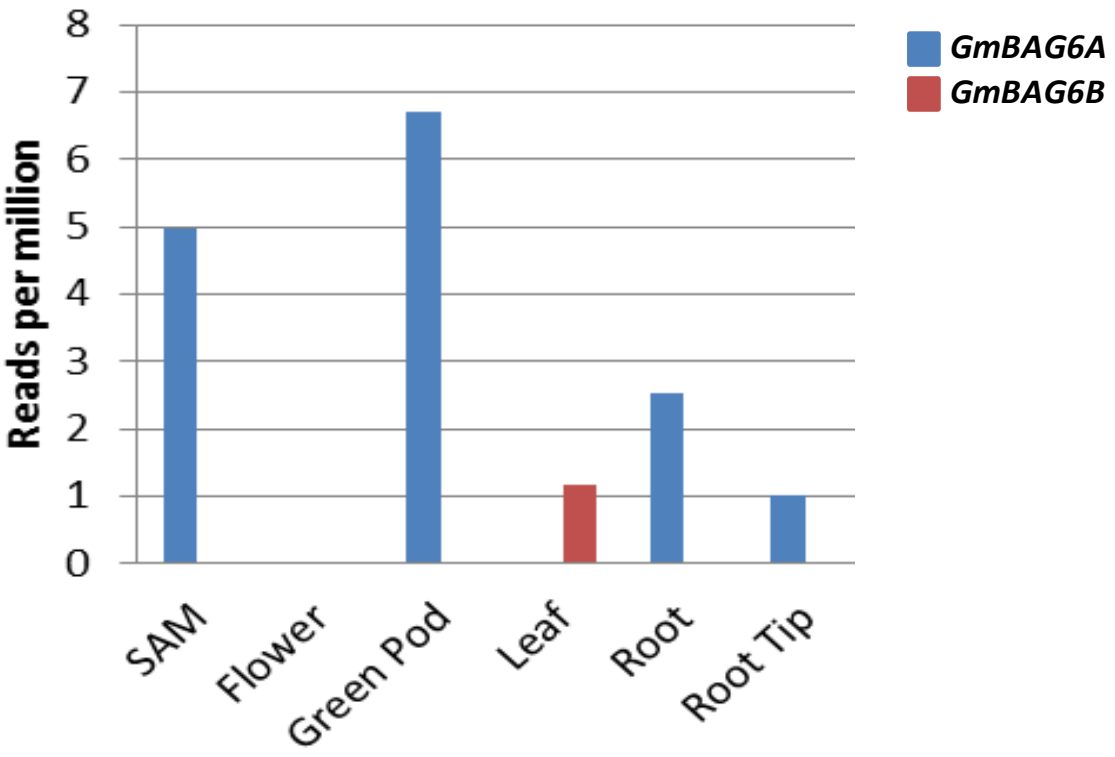


Figure 2.4. *GmBAG6* Illumina Solexa expression profiles in soybean. Data extracted from Libault et al (2010).

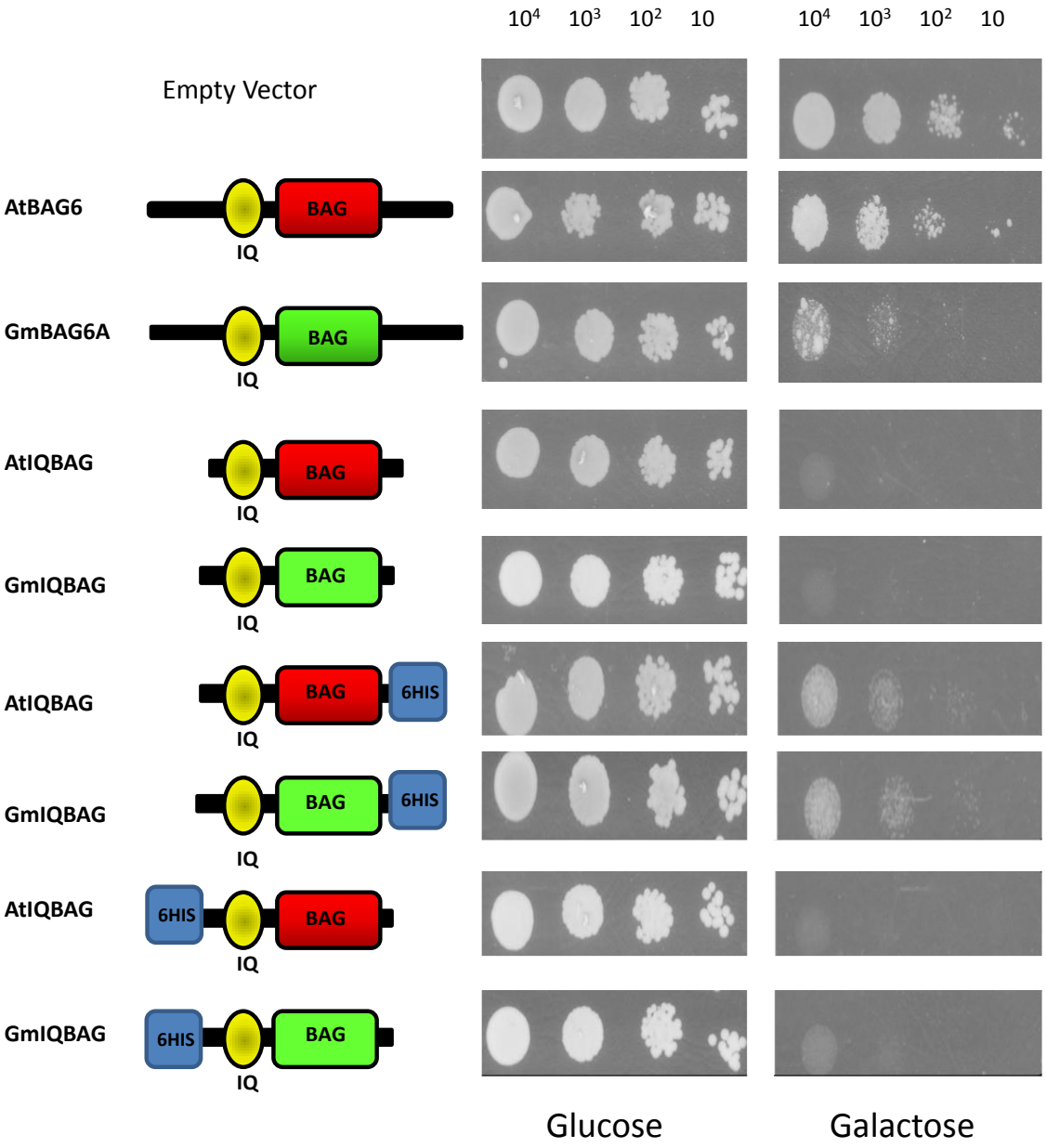


Figure 2.5. Functional analysis of soybean and *Arabidopsis* BAG proteins and IQBAG domains in yeast W303-1A cells. All constructs were cultured in SD-Uracil glucose-based medium to an OD600 of ~ .17. Equal numbers of cells were spotted on minimal SD-Uracil medium plates in the presence of glucose or galactose. Photographs were taken after culturing at 30^0C for 3 days.

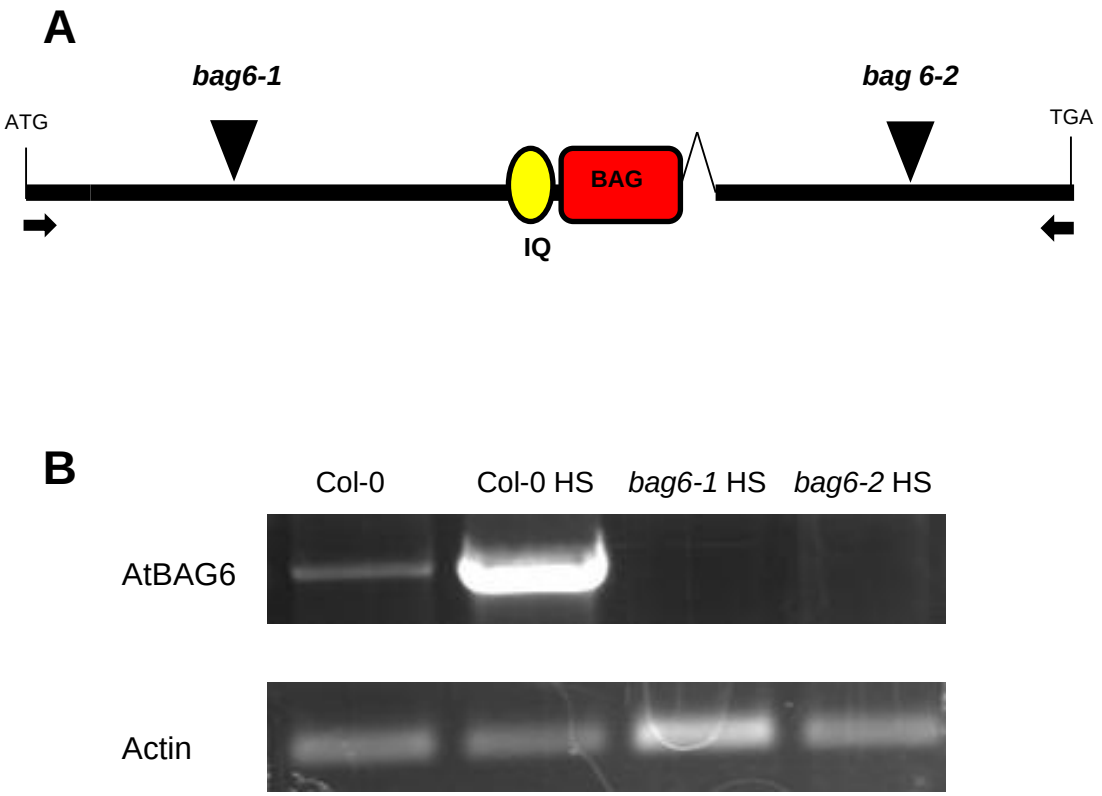


Figure 2.6. SALK *atbag6* mutant T-DNA insertion map and expression profile. (A) Schematic indicating relative position of T-DNA insertion in the *AtBAG6* gene for two *atbag6* mutant lines. (B) Top:PCR using *AtBAG6* primers (denoted by arrows in A) on cDNA generated from Col-0 wild type and heat-shocked (HS) tissues of Col-0, *bag6-1* and *bag6-2*. Bottom: Amplfication of Actin.

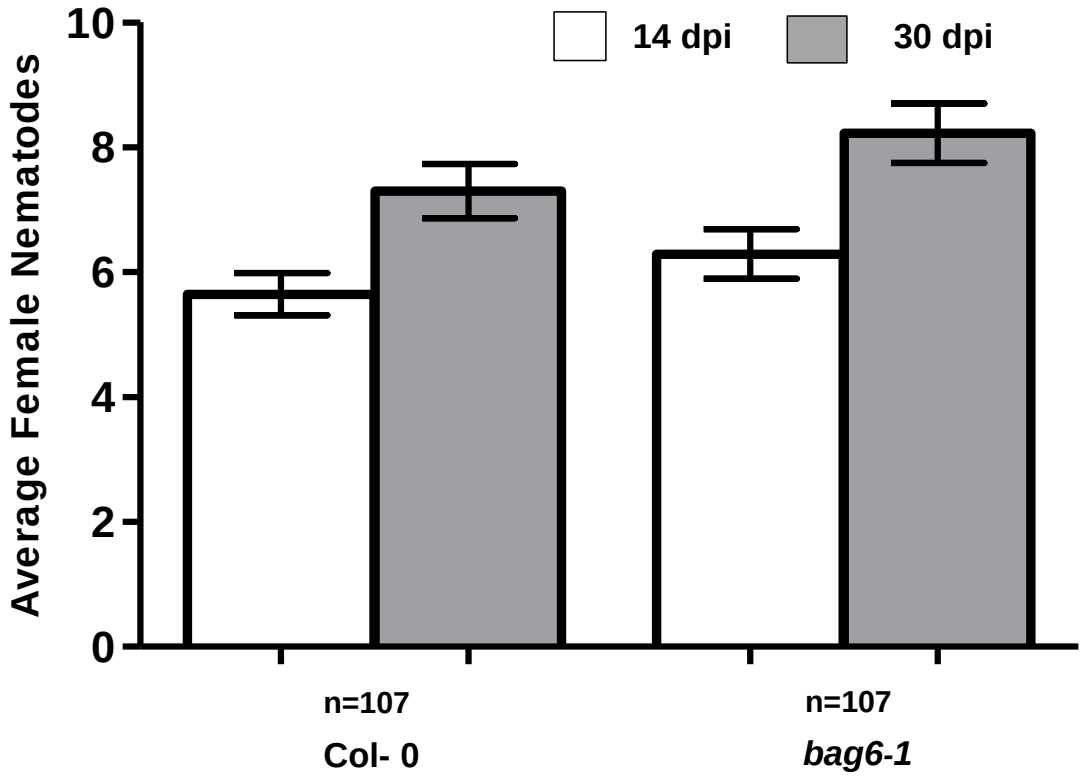


Figure 2.7. Nematode infection assay of *bag6* mutant. Individuals from *bag6-1* T-DNA knockout line and wildtype Columbia-0 were infected in 12-well plates with 175 infective J2 of *H. schachtii*. Fourth-stage juvenile (J4) females were counted at 14 days post-inoculation (dpi) and adult females were counted at 30 dpi. Data represent the mean average. n = sample size and error bars = SE. The graph represents the combined data from three biological replicates.

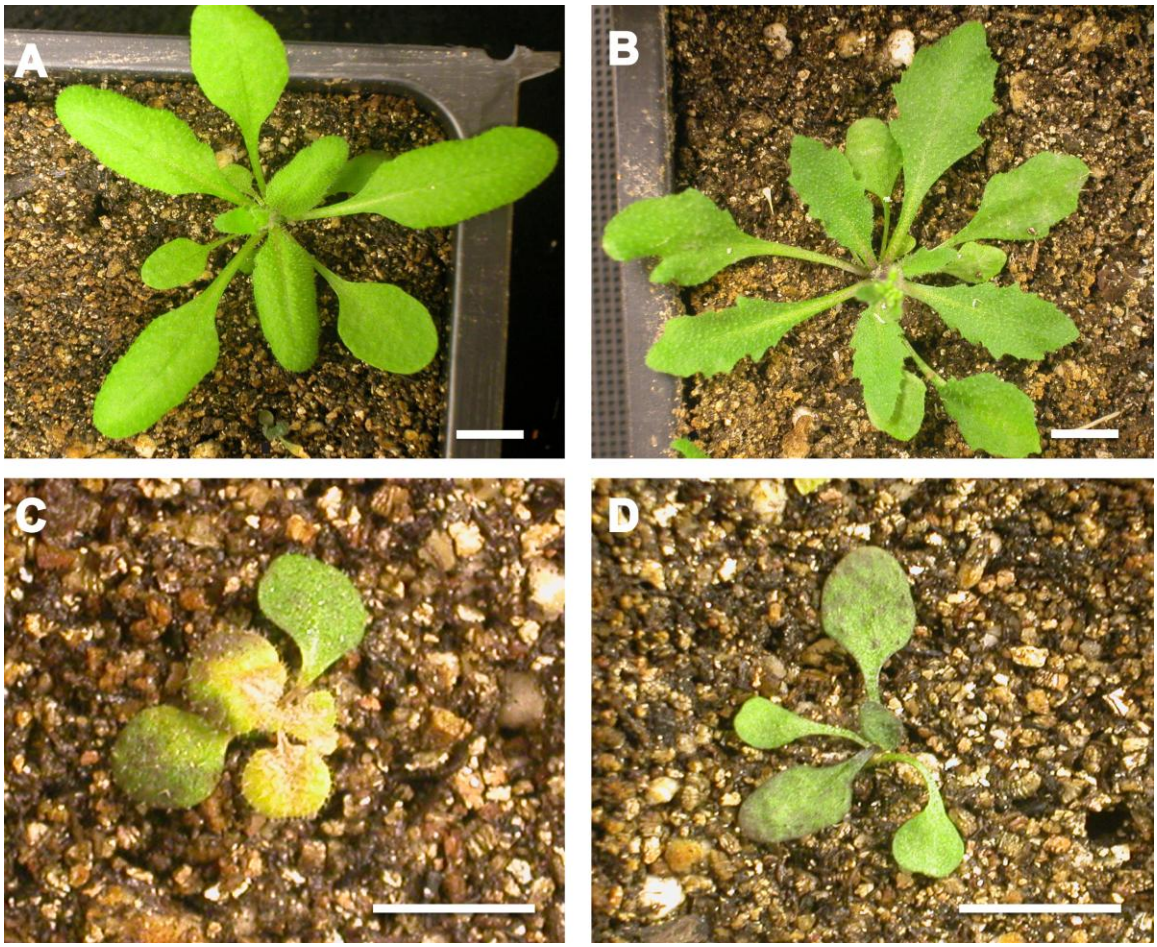


Figure 2.8. Transgenic *Arabidopsis* expressing full-length *AtBAG6* and *AtIQBAG*. (A) Wild-type Col-0 (B) Expression of the full-length AtBAG6 protein in *Arabidopsis* induces weak developmental phenotypes including slightly misshapen leaves and smaller rosettes. (C-D) Expression of the *AtIQBAG* domain in *Arabidopsis* induces severe growth defects including shoot apical meristem termination, anthocyanin accumulation, chlorosis, necrosis, and plant death. Scale bars =1 cm.

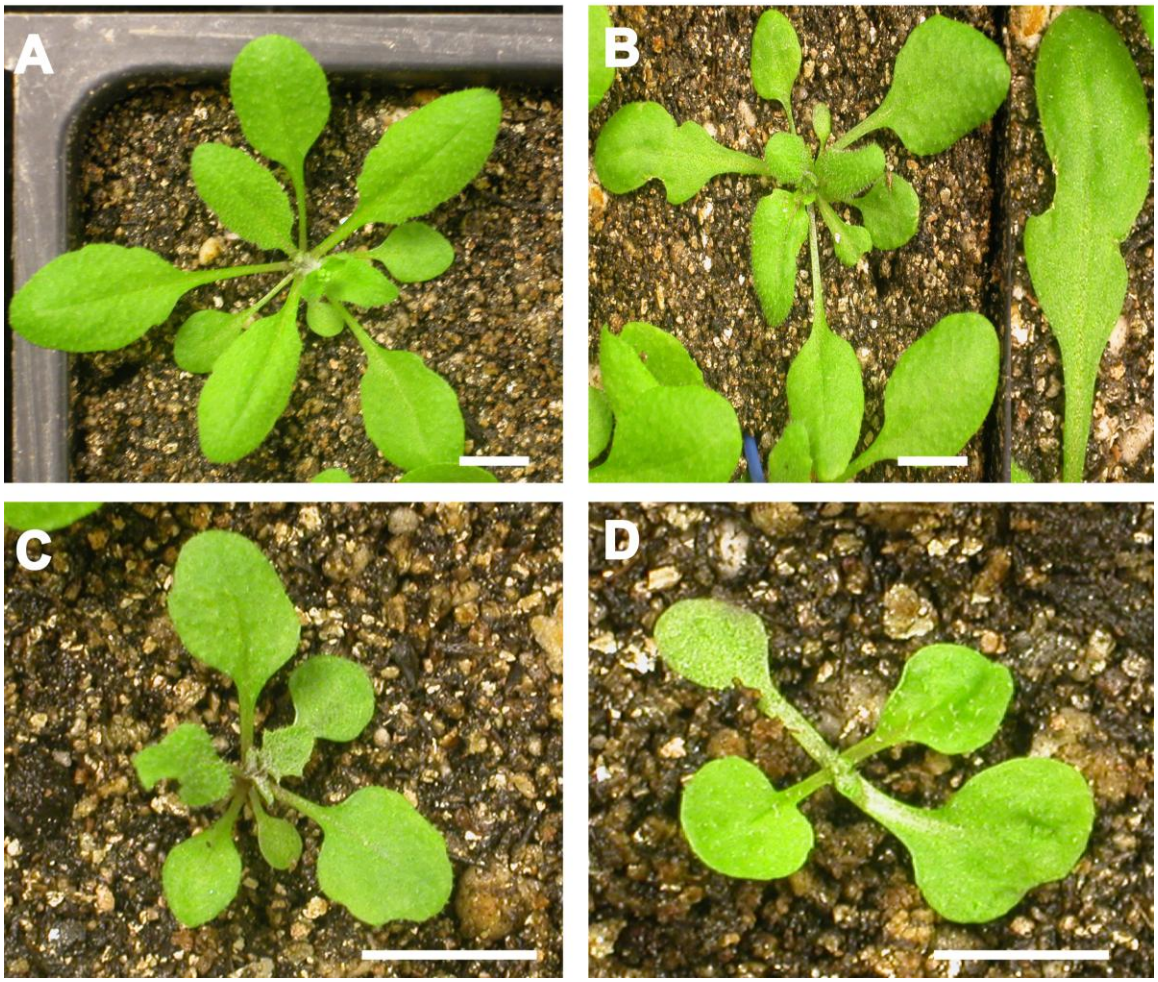


Figure 2.9. Transgenic *Arabidopsis* expressing full-length *GmBAG6A* and *GmIQBAG*. (A) Wild-type Col-0 (B) Expression of the full-length *GmBAG6A* protein in *Arabidopsis* induces weak developmental phenotypes including slightly misshapen leaves and smaller rosettes. (C-D) Expression of the *GmIQBAG* domain in *Arabidopsis* induces weak developmental phenotypes including slightly misshapen leaves and smaller rosettes. Scale bars = 1 cm.

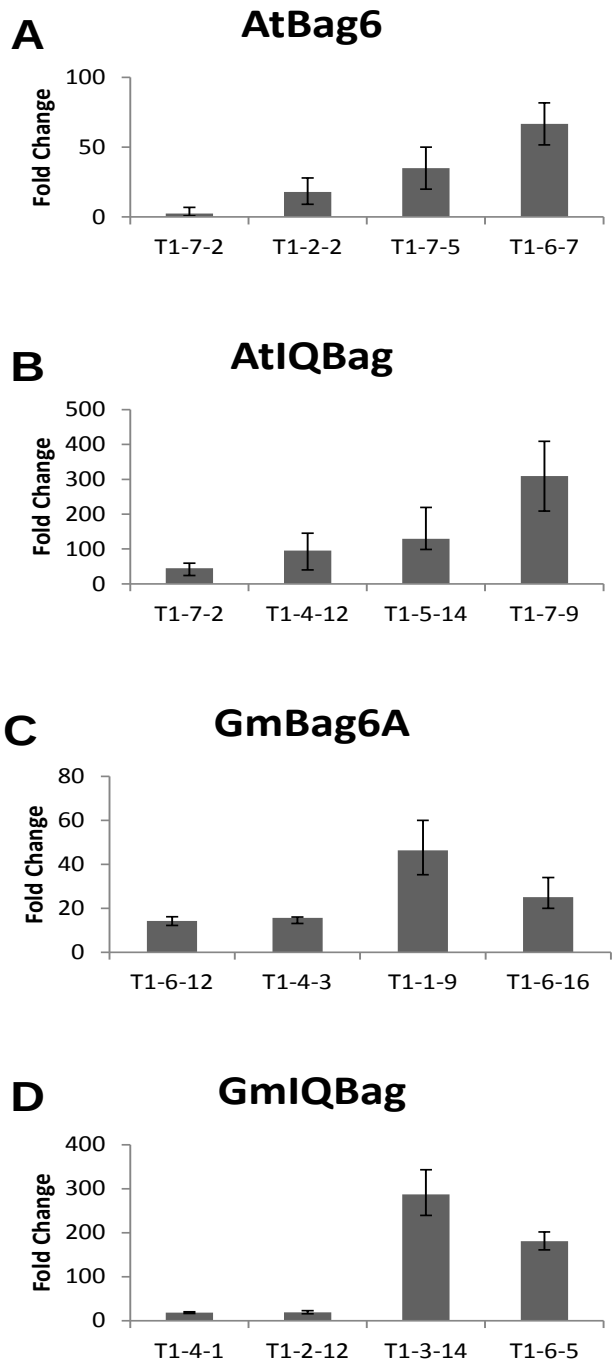
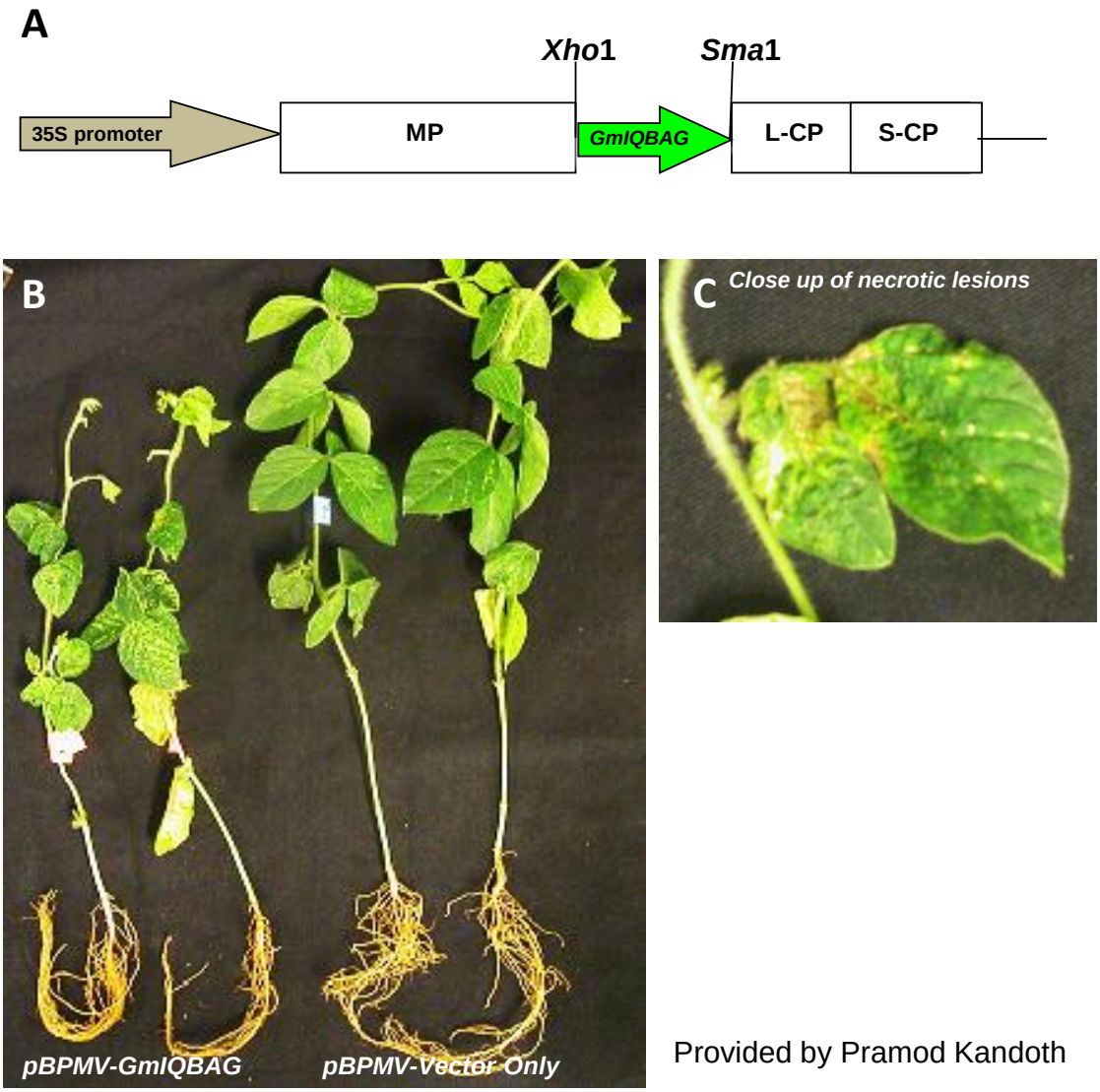


Figure 2.10. Transgene expression level in *BAG6* and *IQBAG* overexpression lines. (A) Real-time quantitative polymerase chain reaction (qPCR) on *Arabidopsis* T_1 plants overexpressing *AtBAG6*, phenotypic severity increases from left to right. (B) qPCR on *Arabidopsis* T_1 plants overexpressing *AtIQBAG*, phenotypic severity increases from left to right. (C) qPCR on *Arabidopsis* T_1 plants overexpressing *GmBAG6A*, phenotypic severity increases from left to right. (D) qPCR on *Arabidopsis* T_1 plants overexpressing *GmIQBAG*, phenotypic severity increases from left to right.



Provided by Pramod Kandoth

Figure 2.11. Overexpression of *GmIQBAG* in soybean using *Bean pod mottle virus* (BPMV). (A) BPMV construct used for overexpression. (B) *pBPMV-GmIQBAG* inoculated soybean compared to BPMV empty vector. Plants inoculated with *pBPMV-GmIQBAG* exhibit severe stunting compared to plants inoculated with BPMV only. (C) Close up of leaves expressing *GmIQBAG* showing necrotic leaf lesions.

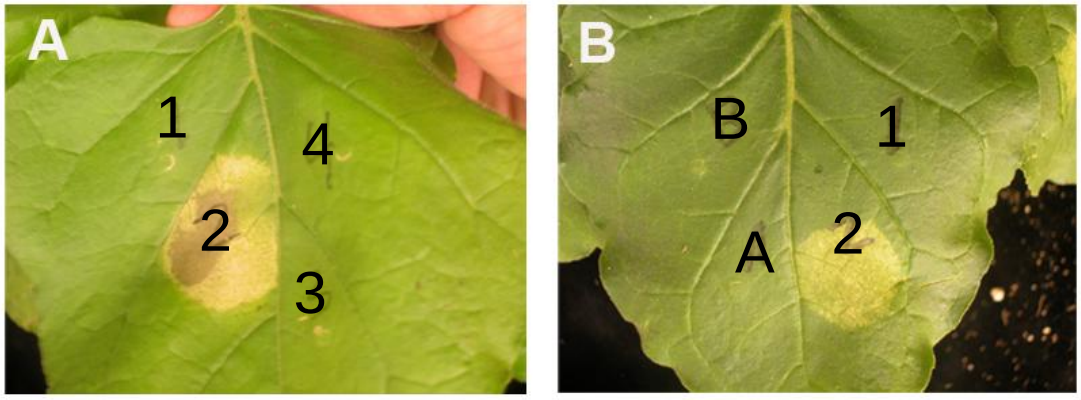


Figure 2.12. Symptoms in *Nicotiana benthamiana* leaves (~ 3-4 week old) agroinfiltrated with various *BAG* gene constructs. (A) 1: Negative control 2X35S-Empty Vector, 2 Positive control, co-infiltration of RX, a NB-LRR protein of potato and the coat protein (CP) from potato virus X (PVX), 3: *AtBAG6*, 4: *GmBAG6A*. (B) 1: Negative control 2X35S-Empty Vector, 2: Positive control co-infiltration of RX2 and CP, A: *AtIQBAG*, B: *GmIQBAG*. Photographs were taken 5 days post-infiltration.

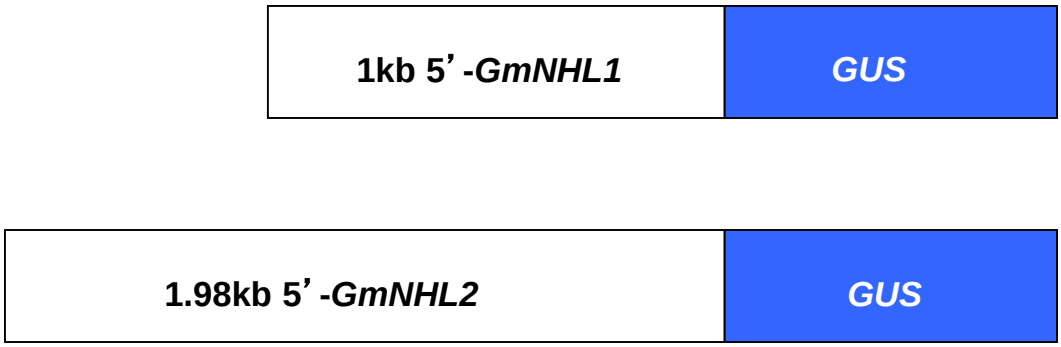


Figure 2.13. *GmNHL1* and *GmNHL2* promoter-GUS reporter gene fusion constructs used in expression studies.

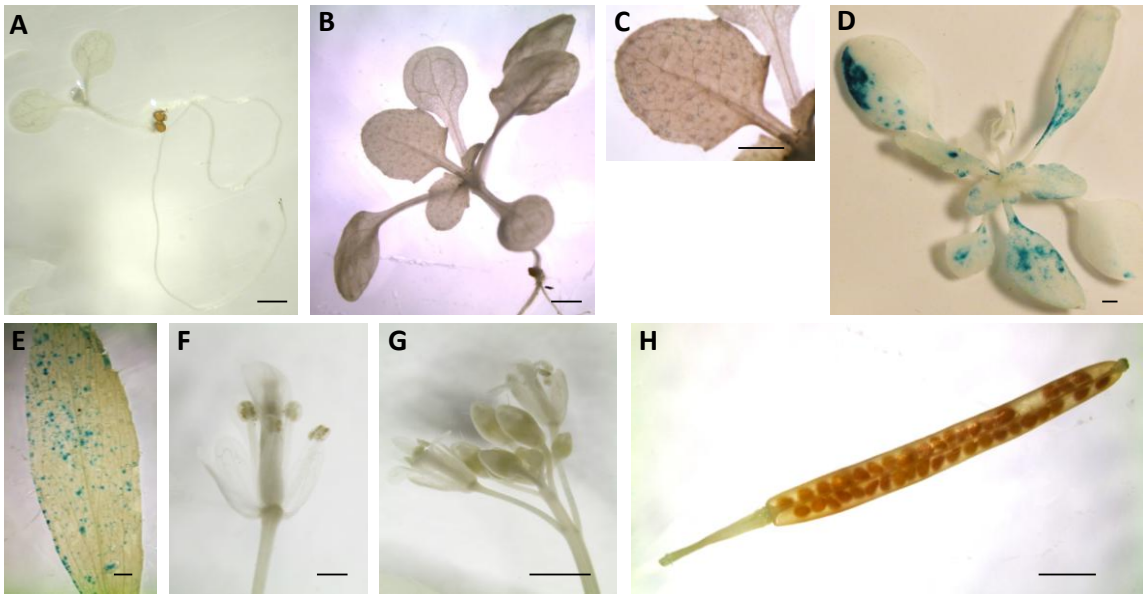


Figure 2.14. *GmNHL1-GUS* expression during plant development. (A) 7-day old seedling, (B-C) aerial rosette of a 14-day old plant and close up of leaf, (D) aerial rosette of a 21-day old plant, (E) cauline leaf, (F) flower, (G) flower cluster, (H) silique. Scale bars = 1mm.

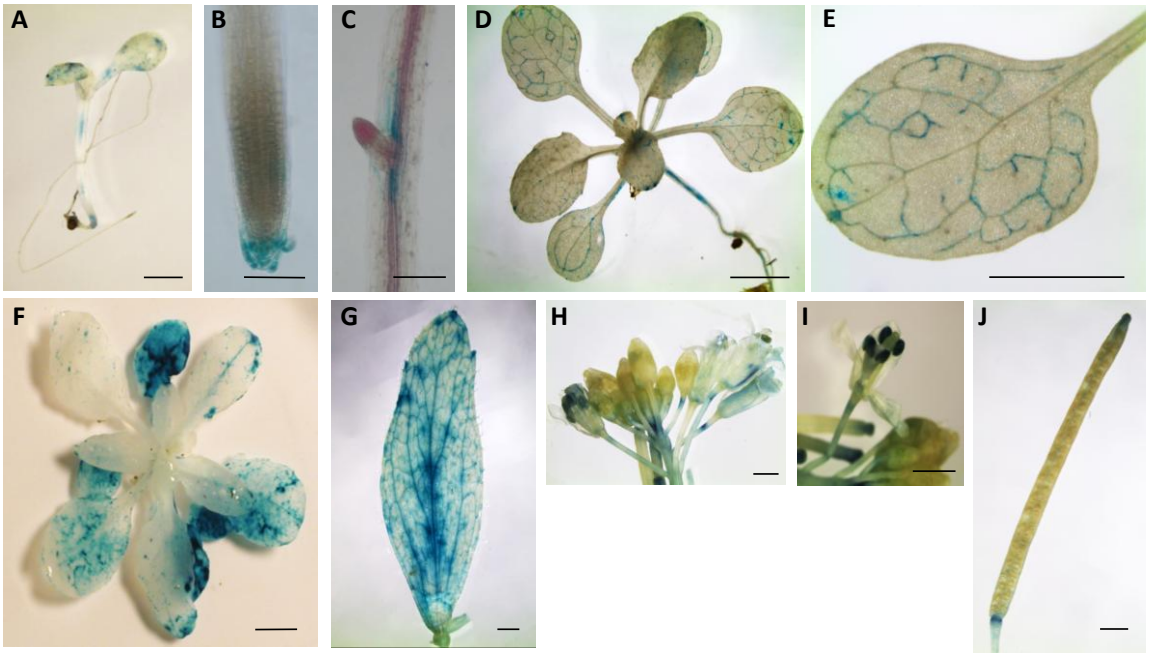


Figure 2.15. *GmNHL2-GUS* expression during plant development. (A) 7-day old seedling, (B) Arabidopsis root tip, (C) emerging lateral roots, (D-E) aerial rosette of 14-day old plant and close up of leaf, (F) aerial rosette of a 21-day old plant, (G) cauline leaf, (H) flower cluster, (I) flower, (L) silique. Scale bars = 1mm.

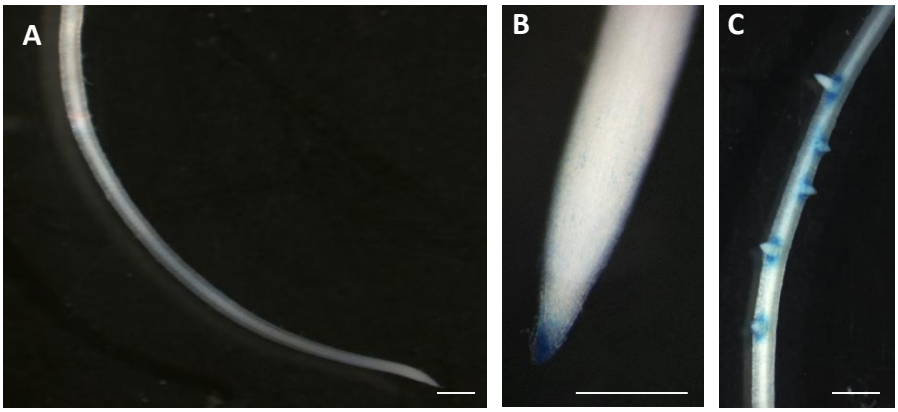


Figure 2.16 *GmNHL1-GUS* and *GmNHL2-GUS* expression in soybean roots. (A) *GmNHL1-GUS* soybean root, (B) *GmNHL2-GUS* root tip, (C) *GmNHL2-GUS* emerging lateral root.

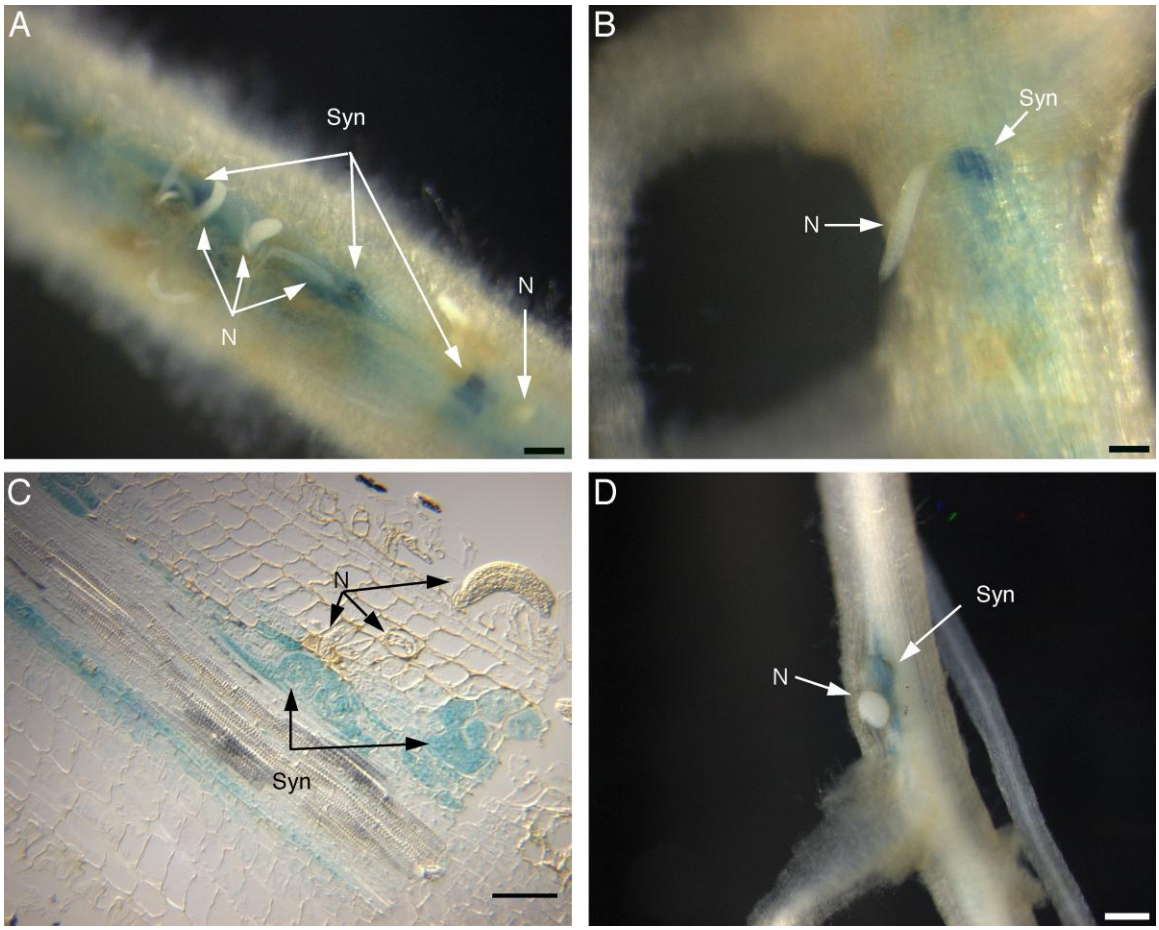


Figure 2.17. *GmNHL1-GUS* expression during nematode infection of Williams 82 transgenic soybean hairy roots. (A-B) 10 days post-inoculation (dpi) with *Heterodera glycines*, SCN. (C) Longitudinal section through root shown in A. (D) 17 dpi with SCN. N = nematode; Syn = syncytium. Scale bars = 100 μm.

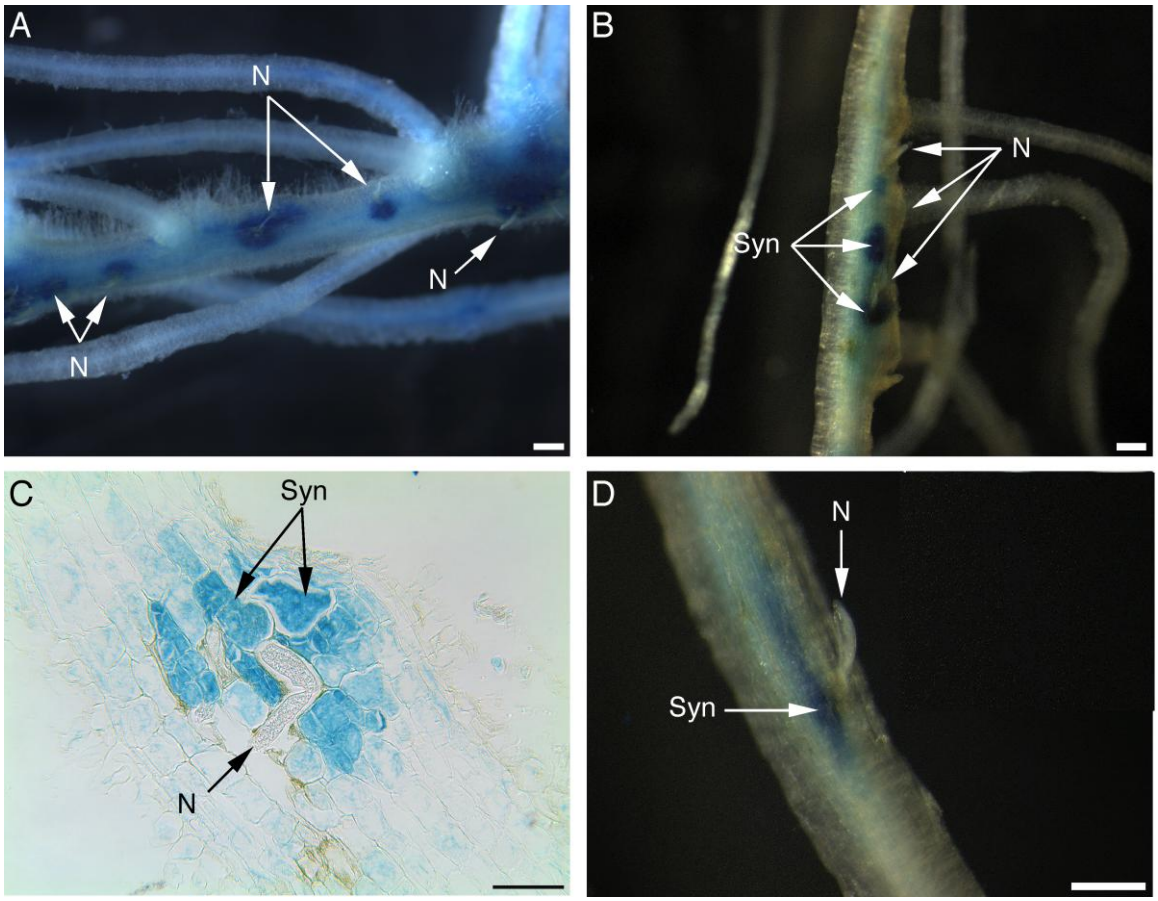


Figure 2.18. *GmNHL2-GUS* expression during nematode infection of Williams 82 transgenic soybean hairy roots. (A) 10 days post-inoculation (dpi) with *Heterodera glycines*, SCN. (B) 15 dpi with SCN. (C) Longitudinal section through root shown in A. (D) 15 days post-infection with SCN. N = nematode; Syn = syncytium. Scale bars = 100 μm.

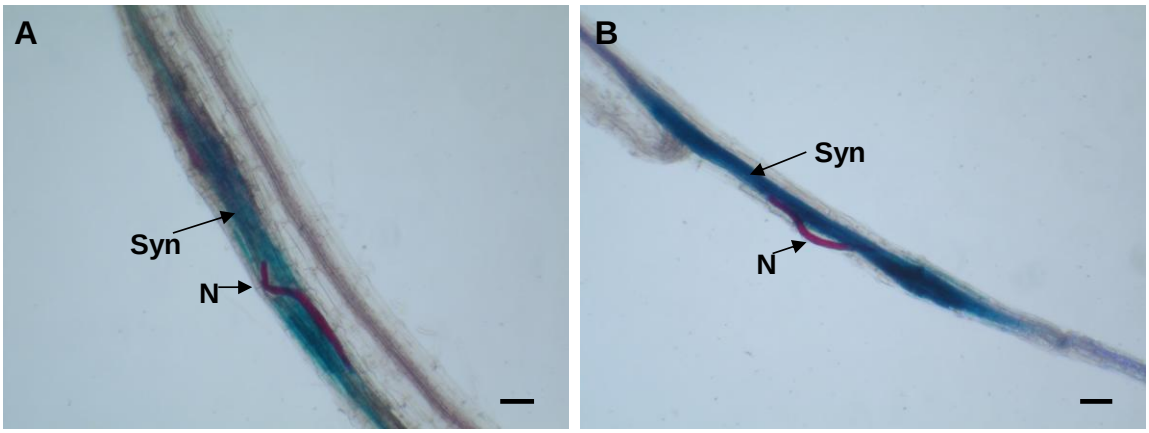


Figure 2.19. *GmNHL-GUS* expression during nematode infection of transgenic *Arabidopsis*. (A) *GmNHL1-GUS* expression in *Arabidopsis* at 2 days post-inoculation (dpi) with *Heterodera schachtii*, BCN. (B) *GmNHL2-GUS* expression in *Arabidopsis* at 2 dpi with BCN. N = nematode; Syn = syncytium. Scale bars = 100 μm. Photos provided by Jianying Wang.

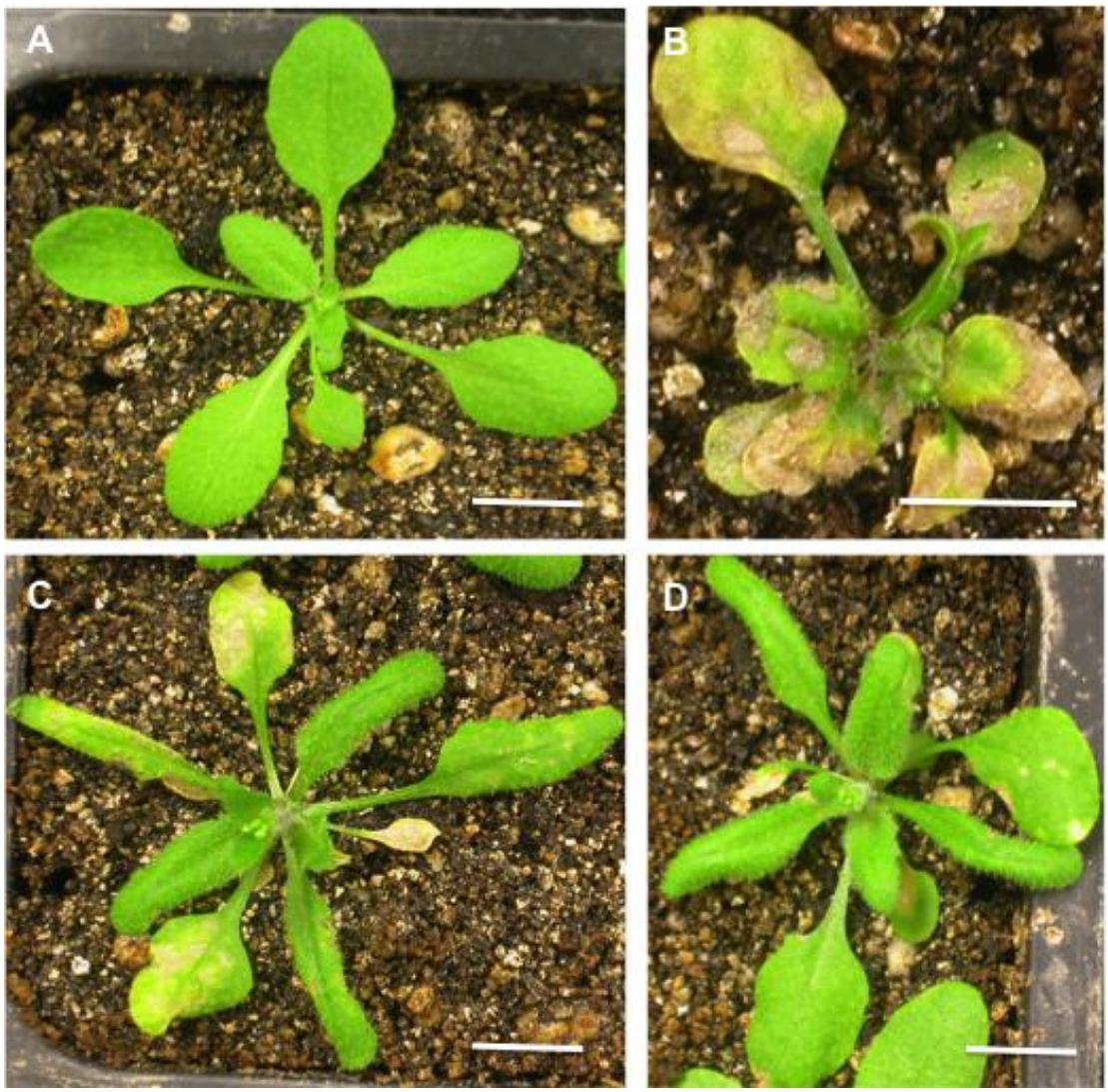


Figure 2.20. Transgenic *Arabidopsis* expressing *AtIQBAG* under control of the *GmNHL1* promoter. (A) Wild-type Col-0 (B-D) Expression of the *AtIQBAG* under the control of the *GmNHL1* promoter in *Arabidopsis* induces sporadic cell death phenotypes in varying degrees. Plants shown are 21 days old. Scale bars = 1 cm.

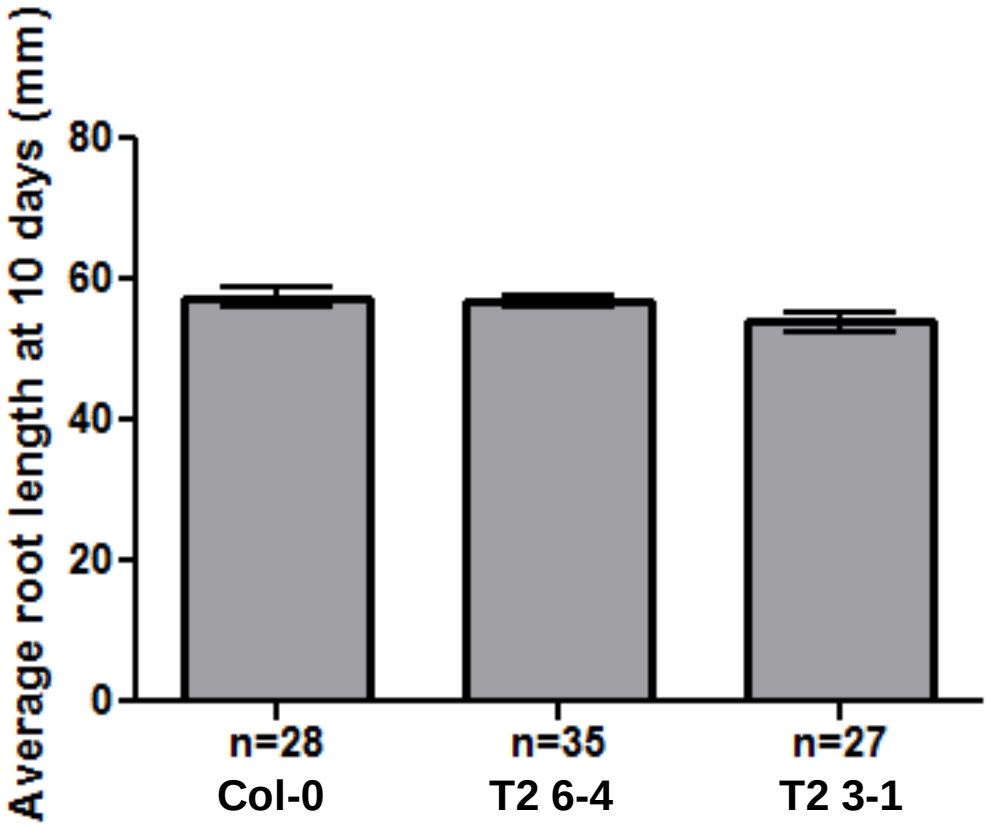


Figure 2.21. Effect of expressing *AtIQBAG* under control of the *GmNHL1* promoter on *Arabidopsis* root growth. Seedlings were grown for 10 days on modified Knops nutrient agar. Roots were photographed at 10 days post sowing and Image SXM software was used to measure root lengths in (mm). Data represent the mean average. n = sample size and error bars = SE.

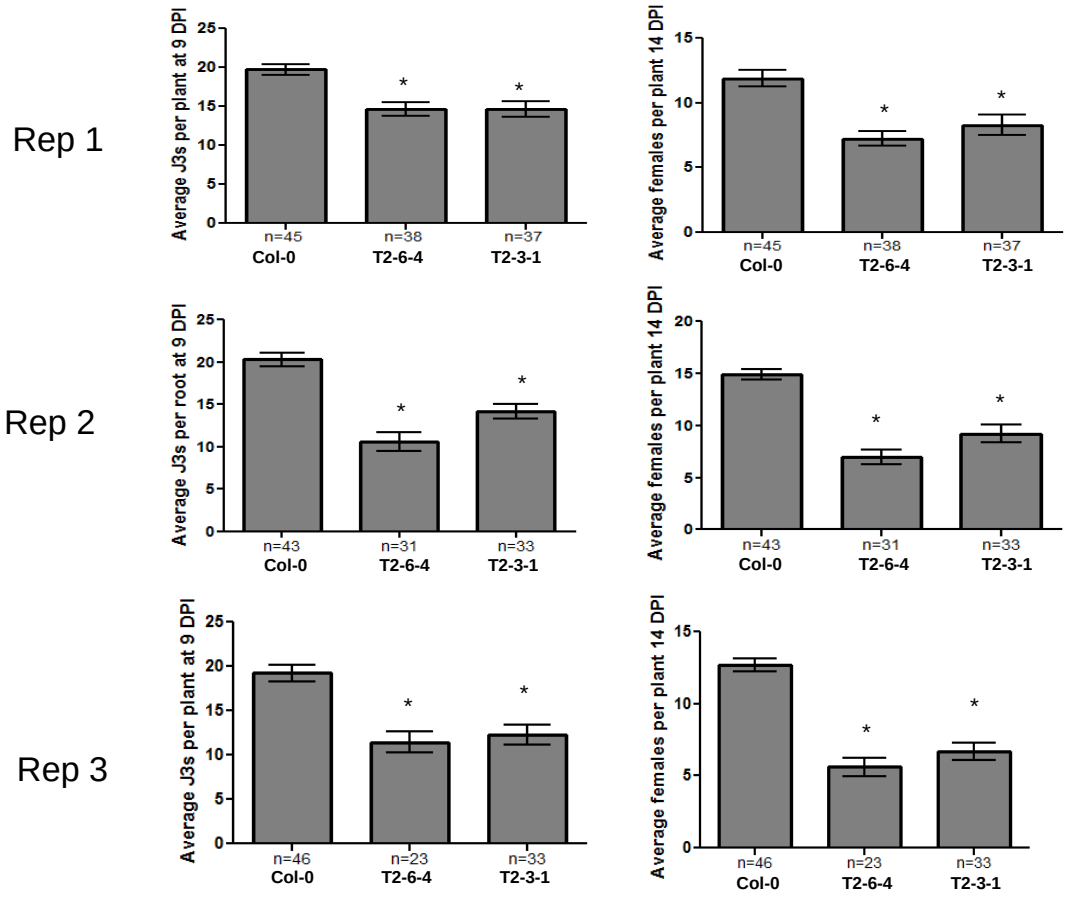


Figure 2.22. *NHL1-AtIQBAG* nematode infection assay. Individuals of Col-0 and two T_2 *NHL1-AtIQBAG* expression lines were infected in 12-well plates with 175 infective J2 of *Heterodera schachtii*. Third-stage juveniles (J3) were counted at 9 days post-inoculation (dpi) and fourth-stage juvenile (J4) female nematodes were counted on individuals at 14 dpi. Three independent experiments were performed showing similar results. Data represent the mean average, n = sample size and error bars = SE. Asterisks indicate statistically significant differences by Student's *t*-test ($P < 0.001$).

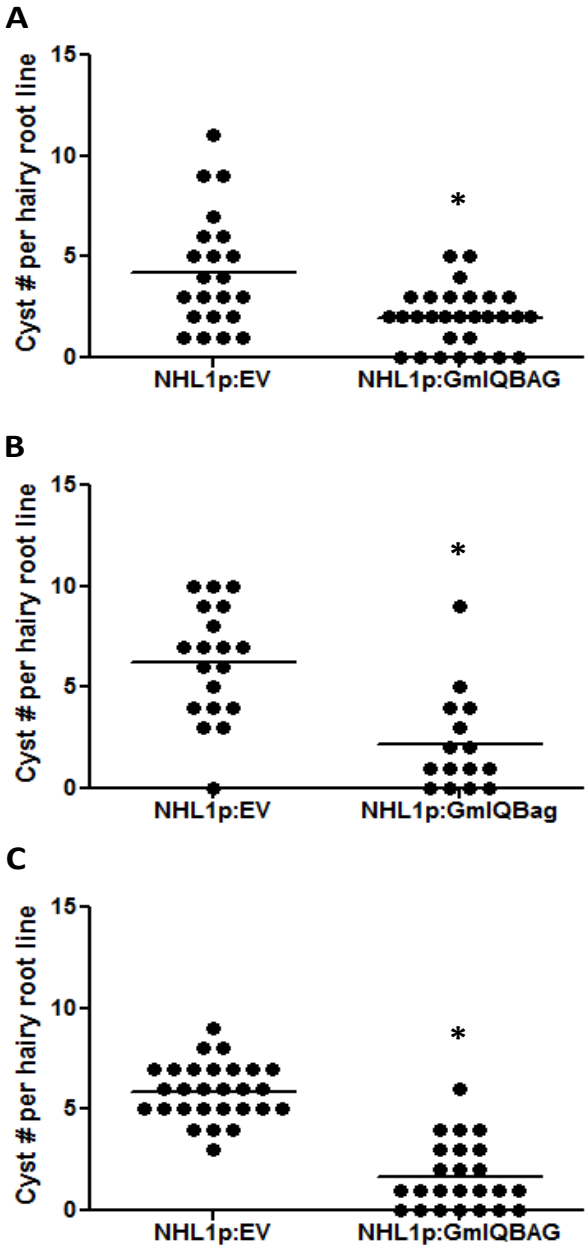


Figure 2.23. SCN development on transgenic *NHL1-GmIQBAG* soybean (cv. Williams 82) hairy roots at 30 days post-inoculation. Hairy roots were generated by transforming in the empty vector (*NHL1p-EV*), used as the susceptible control and *NHL1p-GmIQBAG*. Circles represent the number of cysts on a single hairy root line. (A) Repetition 1 (B) Repetition 2 (C) Repetition 3. Asterisks indicate statistically significant differences by Student's *t*-test (*P* < 0.001).

Chapter 3

Conclusion and Future Perspectives

In the U.S today, SCN continues to reign as the most important pathogen of soybean. Despite the planting of resistant soybean cultivars, SCN causes an estimated $1.286 billion dollars annually in yield losses. Unfortunately, very little is known about the molecular mechanisms underlying resistance. This lack of understanding continues to be a major hurdle in the progress toward enhancing the effectiveness and durability of natural plant resistance and enabling the design of novel strategies for resistance through biotechnological approaches.

As this thesis was being written, the genes underlying two major SCN resistance QTL, *Rhg1 and Rhg4*, were reported (Cook et al, 2012; Liu et al., 2012). These discoveries will no doubt shed new light on our understanding of SCN resistance in soybean. The focus now will be to unravel the molecular mechanism of the resistance genes, signaling pathways, and response genes controlled by these loci. Mapping studies allowed for the generation of NILs differing only at the SCN resistance loci (Liu et al., 2011, Mudge, 1999) which are useful for comparative analyses of plant gene expression between resistant and susceptible soybean in response to SCN. NILs are powerful tools to study the effects of specific gene loci with reduced genetic background effects. Our laboratory recently used soybean NILs at the *Rhg1* locus for comparative studies to identify genes involved in SCN resistance (Kandoth et al., 2011). Using laser capture microdissection (LCM), syncytial cells were isolated from SCN-infected soybean roots and hybridized onto the Affymetrix GeneChip Soybean Genome Arrays. Out of the 37,000 genes printed on the chip, 1,447 genes were differentially expressed between NIL-R and NIL-S lines. From this analysis, 828

genes were found to be upregulated in syncytia of the resistant line compared to the susceptible line. One gene, coding for a BCL2-associated athanogene (BAG) domain protein with highest homology to the *Arabidopsis* BAG6 protein previously shown to induce PCD in yeast and plants, was selected for further study and became the focus of this thesis.

The soybean *BAG* gene was the most highly up-regulated gene in syncytia of the resistant line (87-fold) (Kandoth et al., 2011). The full-length 3,387-bp *GmBAG6A* cDNA sequence was cloned and contained an open reading frame encoding an 1129 amino acid protein. We found that expression of *GmBAG6A* in yeast cells induced cell death. In addition, expression of just the region of the protein containing the IQ and BAG domains (GmIQBAG) was sufficient for cell death and a more robust cell death phenotype was observed with the truncated form of the protein. To further characterize the function of *GmBAG6A*, overexpression studies were conducted in *Arabidopsis*, soybean, and tobacco. Interestingly, unlike *AtBAG6*, *GmBAG6A* did not cause significant dwarfism and induction of disease-like necrotic lesions in *Arabidopsis*, which were previously observed in the *AtBAG6* overexpression plants (Kang et al., 2006). This was true for plants expressing either the full-length *GmBAG6A* or the *GmIQBAG* domain only. However, as observed in yeast, plants expressing the *AtIQBAG* domain exhibited much more severe cell death phenotypes than plants expressing the full-length AtBAG6 protein. It is interesting to note the difference in transgene expression levels between the full-length and truncated versions of the protein. The expression level of the IQBAG domain construct was significantly higher

than the expression level of the full-length *BAG6* construct in all plants tested, suggesting that regions outside the IQBAG domain may be important for *BAG6* gene regulation. Although only weak phenotypes were observed in *Arabidopsis* plants expressing the *GmIQBAG* domain, when the *GmIQBAG* domain was expressed in soybean, plants were stunted and sporadic disease-like necrotic lesions formed on leaves. No cell death phenotypes were observed when either *AtBAG6* or *GmBAG6A* was expressed in tobacco leaves. These findings suggest that the reason for this difference may be due to species-specific interactions of plant BAG6 proteins with binding partners. Taken together, the results from these studies indicate that the GmBAG6A protein can induce cell death in yeast and plants similar to *AtBAG6.*

The robust cell death phenotype elicited by overexpression of either the *AtIQBAG* domain in *Arabidopsis* or the *GmIQBAG* domain in soybean led us to test its potential use in engineering a novel form of resistance to nematodes by specifically targeting the feeding cells for degeneration. To test this, the nematode-inducible promoter *GmNHL1*, which was found to be induced in syncytia during nematode infection, and had very low, if any, background expression in root tissues in the absence of nematode infection, was selected for targeted *IQBAG* expression studies. *Arabidopsis* plants expressing *NHL1-AtIQBAG* grew similar to wild-type plants, except that sporadic necrotic leaf lesions were observed on rosette leaves by 14 days post germination consistent with the sporadic GUS expression pattern observed in leaves of the transgenic *Arabidopsis* expressing *NHL1-GUS*. On the otherhand, root growth of the

Arabidopsis NHL1-AtIQBAG lines and soybean *NHL1-GmIQBAG* hairy root lines was unaffected. In nematode infection assays of the transgenic *Arabidopsis* expressing *GmNHL1-AtIQBAG,* nematode development was reduced by 35-50%. Similar results were observed in soybean hairy roots generated from the SCN-susceptible soybean cv. Williams 82 where a 45-70% reduction in nematode development was measured compared with control lines. These results demonstrate that specific expression of the IQBAG domain in nematode feeding sites could be a viable means of nematode control.

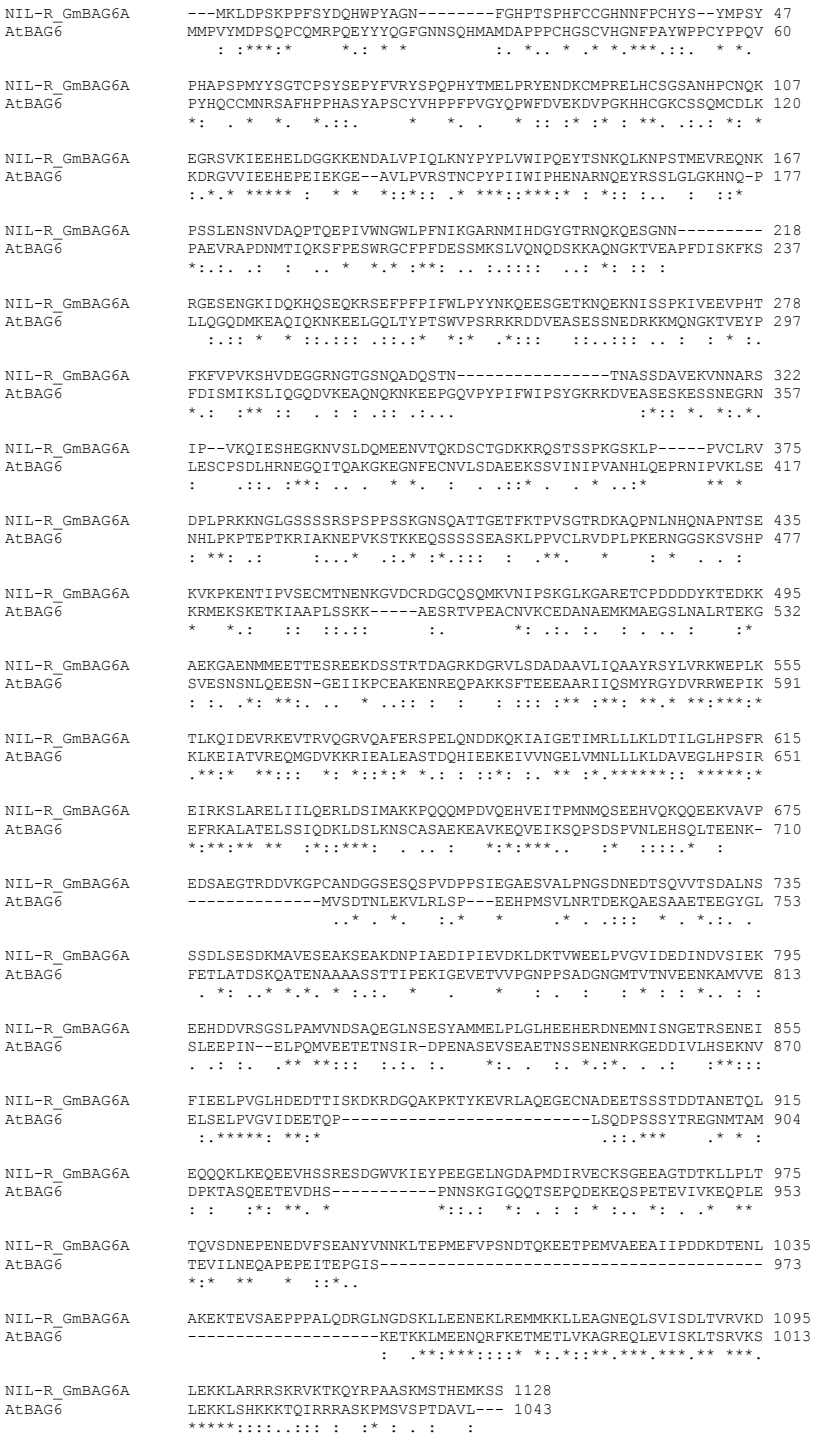
Even though the promoter is not ideal for use in commercial lines due to its sporadic expression in the vegetative tissue, it nonetheless shows the potential benefits and efficacy in using *IQBAG* under the control of a more stringent nematode-inducible promoter. Utilizing this new knowledge gained about *GmBAG6A* and its role in cell death offers promise in extending beyond just soybean cyst nematode control, but offers the possibility of a broader form of nematode control. This type of transgenic approach contrasts with the specificity of resistance genes, which are normally only protective for one or select nematode species. Combining natural and transgenic resistance shows promise in bestowing a durable control method for protecting soybeans and other crop species against a broad range of plant-parasitic nematodes. Improved technology for controlling nematodes has the potential to dramatically improve annual yields for most crops impacted by sedentary endoparasitic nematodes. Thus, the potential to use an endogenous plant gene for nematode control substantiates further research.

Further research will be necessary to answer the many questions that still remain including the exact role, if any, of *GmBAG6A* in soybean resistance to SCN. Gene silencing approaches in soybean, including RNAi and VIGS, have not been able to conclusively demonstrate a role for BAG6 in SCN resistance. Unfortunately, there are several limitations with these approaches including the fact that both approaches only achieve partial silencing and this can be further complicated by gene redundancy in soybean. We also observed interesting differences in transgene expression level and correspondingly in phenotype severity between full-length and truncated BAG6 proteins. Investigating how these proteins are differentially regulated would be interesting to explore further. More research is needed to determine how the different BAG protein domains contribute to the important and diverse roles these proteins play to maintain the homeostasis of the cell. Moreover, determining where BAG6 proteins localize in cells, what their binding partners are, the signaling cascades they are involved in, and how they mechanistically function are all questions that need to be addressed to gain a better understanding of this important class of plant proteins.

APPENDIX

Figures

Figure A1. Alignment of the amino acid sequence of GmBAG6A NIL-7923R with AtBAG6.



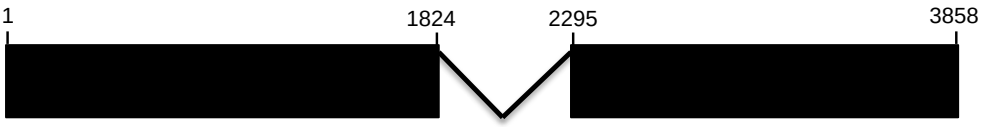


Figure A2. *GmBAG6A* gene model.

Figure A3. Alignment of the coding regions of *GmBAG6A* from NIL-7923R, NIL-7923S and Williams 82.

```
NIL-R_GmBAG6A      ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCCTATGCCGGC 60
NIL-S_GmBAG6A      ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCCTATGCCGGC 60
Williams           ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCCTATGCCGGC 60
                   ************************************************************

NIL-R_GmBAG6A      AATTTTGGGCACCCTACTTCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCAT 120
NIL-S_GmBAG6A      AATTTTGGGCACCCTACTTCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCAT 120
Williams           AATTTTGGGCACCCTACTTCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCAT 120
                   ************************************************************

NIL-R_GmBAG6A      TATAGCTACATGCCTTCATATCCTCATGCCCCTTCTCCAATGTACTATTCTGGAACTTGT 180
NIL-S_GmBAG6A      TATAGCTACATGCCTTCATATCCTCATGCCCCTTCTCCAATGTACTATTCTGGAACTTGT 180
Williams           TATAGCTACATGCCTTCATATCCTCATGCCCCTTCTCCAATGTACTATTCTGGAACTTGT 180
                   ************************************************************

NIL-R_GmBAG6A      CCTTCATATAGTGAACCATATTTTGTTCGTTATTCCCCACAACCACATTATACCATGGAG 240
NIL-S_GmBAG6A      CCTTCATATAGTGAACCATATTTTGTTCGTTATTCCCCACAACCACATTATACCATGGAG 240
Williams           CCTTCATATAGTGAACCATATTTTGTTCCTTATTCCCCACAACCACATTATACCATGGAG 240
                   ***************************** ******************************

NIL-R_GmBAG6A      CTGCCTAGGTATGAAAATGACAAATGCATGCCCCGAGAGCTTCATTGTTCTGGTTCTGCT 300
NIL-S_GmBAG6A      CTGCCTAGGTATGAAAATGACAAATGCATGCCCCGAGAGCTTCATTGTTCTGGTTCTGCT 300
Williams           CTGCCTAGGTATGAAAATGACAAATGCATGCCCCGAGAGCTTCATTGTTCTGGTTCTGCT 300
                   ************************************************************

NIL-R_GmBAG6A      AATCATCCATGCAACCAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAGCATGAACTGGAT 360
NIL-S_GmBAG6A      AATCATCCATGCAACCAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAGCATGAACTGGAT 360
Williams           AATCATCCATGCAACCAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAGCATGAACTGGAT 360
                   ************************************************************

NIL-R_GmBAG6A      GGTGGAAAGAAAGAGAATGATGCTTTGGTCCCAATTCAGCTCAAGAATTATCCATATCCC 420
NIL-S_GmBAG6A      GGTGGAAAGAAAGAGAATGATGCTTTGGTCCCAATTCAGCTCAAGAATTATCCATATCCC 420
Williams           GGTGGAAAGAAAGAGAATGATGCTTTGGTCCCAATTCAGCTCAAGAATTATCCATATCCC 420
                   ************************************************************

NIL-R_GmBAG6A      TTAGTTTGGATTCCACAGGAGTACACAAGTAACAAACAGCTGAAAAATCCTAGTACAATG 480
NIL-S_GmBAG6A      TTAGTTTGGATTCCACAGGAGTACACAAGTAACAAACAGCTGAAAAATCCTAGTACAATG 480
Williams           TTAGTTTGGATTCCACAGGAGTACACAAGTAACAAACAGCTGAAAAATCCTAGTACAATG 480
                   ************************************************************

NIL-R_GmBAG6A      GAAGTTCGTGAACAAAACAAGCCTTCTAGTCTTGAGAATTCTAATGTTGATGCACAGCCA 540
NIL-S_GmBAG6A      GAAGTTCGTGAACAAAACAAGCCTTCTAGTCTTGAGAATTCTAATGTTGATGCACAGCCA 540
Williams           GAAGTTCGTGAACAAAACAAGCCTTCTAGTCTTGAGAATTCTAATGTTGATGCGCAGCCA 540
                   ************************************************** ******

NIL-R_GmBAG6A      ACACAGGAGCCTATAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCGGAAC 600
NIL-S_GmBAG6A      ACACAGGAGCCTATAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCGGAAC 600
Williams           ACACAGGAGCCTATAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCGGAAC 600
                   ************************************************************

NIL-R_GmBAG6A      ATGATTCACGATGGATATGGAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGG 660
NIL-S_GmBAG6A      ATGATTCACGATGGATATGGAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGG 660
Williams           ATGATTCACGATGGATATGGAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGG 660
                   ************************************************************

NIL-R_GmBAG6A      GAATCTGAAAATGGAAAAATAGACCAGAAACATCAAAGTGAACAGAAGAGGTCAGAATTC 720
NIL-S_GmBAG6A      GAATCTGAAAATGGAAAAATAGACCAGAAACATCAAAGTGAACAGAAGAGGTCAGAATTC 720
Williams           GAATCTGAAAATGGAAAAATAGACCAGAAACATCAAAGTGAACAGAAGAGGTCAGAATTC 720
                   ************************************************************

NIL-R_GmBAG6A      CCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCAGGAGGAGAGTGGAGAGACTAAG 780
NIL-S_GmBAG6A      CCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCAGGAGGAGAGTGGAGAGACTAAG 780
Williams           CCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCAGGAGGAGAGTGGAGAGACTAAG 780
                   ************************************************************

NIL-R_GmBAG6A      AACCAGGAGAAAAACATTTCTTCACCAAAAATTGTTGAGGAGGTACCCCATACATTCAAA 840
NIL-S_GmBAG6A      AACCAGGAGAAAAACATTTCTTCACCAAAAATTGTTGAGGAGGTACCCCATACATTCAAA 840
Williams           AACCAGGAGAAAAACATTTCTTCACCAAAAATTGTTGAGGAGGTACCCCATACATTCAAA 840
                   ************************************************************

NIL-R_GmBAG6A      TTTGTTCCAGTGAAGTCTCATGTTGATGAAGGTGGTAGGAACGGAACCGGATCAAATCAA 900
NIL-S_GmBAG6A      TTTGTTCCAGTGAAGTCTCATGTTGATGAAGGTGGTAGGAACGGAACCGGATCAAATCAA 900
Williams           TTTGTTCCAGTGAAGTCTCATGTTGATGAAGGTGGTAGGAACAGAACCGGATCAAATCAA 900
                   ****************************************** ****************

NIL-R_GmBAG6A      GCTGATCAATCCACAAATACAAATGCTTCTTCGGATGCTGTAGAGAAAGTGAATAATGCC 960
NIL-S_GmBAG6A      GCTGATCAATCCACAAATACAAATGCTTCTTCGGATGCTGTAGAGAAAGTGAATAATGCC 960
Williams           GCTGATCAATCCACAAATACAAATGCTTCTTCGGATGCTGTAGAGAAAGTGAATAATGCC 960
                   ************************************************************

NIL-R_GmBAG6A      AGAAGCATACCTGTGAAGCAGATAGAATCCCACGAAGGAAAAAATGTTTCTCTCGATCAA 1020
NIL-S_GmBAG6A      AGAAGCATACCTGTGAAGCAGATAGAATCCCACGAAGGAAAAAATGTTTCTCTCGATCAA 1020
Williams           AGAAGCATACCTGTGAAGCAGATAGAATCCCACGAAGGAAAAAATGTTTCTCTCGATCAA 1020
                   ************************************************************
```

```
NIL-R_GmBAG6A       ATGGAAGAGAATGTGACACAAAAGGACTCTTGCACTGGGGACAAAAAGAGACAATCTACA 1080
NIL-S_GmBAG6A       ATGGAAGAGAATGTGACACAAAAGGACTCTTGCACTGGGGACAAAAAGAGACAATCTACA 1080
Williams            ATGGAAGAGAATGTGACACAAAAGGACTCTTGCACTGGGGACAAAAAGAGACAATCTACA 1080
                    ************************************************************


NIL-R_GmBAG6A       TCTTCACCTAAAGGATCCAAGTTACCTCCGGTTTGTCTGAGAGTTGATCCACTACCAAGG 1140
NIL-S_GmBAG6A       TCTTCACCTAAAGGATCCAAGTTACCTCCGGTTTGTCTGAGAGTTGATCCACTACCAAGG 1140
Williams            TCTTCACCTAAAGGATCCAAGTTACCTCCGGTTTGTCTGAGAGTTGATCCACTACCAAGG 1140
                    ************************************************************

NIL-R_GmBAG6A       AAGAAAAATGGCCTCGGGAGTTCGAGTTCGAGGTCCCCAAGTCCACCTTCATCAAAAGGG 1200
NIL-S_GmBAG6A       AAGAAAAATGGCCTCGGGAGTTCGAGTTCGAGGTCCCCAAGTCCACCTTCATCAAAAGGG 1200
Williams            AAGAAAAATGGCCACGGGAGTTCGAGTTCGAGGTCCCCAAGTCCACCTTCATCAAAAGGG 1200
                    ************ ************************************************

NIL-R_GmBAG6A       AATTCCCAAGCTACAACTGGTGAAACATTCAAGACTCCTGTGAGTGGCACACGTGACAAG 1260
NIL-S_GmBAG6A       AATTCCCAAGCTACAACTGGTGAAACATTCAAGACTCCTGTGAGTGGCACACGTGACAAG 1260
Williams            AATTCCCAAGCTACAACTGGTGAAACATTCAAGACTCCTGTGAGTGGCACACATGACAAG 1260
                    ************************************************** *******

NIL-R_GmBAG6A       GCTCAGCCAAATTTGAATCATCAGAATGCTCCAAACACCAGTGAGAAAGTTAAACCAAAG 1320
NIL-S_GmBAG6A       GCTCAGCCAAATTTGAATCATCAGAATGCTCCAAACACCAGTGAGAAAGTTAAACCAAAG 1320
Williams            GCTCAGCCAAATTTGAATCATCAGAATGCTCCAAACACCAGTGAGAAAGTTAAACCAAAG 1320
                    ************************************************************

NIL-R_GmBAG6A       GAGAACACCATTCCGGTGTCAGAATGCATGACTAATGAAAACAAGGGTGTTGACTGTAGG 1380
NIL-S_GmBAG6A       GAGAACACCATTCCGGTGTCAGAATGCATGACTAATGAAAACAAGGGTGTTGACTGTAGG 1380
Williams            GAGAACACCATTCCGGTGTCAGAATGCATGACTAATGAAAACAAGGGTGTTGACTGTAGG 1380
                    ************************************************************

NIL-R_GmBAG6A       GATGGATGTCAGAGCCAGATGAAAGTAAACATACCCAGTAAAGGTCTGAAAGGGGCAAGG 1440
NIL-S_GmBAG6A       GATGGATGTCAGAGCCAGATGAAAGTAAACATACCCAGTAAAGGTCTGAAAGGGGCAAGG 1440
Williams            GATGGATGTCAGAGCCAGATGAAAGTAAACATACCCAGTAAAGGTCTGAAAGGGGCAAGG 1440
                    ************************************************************

NIL-R_GmBAG6A       GAAACATGTCCAGATGATGATGACTATAAGACTGAAGATAAAAAGGCAGAGAAAGGAGCA 1500
NIL-S_GmBAG6A       GAAACATGTCCAGATGATGATGACTATAAGACTGAAGATAAAAAGGCAGAGAAAGGAGCA 1500
Williams            GAAACATGTCCAGATGATGATGACTATAAGACTGAAGATAAAAAGGCAGAGAAAGGAGCA 1500
                    ************************************************************

NIL-R_GmBAG6A       GAAAATATGATGGAGGAAACTACTGAATCAAGGGAAGAGAAGGATTCAAGCACACGAACT 1560
NIL-S_GmBAG6A       GAAAATATGATGGAGGAAACTACTGAATCAAGGGAAGAGAAGGATTCAAGCACACGAACT 1560
Williams            GAAAATATGATGGAGGAAACTACTGAATCAAGGGAAGAGAAGGATTCAAGCACACGAACT 1560
                    ************************************************************

NIL-R_GmBAG6A       GATGCAGGTCGAAAAGATGGAAGAGTTTTGTCAGATGCAGATGCTGCTGTTTTGATACAA 1620
NIL-S_GmBAG6A       GATGCAGGTCGAAAAGATGGAAGAGTTTTGTCAGATGCAGATGCTGCTGTTTTGATACAA 1620
Williams            GATGCGGGTCGAAAAGATGGAAGAGTTTTGTCAGATGCAGATGCTGCTGTTTTGATACAA 1620
                    ***** ********************************************************

NIL-R_GmBAG6A       GCTGCATATCGCAGTTATCTAGTTAGAAAATGGGAACCGTTGAAGACGTTGAAGCAGATA 1680
NIL-S_GmBAG6A       GCTGCATATCGCAGTTATCTAGTTAGAAAATGGGAACCGTTGAAGACGTTGAAGCAGATA 1680
Williams            GCTGCATATCGCAGTTATCTAGTTAGAAAATGGGAACCGTTGAAGAAGTTGAAGCAGATA 1680
                    ********************************************** **************

NIL-R_GmBAG6A       GATGAAGTCAGGAAGGAGGTGACTCGTGTTCAAGGCCGTGTTCAAGCTTTTGAGAGATCT 1740
NIL-S_GmBAG6A       GATGAAGTCAGGAAGGAGGTGACTCGTGTTCAAGGCCGTGTTCAAGCTTTTGAGAGATCT 1740
Williams            GATGAAGTCAGGAAGGAGGTGACTCGTGTTCAAGGCCGTGTTCAAGCTTTTGAGAGATCT 1740
                    ************************************************************

NIL-R_GmBAG6A       CCCGAACTTCAAAATGATGACAAACAAAAAATTGCAATTGGAGAGACCATAATGAGACTC 1800
NIL-S_GmBAG6A       CCCGAACTTCAAAATGATGACAAACAAAAAATTGCAATTGGAGAGACCATAATGAGACTC 1800
Williams            CCCGAACTTCAAAATGATGACAAACAAAAAATTGCAATTGAAGAGACCATAATGAAACTC 1800
                    ****************************************** ************** ****

NIL-R_GmBAG6A       CTGCTGAAGTTGGATACTATACTGGGTTTGCATCCAAGTTTCAGGGAGATCAGAAAATCC 1860
NIL-S_GmBAG6A       CTGCTGAAGTTGGATACTATACTGGGTTTGCATCCAAGTTTCAGGGAGATCAGAAAATCC 1860
Williams            CTGCTGAAGTTGGATACTATACTGGGTTTGCATCCAAGTTTCAGGGAGATCAGAAAATCC 1860
                    ************************************************************

NIL-R_GmBAG6A       TTGGCTAGGGAGCTCATAATCTTGCAAGAAAGGCTTGATTCTATAATGGCCAAGAAACCT 1920
NIL-S_GmBAG6A       TTGGCTAGGGAGCTCATAATCTTGCAAGAAAGGCTTGATTCTATAATGGCCAAGAAACCT 1920
Williams            TTGGCTAGGGAGCTCATAATCTTGCAAGAAAGGCTTGATTCTATAATGGCCAAGAAACCT 1920
                    ************************************************************

NIL-R_GmBAG6A       CAGCAGCAGATGCCGGATGTTCAGGAACACGTTGAAATCACTCCAATGAACATGCAGAGT 1980
NIL-S_GmBAG6A       CAGCAGCAGATGCCGGATGTTCAGGAACACGTTGAAATCACTCCAATGAACATGCAGAGT 1980
Williams            CAGCAGCAGATGCCGGATGTTCAGGAACACGTTGAAATCACTCCAATGAACATGCAGAGT 1980
                    ************************************************************

NIL-R_GmBAG6A       GAAGAACATGTGCAAAAGCAGCAAGAAGAAAAGGTTGCTGTACCAGAGGATTCAGCTGAA 2040
NIL-S_GmBAG6A       GAAGAACATGTGCAAAAGCAGCAAGAAGAAAAGGTTGCTGTACCAGAGGATTCAGCTGAA 2040
Williams            GAAGAACATGTGCAAAAGCAGCAAGAAGAAAAGGTTGCTGTACCAGAGGATTCAGCTGAA 2040
                    ************************************************************

NIL-R_GmBAG6A       GGCACTAGGGATGATGTAAAAGGTCCTTGTGCTAATGATGGTGGAAGTGAATCTCAGTCA 2100
NIL-S_GmBAG6A       GGCACTAGGGATGATGTAAAAGGTCCTTGTGCTAATGATGGTGGAAGTGAATCTCAGTCA 2100
Williams            GGCACTAGGGATGATGTAAAAGGTCCTTGTGCTAATGATGGTGGAAGTGAATCTCAGTCA 2100
                    ************************************************************

NIL-R_GmBAG6A       CCAGTTGATCCTCCATCAATTGAGGGAGCAGAGTCTGTTGCACTTCCAAATGGCTCAGAT 2160
NIL-S_GmBAG6A       CCAGTTGATCCTCCATCAATTGAGGGAGCAGAGTCTGTTGCACTTCCAAATGGCTCAGAT 2160
```

```
Williams          CCAGTTGATCCTCCATCAAATGAGGGAGCAGAGTCTGTTGCACTTCCAAATGGCTCAGAT 2160
                  ****************** ****************************************


NIL-R_GmBAG6A     AATGAGGACACCAGCCAAGTGGTTACATCTGATGCATTGAATTCTTCAAGTGATCTGTCT 2220
NIL-S_GmBAG6A     AATGAGGACACCAGCCAAGTGGTTACATCTGATGCATTGAATTCTTCAAGTGATCTGTCT 2220
Williams          AATGAGGACACCAGCCAAGTGGTTACATCTGATGCATTGAATTCTTCAAGTGATCTGTCT 2220
                  ************************************************************


NIL-R_GmBAG6A     GAGAGTGACAAAATGGCTGTGGAATCCGAAGCTAAATCAGAAGCGAAAGACAATCCGATT 2280
NIL-S_GmBAG6A     GAGAGTGACAAAATGGCTGTGGAATCCGAAGCTAAATCAGAAGCGAAAGACAATCCGATT 2280
Williams          GAGAGTGACAAAATGGCTGTGGAATCCGAAGCTAAATCAGAAGTGAAAGACAATCCGATT 2280
                  ********************************************* ****************


NIL-R_GmBAG6A     GCGGAAGACATTCCCATTGAGGTTGATAAATTGGACAAGACTGTTTGGGAAGAATTGCCT 2340
NIL-S_GmBAG6A     GCGGAAGACATTCCCATTGAGGTTGATAAATTGGACAAGACTGTTTGGGAAGAATTGCCT 2340
Williams          GCGGAAGACATTCCCATTGAGGTTGATAAATTGGACAAGACTGTTTGGGAAGAATTGCCT 2340
                  ************************************************************


NIL-R_GmBAG6A     GTGGGAGTTATTGATGAAGATATCAATGATGTTAGTATTGAGAAGGAAGAACATGATGAT 2400
NIL-S_GmBAG6A     GTGGGAGTTATTGATGAAGATATCAATGATGTTAGTATTGAGAAGGAAGAACATGATGAT 2400
Williams          GTGGGAGTTATTGATGAAGATATCAATGATGTTAGTATTGAGAAGGAAGAACATGATGAT 2400
                  ************************************************************


NIL-R_GmBAG6A     GTTAGATCGGGAAGTCTCCCAGCCATGGTGAATGATTCGGCACAAGAAGGATTAAATTCA 2460
NIL-S_GmBAG6A     GTTAGATCGGGAAGTCTCCCAGCCATGGTGAATGATTCGGCACAAGAAGGATTAAATTCA 2460
Williams          ATTAGATCGGGAAGTCTCCCAGCCATGGTGAATGATTCGGCACAAGAAGGATTAAATTCA 2460
                   ************************************************************


NIL-R_GmBAG6A     GAGAGCTATGCAATGATGGAACTGCCATTGGGATTACATGAGGAGCATGAAAGGGACAAT 2520
NIL-S_GmBAG6A     GAGAGCTATGCAATGATGGAACTGCCATTGGGATTACATGAGGAGCATGAAAGGGACAAT 2520
Williams          GAGAGCTATGCAATGATGGAACTGCCATTGGGATTACATGAGGAGCATGAAAGGGACAAT 2520
                  ************************************************************


NIL-R_GmBAG6A     GAAATGAATATTTCTAATGGAGAAACACGGTCTGAGAATGAGATATTTATTGAGGAGCTT 2580
NIL-S_GmBAG6A     GAAATGAATATTTCTAATGGAGAAACACGGTCTGAGAATGAGATATTTATTGAGGAGCTT 2580
Williams          GAAATGAATATTTCTAATGGAGAAACACGGTCTGAGAATGAGATATTTATTGAGGAGCTT 2580
                  ************************************************************


NIL-R_GmBAG6A     CCTGTGGGACTGCACGATGAAGATACAACAATATCTAAAGATAAGAGGGATGGTCAAGCT 2640
NIL-S_GmBAG6A     CCTGTGGGACTGCACGATGAAGATACAACAATATCTAAAGATAAGAGGGATGGTCAAGCT 2640
Williams          CCTGTGGGACTGCACGATGAAGATACAACAATATCTAAAGATAAGAGGGATGGTCAAGCT 2640
                  ************************************************************


NIL-R_GmBAG6A     AAGCCTAAAACATATAAAGAGGTTCGACTAGCTCAAGAAGGGGAATGCAATGCAGATGAG 2700
NIL-S_GmBAG6A     AAGCCTAAAACATATAAAGAGGTTCGACTAGCTCAAGAAGGGGAATGCAATGCAGATGAG 2700
Williams          AAGCCTAAAACATATAAAGAGGTTCGACTAGCTCAAGAAGGGGAATGCAATGCAGATGAG 2700
                  ************************************************************


NIL-R_GmBAG6A     GAAACAAGTTCTTCCACAGATGACACTGCCAACGAAACTCAACTAGAGCAACAGCAGAAG 2760
NIL-S_GmBAG6A     GAAACAAGTTCTTCCACAGATGACACTGCCAACGAAACTCAACTAGAGCAACAGCAGAAG 2760
Williams          GAAACAAGTTCTTCCACAGATGACACTGCCAACGAAACTCAACTAGAGCAACAGCAGAAG 2760
                  ************************************************************


NIL-R_GmBAG6A     CTGAAAGAGCAAGAAGAGGTGCATTCTTCTAGGGAATCAGATGGCTGGGTAAAAATTGAG 2820
NIL-S_GmBAG6A     CTGAAAGAGCAAGAAGAGGTGCATTCTTCTAGGGAATCAGATGGCTGGGTAAAAATTGAG 2820
Williams          CTGAAAGAGCAAGAAGAGGTGCATTATTCTAGGGAATCAGATGGCTGGGTAAAAATTGAG 2820
                  ************************* ****************************************


NIL-R_GmBAG6A     TACCCGGAAGAAGGTGAACTCAATGGTGATGCACCAATGGATATAAGAGTTGAGTGCAAG 2880
NIL-S_GmBAG6A     TACCCGGAAGAAGGTGAACTCAATGGTGATGCACCAATGGATATAAGAGTTGAGTGCAAG 2880
Williams          TACCCGGAAGAAGGTGAACTCAATGGTGATGCACCAATGGATATAAGAGTTGAGTGCAAG 2880
                  ************************************************************


NIL-R_GmBAG6A     TCAGGTGAGGAAGCTGGAACTGATACTAAGTTGCTTCCTTTAACAACACAAGTCAGTGAT 2940
NIL-S_GmBAG6A     TCAGGTGAGGAAGCTGGAACTGATACTAAGTTGCTTCCTTTAACAACACAAGTCAGTGAT 2940
Williams          TCAGGTGAGGAAGCTGGAACTGATACTAAGTTGCTTCCTTTAACAACACAAGTCAGTGAT 2940
                  ************************************************************


NIL-R_GmBAG6A     AATGAACCAGAAAATGAAGATGTATTCTCAGAAGCAAATTATGTAAATAACAAATTAACC 3000
NIL-S_GmBAG6A     AATGAACCAGAAAATGAAGATGTATTCTCAGAAGCAAATTATGTAAATAACAAATTAACC 3000
Williams          AATGAACCAGAAAATGAAGATGTATTCTCAGAAGCAAATTATGTAAATAACAAATTAACC 3000
                  ************************************************************


NIL-R_GmBAG6A     GAGCCAATGGAGTTTGTACCTTCCAATGACACACAGAAGGAGGAGACACCAGAGATGGTT 3060
NIL-S_GmBAG6A     GAGCCAATGGAGTTTGTACCTTCCAATGACACACAGAAGGAGGAGACACCAGAGATGGTT 3060
Williams          GAGCCAATGGAGTTTGTACCTTCCAATGACACACAGAAGGAGGAGACACCAGAGATGGTT 3060
                  ************************************************************


NIL-R_GmBAG6A     GCTGAAGAGGCAATTATCCCTGATGATAAAGACACAGAAAATTTGGCCAAAGAGAAAACT 3120
NIL-S_GmBAG6A     GCTGAAGAGGCAATTATCCCTGATGATAAAGACACAGAAAATTTGGCCAAAGAGAAAACT 3120
Williams          GCTGAAGAGGCAATTATCCCTGATGATAAAGACACAGAAAATTTGGCCAAAGAGAAAACT 3120
                  ************************************************************


NIL-R_GmBAG6A     GAAGTATCTGCAGAACCACCACCTGCACTGCAAGACCGAGGGTTAAACGGTGACTCGAAG 3180
NIL-S_GmBAG6A     GAAGTATCTGCAGAACCACCACCTGCACTGCAAGACCGAGGGTTAAACGGTGACTCGAAG 3180
Williams          GAAGTATCTGCAGAACCACCACCTGCATTGCAAGACCGAGGGTTAAACGGTGACTCGAAG 3180
                  *************************** ****************************************


NIL-R_GmBAG6A     TTATTAGAAGAGAATGAGAAGTTAAGGGAGATGATGAAGAAGTTGCTTGAAGCCGGGAAT 3240
NIL-S_GmBAG6A     TTATTAGAAGAGAATGAGAAGTTAAGGGAGATGATGAAGAAGTTGCTTGAAGCCGGGAAT 3240
Williams          TTATTAGAAGAGAATGAGAAGTTAAGGGAGATGATGAAGAAGTTGCTTGAAGCCGGGAAT 3240
                  ************************************************************


NIL-R_GmBAG6A     GAACAGTTAAGCGTGATATCAGATTTGACTGTCAGAGTGAAGGACTTGGAGAAGAAATTA 3300
```

```
NIL-S_GmBAG6A       GAACAGTTAAGCGTGATATCAGATTTGACTGTCAGAGTGAAGGACTTGGAGAAGAAATTA 3300
Williams            GAACAGTTAAGCGTGATATCAGATTTGACTGTCAGAGTGAAGGACTTGGAGAAGAAATTA 3300
                    ************************************************************


NIL-R_GmBAG6A       GCCAGGAGAAGGAGTAAGAGAGTGAAGACAAAACAGTATAGACCCGCAGCTTCCAAAATG 3360
NIL-S_GmBAG6A       GCCAGGAGAAGGAGTAAGAGAGTGAAGACAAAACAGTATAGACCCGCAGCTTCCAAAATG 3360
Williams            GCCAGGAGAAGGAGTAAGAGAGTGAAGACAAAACAGTATAGACCCGCAGCTTCCAAAATG 3360
                    ************************************************************

NIL-R_GmBAG6A       TCTACCCATGAAATGAAATCCTCCTAA 3387
NIL-S_GmBAG6A       TCTACCCATGAAATGAAATCCTCCTAA 3387
Williams            TCTACCCATGAAATGAAATCCTCCTAA 3387
                    ***************************
```

Figure A4. Amino-acid alignment of GmBAG6A from NIL-7923R, NIL-7923S and Williams 82.

```
NIL-R_GmBAG6A      MKLDPSKPPFSYDQHWPYAGNFGHPTSPHFCCGHNNFPCHYSYMPSYPHAPSPMYYSGTC 60
NIL-S_GmBAG6A      MKLDPSKPPFSYDQHWPYAGNFGHPTSPHFCCGHNNFPCHYSYMPSYPHAPSPMYYSGTC 60
Williams           MKLDPSKPPFSYDQHWPYAGNFGHPTSPHFCCGHNNFPCHYSYMPSYPHAPSPMYYSGTC 60
                   ************************************************************

NIL-R_GmBAG6A      PSYSEPYFVRYSPQPHYTMELPRYENDKCMPRELHCSGSANHPCNQKEGRSVKIEEHELD 120
NIL-S_GmBAG6A      PSYSEPYFVRYSPQPHYTMELPRYENDKCMPRELHCSGSANHPCNQKEGRSVKIEEHELD 120
Williams           PSYSEPYFVPYSPQPHYTMELPRYENDKCMPRELHCSGSANHPCNQKEGRSVKIEEHELD 120
                   ********* ************************************************

NIL-R_GmBAG6A      GGKKENDALVPIQLKNYPYPLVWIPQEYTSNKQLKNPSTMEVREQNKPSSLENSNVDAQP 180
NIL-S_GmBAG6A      GGKKENDALVPIQLKNYPYPLVWIPQEYTSNKQLKNPSTMEVREQNKPSSLENSNVDAQP 180
Williams           GGKKENDALVPIQLKNYPYPLVWIPQEYTSNKQLKNPSTMEVREQNKPSSLENSNVDAQP 180
                   ************************************************************

NIL-R_GmBAG6A      TQEPIVWNGWLPFNIKGARNMIHDGYGTRNQKQESGNNRGESENGKIDQKHQSEQKRSEF 240
NIL-S_GmBAG6A      TQEPIVWNGWLPFNIKGARNMIHDGYGTRNQKQESGNNRGESENGKIDQKHQSEQKRSEF 240
Williams           TQEPIVWNGWLPFNIKGARNMIHDGYGTRNQKQESGNNRGESENGKIDQKHQSEQKRSEF 240
                   ************************************************************

NIL-R_GmBAG6A      PFPIFWLPYYNKQEESGETKNQEKNISSPKIVEEVPHTFKFVPVKSHVDEGGRNGTGSNQ 300
NIL-S_GmBAG6A      PFPIFWLPYYNKQEESGETKNQEKNISSPKIVEEVPHTFKFVPVKSHVDEGGRNGTGSNQ 300
Williams           PFPIFWLPYYNKQEESGETKNQEKNISSPKIVEEVPHTFKFVPVKSHVDEGGRNRTGSNQ 300
                   ************************************************** *****

NIL-R_GmBAG6A      ADQSTNTNASSDAVEKVNNARSIPVKQIESHEGKNVSLDQMEENVTQKDSCTGDKKRQST 360
NIL-S_GmBAG6A      ADQSTNTNASSDAVEKVNNARSIPVKQIESHEGKNVSLDQMEENVTQKDSCTGDKKRQST 360
Williams           ADQSTNTNASSDAVEKVNNARSIPVKQIESHEGKNVSLDQMEENVTQKDSCTGDKKRQST 360
                   ************************************************************

NIL-R_GmBAG6A      SSPKGSKLPPVCLRVDPLPRKKNGLGSSSSRSPSPPSSKGNSQATTGETFKTPVSGTRDK 420
NIL-S_GmBAG6A      SSPKGSKLPPVCLRVDPLPRKKNGLGSSSSRSPSPPSSKGNSQATTGETFKTPVSGTRDK 420
Williams           SSPKGSKLPPVCLRVDPLPRKKNGHGSSSSRSPSPPSSKGNSQATTGETFKTPVSGTHDK 420
                   ********************** ********************************:**

NIL-R_GmBAG6A      AQPNLNHQNAPNTSEKVKPKENTIPVSECMTNENKGVDCRDGCQSQMKVNIPSKGLKGAR 480
NIL-S_GmBAG6A      AQPNLNHQNAPNTSEKVKPKENTIPVSECMTNENKGVDCRDGCQSQMKVNIPSKGLKGAR 480
Williams           AQPNLNHQNAPNTSEKVKPKENTIPVSECMTNENKGVDCRDGCQSQMKVNIPSKGLKGAR 480
                   ************************************************************

NIL-R_GmBAG6A      ETCPDDDDYKTEDKKAEKGAENMMEETTESREEKDSSTRTDAGRKDGRVLSDADAAVLIQ 540
NIL-S_GmBAG6A      ETCPDDDDYKTEDKKAEKGAENMMEETTESREEKDSSTRTDAGRKDGRVLSDADAAVLIQ 540
Williams           ETCPDDDDYKTEDKKAEKGAENMMEETTESREEKDSSTRTDAGRKDGRVLSDADAAVLIQ 540
                   ************************************************************

NIL-R_GmBAG6A      AAYRSYLVRKWEPLKTLKQIDEVRKEVTRVQGRVQAFERSPELQNDDKQKIAIGETIMRL 600
NIL-S_GmBAG6A      AAYRSYLVRKWEPLKTLKQIDEVRKEVTRVQGRVQAFERSPELQNDDKQKIAIGETIMRL 600
Williams           AAYRSYLVRKWEPLKKLKQIDEVRKEVTRVQGRVQAFERSPELQNDDKQKIAIEETIMKL 600
                   ***************.************************************ ****:*

NIL-R_GmBAG6A      LLKLDTILGLHPSFREIRKSLARELIILQERLDSIMAKKPQQQMPDVQEHVEITPMNMQS 660
NIL-S_GmBAG6A      LLKLDTILGLHPSFREIRKSLARELIILQERLDSIMAKKPQQQMPDVQEHVEITPMNMQS 660
Williams           LLKLDTILGLHPSFREIRKSLARELIILQERLDSIMAKKPQQQMPDVQEHVEITPMNMQS 660
                   ************************************************************

NIL-R_GmBAG6A      EEHVQKQQEEKVAVPEDSAEGTRDDVKGPCANDGGSESQSPVDPPSIEGAESVALPNGSD 720
NIL-S_GmBAG6A      EEHVQKQQEEKVAVPEDSAEGTRDDVKGPCANDGGSESQSPVDPPSIEGAESVALPNGSD 720
Williams           EEHVQKQQEEKVAVPEDSAEGTRDDVKGPCANDGGSESQSPVDPPSNEGAESVALPNGSD 720
                   ********************************************* *************

NIL-R_GmBAG6A      NEDTSQVVTSDALNSSSDLSESDKMAVESEAKSEAKDNPIAEDIPIEVDKLDKTVWEELP 780
NIL-S_GmBAG6A      NEDTSQVVTSDALNSSSDLSESDKMAVESEAKSEAKDNPIAEDIPIEVDKLDKTVWEELP 780
Williams           NEDTSQVVTSDALNSSSDLSESDKMAVESEAKSEVKDNPIAEDIPIEVDKLDKTVWEELP 780
                   ********************************.***************************

NIL-R_GmBAG6A      VGVIDEDINDVSIEKEEHDDVRSGSLPAMVNDSAQEGLNSESYAMMELPLGLHEEHERDN 840
NIL-S_GmBAG6A      VGVIDEDINDVSIEKEEHDDVRSGSLPAMVNDSAQEGLNSESYAMMELPLGLHEEHERDN 840
Williams           VGVIDEDINDVSIEKEEHDDIRSGSLPAMVNDSAQEGLNSESYAMMELPLGLHEEHERDN 840
                   ********************:***************************************
```

Role in Other Projects Related to SCN Resistance

As part of an ongoing effort to clone the SCN resistance gene at the *Rhg4* locus in soybean, I assisted with the SCN phenotyping and genotyping of thousands of recombinant inbred lines (RILs) generated from crosses between resistant and susceptible soybean lines for positional cloning studies. In addition, I assisted with the SCN phenotyping and genotyping of TILLING mutants in candidate genes mapped to the *Rhg4* locus. I collected all plant tissues and assisted in phenotyping the plants, which involved counting the number of female cysts able to grow on each line. Due to the high number of plants and labor involved in manually counting the cysts, an automated counting method using a fluorescence-based imaging system was explored. Although the imager was not intended for cyst counting, I optimized a method for automated counting of SCN females using the fluorescence-based image analysis system. For optimal automated imaging, females were washed from roots at 30 days post-inoculation into small Petri dishes. Using a Kodak Image Station 4000MM Pro, the Petri dishes were scanned using excitation and emission wavelengths of 470 nm and 535 nm, respectively. Fluorescent images were captured and analyzed with Carestream Molecular Imaging Software for automated counting. We demonstrated that the automated fluorescent-based imaging system is just as accurate ($r^2 \geq 0.95$) and more efficient (>50% faster) than manual counting under a microscope. This method can greatly improve the consistency and turnaround of data while reducing the time and labor commitment associated with SCN female counting. This method is currently being utilized by our SCN breeding

program, and is being explored by other academic researchers and various companies.

The results of this work were published in:

*Brown S, ***Yeckel G**, *Heinz R, Clark K, Sleper D, and Mitchum MG. A high-throughput automated technique for counting females of *Heterodera glycines* using a fluorescence-based imaging system. *Journal of Nematology* 2010;42(3):201-206. *equal contributing authors

*Liu, S, *Kandoth, P, Warren S, **Yeckel G**, Heinz R, Alden J, Yang C, Jamai A, El-Mellouki T, Juvale P, Hill J, Baum TJ, Cianzio S, Whitham S, Korkin D, †Mitchum MG, †Meksem K. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 2012*:* doi:10.1038/nature11651.*equal contributing authors, †co-corresponding authors

Data presented in this chapter was published in:

*Brown S, *__Yeckel G__, *Heinz R, Clark K, Sleper D, and Mitchum MG. A high-throughput automated technique for counting females of *Heterodera glycines* using a fluorescence-based imaging system. *Journal of Nematology* 2010;42(3):201-206. *equal contributing authors

Abstract

The soybean cyst nematode (SCN), *Heterodera glycines*, is the most damaging pathogen of soybean. Methods to phenotype soybean varieties for resistance to SCN are currently very laborious and time consuming. Streamlining a portion of this phenotyping process could increase productivity and accuracy. Here we report an automated method to count SCN females using a fluorescence-based imaging system that is well suited to high-throughput SCN phenotyping methods used in greenhouse screening. For optimal automated imaging, females were washed from roots at 30 days post-inoculation into small Petri dishes. Using a Kodak Image Station 4000MM Pro, the Petri dishes were scanned using excitation and emission wavelengths of 470 nm and 535 nm, respectively. Fluorescent images were captured and analyzed with Carestream Molecular Imaging Software for automated counting. We demonstrate that the automated fluorescent-based imaging system is just as accurate ($r^2 \geq 0.95$) and more efficient (>50% faster) than manual counting under a microscope. This method can greatly improve the consistency and turnaround of data while reducing the time and labor commitment associated with SCN female counting.

Introduction

The soybean cyst nematode (SCN), *Heterodera glycines* Ichinohe, is a serious economic threat to soybean (*Glycine max* (L.) Merr.) producers. Yield suppression attributed to *H. glycines* resulted in an estimated $1.2 billion in losses to U.S. soybean producers annually from 2006 to 2009, causing more yield loss than any other soybean pest (Koenning & Wrather, 2010). Planting resistant soybean cultivars is the primary means to limit soybean yield suppression due to *H. glycines*. However, greater than 90% of commercially available soybean cultivars derive their resistance from a single genetic source, plant introduction (PI) 88788. Consequently, monoculture of a single type of resistance has led to an increase in SCN field populations virulent on PI 88788 (Mitchum et al., 2007, Niblack et al., 2008) This has intensified the need for breeders to identify additional sources of resistance to broaden the genetic base of H. glycines resistance in soybean cultivars. The standard calculation for determining SCN resistance levels of soybean germplasm is the Female Index, where FI = (mean number of females on a test soybean line) / (mean number of females on the standard susceptible) × 100 . Determination of resistance is then based on the FI. The FI is also used to determine the virulence profiles of genetically diverse field populations of SCN using the HG Type test (Niblack et al., 2002) to screen mapping populations for the identification and cloning of SCN resistance genes, and to evaluate mutant and transgenic soybean lines for SCN resistance or susceptibility. Thus, SCN phenotyping of soybean is a widespread practice in SCN research. Screening is typically conducted in a greenhouse in

containers with volumes between 100 and 500 cm^3 to accommodate growth of a 30-day-old soybean plant (Niblack et al., 2002). The traditional SCN phenotyping method employs manual counting of SCN females washed from roots at 30 days post-inoculation (dpi) under a microscope (Niblack et al., 2002; 2009). Manual counting is a labor intensive, time consuming, and cumbersome process. In addition, the small size and varied color of SCN females mixed with fragments of root debris requires a trained eye for accurate counting. Thus, counting discrepancies among individuals is a common contributing factor to experimental variability.

In recent years, fluorescence-based counting methods have been explored (Lu et al., 2005). The imaging system developed by Pioneer Hi-Bred International, Inc and described by (Lu et al., 2005) employs an enclosed lighting system equipped with a camera attached to a computer. The system was optimized for imaging individual Petri dishes containing SCN-infected soybean root explants grown on Gamborg's B5 medium. An excitation wavelength of 570 nm was found to be optimal to achieve high contrast between the SCN females and root tissues. Manual counts are made from a fluorescent image that is captured and displayed on a computer screen. Manual counting of females from the fluorescent images was found to be a reliable (≤14% difference) and more efficient (60% faster) method than manual counting of females in a dish.

Nonetheless, the method is not suitable for high-throughput SCN phenotyping of soybean in greenhouse screens and has limited efficiency due to a lack of automated counting.

The objective of this study was to develop a method for automated counting of SCN females using a fluorescence-based image analysis system that is more suited to high-throughput SCN phenotyping methods used in greenhouse screening.

Materials and Methods

Plant materials and growth conditions: Greenhouse screening methods for assessing soybean cyst nematode resistance in this study were conducted in accordance with the SCE08 protocol (Niblack et al., 2009). Seeds of *Glycine max* differing in their resistance to SCN were germinated in rolls of germination paper for two days at 27°C to a radical length of 4-6 cm prior to transplanting. Each soybean seedling was transplanted into a 6 cm hole made in sterile river sand in 100 cm^3 thin-walled polyvinyl carbonate (PVC) tubes (15.2 cm in length, 2.9 cm inside diameter) placed into a 5.7 liter Bains Marie plastic container (Continental Carlisle, Oklahoma City, Oklahoma) and inoculated with SCN eggs as described below. The containers were immersed in a temperature-controlled water bath set at 27.5°C in the greenhouse. The plants were maintained for 30 days under 12-hour supplemental lighting in the greenhouse and fertilized weekly with Miracle-Gro (Scotts Miracle-Gro, Marysville, Ohio).

Nematode inoculation and harvesting: Cysts of *Heterodera glycines* were extracted from infested soil by flotation in water and collected on a no. 60-mesh (250-µm) sieve. Harvested cysts were crushed gently using a drill press and the eggs collected on a no. 500 (25-µm) sieve (Faghihi and Ferris, 2000). Eggs were further purified by centrifugal flotation on a sucrose density gradient. Individual

soybean seedlings were inoculated with 1500-2000 eggs in one ml of water. At 30 days post-inoculation, the tops of the soybean plants were removed and each PVC tube was soaked in a liter of water until the tube could easily be removed from the root. Each root system was gently agitated in the water until the root was free of sand. For automated imaging of SCN females directly on roots, the root system was placed into a plastic cup to allow excess water to drain from the roots and air dried for 30 minutes. Roots were blotted gently with paper towels prior to imaging. Alternatively, SCN females were blasted with a water sprayer on a no. 20-mesh (750 µm) over no. 60-mesh stack of sieves. Females with a small amount of root debris were then rinsed into a labeled 5.5 cm diameter Petri dish.

Imaging system: A Kodak Image Station 4000MM Pro (Carestream Health, Inc. Rochester, NY) equipped with a cooled 4-million pixel CCD camera, 16-bit imaging and an automated precision lens, paired with Carestream Molecular Imaging Software version 5.0.2 (Carestream Health, Inc. Rochester, NY) was used for imaging and automated counting. Fluorescent excitation and emission filters of 470 nm and 535 nm, respectively, were used for all samples.

Imaging and automated counting of SCN females in Petri dishes and on roots: For imaging and automated counting of SCN females in Petri dishes, three Petri dishes with SCN females were placed on the center of the scanning bed in a triangular formation. The Petri plates were then scanned and fluorescent images captured using the following settings: Field of View (FOV): 111.73 mm2; Signal to Noise: 16; Shape of reference: Oval; Width:15, Height:15; Exposure time: 30 seconds.

For imaging and automated counting of SCN females on roots, root systems were individually placed on the scanner. Plastic tweezers were used as needed to aid in spreading out the roots and the root system was then weighted down by a black plate to improve focus. Roots were then scanned and fluorescent images captured using the following settings: Field of View (FOV): 100.30 mm2; Signal to Noise: 15; Shape of reference: Oval; Width:15 Height:15; Exposure time: 30 seconds.

The signal to noise setting compares the level of a desired signal (e.g., fluorescence of SCN females) to the level of background noise. The higher the signal to noise ratio (range 1-20), the more stringent the search. The shape of reference sets the reference region of interest to an oval. Width and height settings establish the number of pixels scanned per major and minor axes of the interior pixels of the oval. The program creates an oval around peaks of intensity emitted from the images using the width and height settings. The program then uses the signal intensity found in that area to calculate the signal emanating from the region of interest.

To obtain automated counts of SCN females in Petri dishes, the area of each dish was outlined. Root images were counted by selecting the entire image for analysis. Fluorescent females were counted within the selected area using the automated region of interest (Auto ROI) function of the imaging software with the aforementioned settings, which uses light intensity at the emitting wavelength of 535 nm and shape, to select the females. Auto ROI counts were recorded and saved with each image.

Statistical analysis: A correlation analysis was performed with the statistical computing package SAS (SAS Institute, Inc, Cary, NC). The relationship between manual SCN female counts and automated SCN females counts either directly on roots or in Petri dishes was determined by plotting manual count data for individual plants against automated count data for each experiment. SAS was used to obtain all graphs, correlation coefficients r, and the estimated regression coefficients.

Results

Optimization of an automated fluorescence-based imaging system to count *H. glycines* females: The fluorescent imaging system was tested using multiple combinations of wavelengths to identify the best setting to achieve the greatest contrast between SCN females and soybean roots. An excitation filter of 470 nm and emission filter of 535 nm was found to be the best combination available.

To optimize the scanner for counting SCN females in Petri dishes, dishes with known cyst counts were placed on the scanning bed in various formations and imaged using different settings. The scanning bed can hold up to six 5.5 cm Petri dishes simultaneously; however, a triangular formation of three Petri dishes set to a FOV of 111.73 mm2, which focused the camera on an area slightly bigger than the area occupied by the three Petri dishes, was found to be optimal (Figure A5, A-B). A signal to noise ratio set to 16 provided the most accurate result. This setting was stringent enough to avoid picking up any erroneous

background noise created by root debris, but not too stringent to underestimate the SCN female counts.

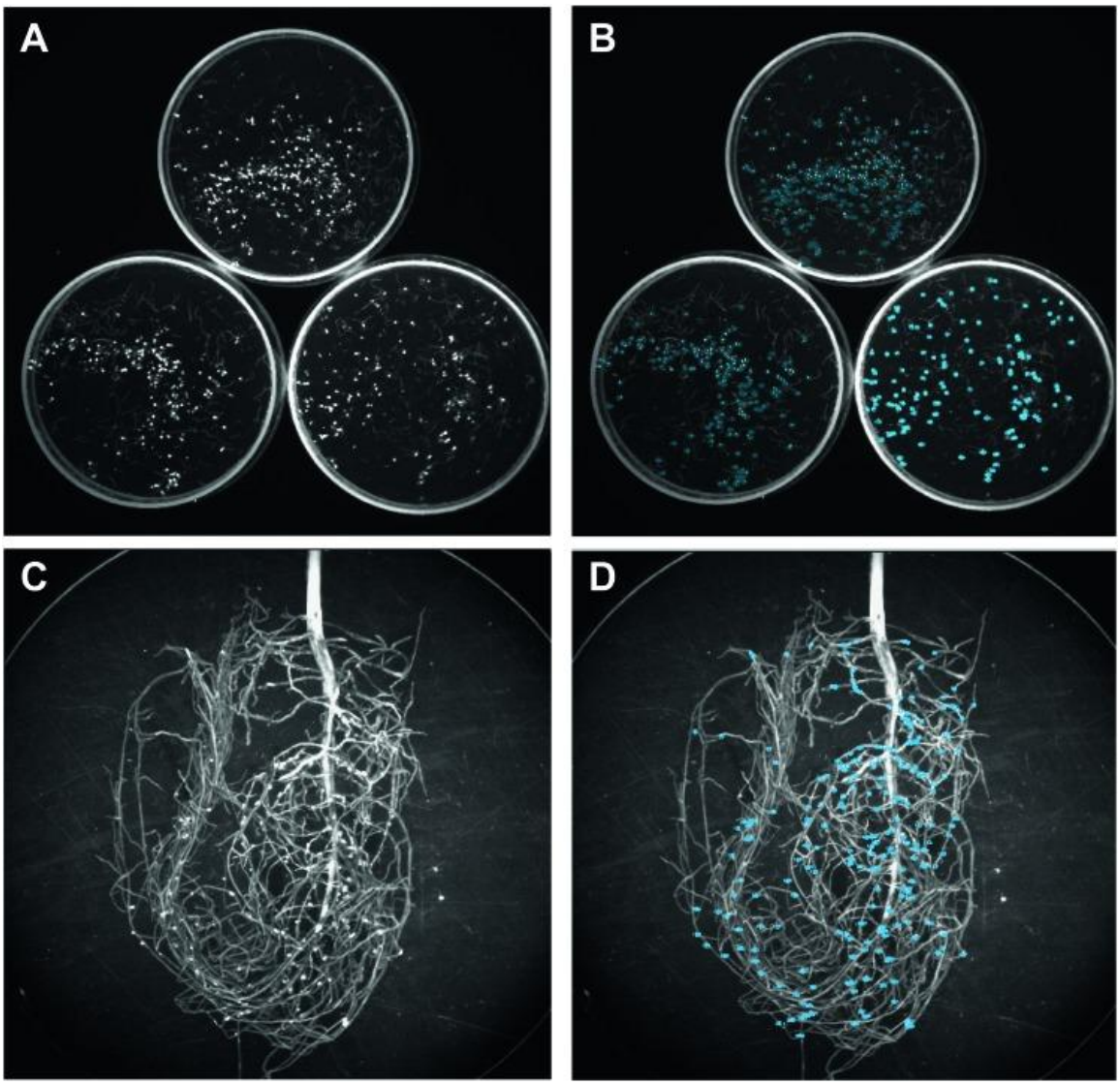


Figure A5. Fluorescent cyst images captured by the Kodak Image Station 4000MM Pro and analyzed with Carestream Molecular Imaging Software. (A) Petri dish image with fluorescing SCN females prior to automated region of interest (Auto ROI) search. (B) Petri dish after Auto ROI search, blue circles represent identified SCN females. (C) Whole root image with fluorescing SCN females prior to Auto ROI search (D) Whole root image after Auto ROI search, blue circles represent identified SCN females.

To optimize the scanner for counting SCN females directly on roots, individual root systems were placed on the scanning bed and imaged using different settings (Figure A5, C-D). Females were then washed from the root systems and counted manually to obtain actual counts for optimizing settings. A FOV of 100.30 mm2, which focused the camera on an area slightly bigger than the area occupied by the root system, was found to be optimal. A signal to noise ratio set to 15 provided the most accurate result.

The size and shape of the cyst were used along with the intensity of the light to obtain the cyst counts in the Auto ROI portion of the software and adjusted to the best fit using the signal to noise setting (Figure A5, B,D).

Automated fluorescence-based counting of SCN females in Petri dishes is accurate: To assess the accuracy of the automated fluorescence-based imaging system for counting *H. glycines* females, we determined the coefficients of determination (r^2) between female numbers obtained using fluorescence-based imaging on roots and in Petri dishes with that of manual counting in Petri dishes under a microscope. The root systems of 77 individual soybean plants were first scanned and females counted using the fluorescence-based imaging system; the females were then blasted from the roots into Petri dishes and again scanned and counted using the imaging system. Lastly, the SCN females in the Petri dishes were manually counted under a microscope. The coefficient of determination between automated fluorescence-based counting of SCN females on roots and manual counting of SCN females in Petri dishes under the microscope was determined to be 0.76 (Figure A6, A). The coefficient of

determination between automated fluorescence-based counting of SCN females in Petri dishes and manual counting of SCN females in Petri dishes under the microscope was determined to be 0.95 (Figure A6, B). Two additional independent experiments comparing fluorescence-based counting of SCN females in Petri dishes with manual counting of SCN females in Petri dishes under the microscope yielded coefficients of determination of 0.99 (n = 105) and 0.97 (n = 89) (data not shown). One additional independent experiment comparing fluorescence-based counting of SCN females on roots with manual counting of SCN females in Petri dishes under the microscope using a larger sample size (n = 212) yielded a coefficient of determination of 0.85 (data not shown). These data indicate a much higher degree of accuracy in female counting using automated fluorescence-based counting of SCN females in Petri dishes versus directly on roots.

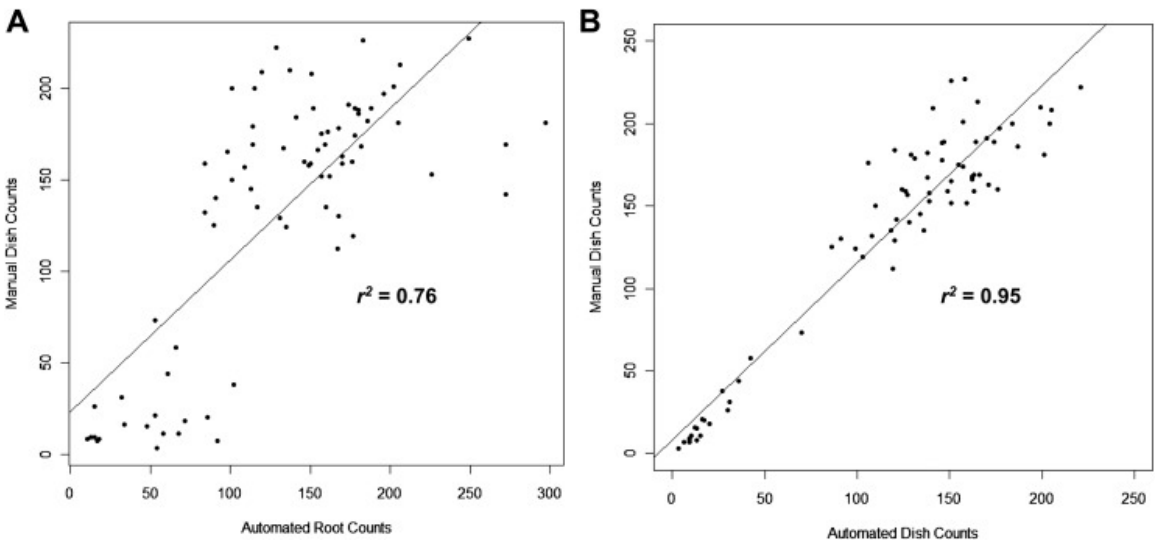


Figure A6. Scatter plots depicting the relationship between manual counts of SCN females in Petri dishes and automated SCN female counts directly on roots. (A) or automated SCN female counts in Petri dishes (B). Each point corresponds to counts from an individual plant. Coefficients of determination (r2) are indicated.

Automated fluorescence-based counting is more efficient than manual counting: To determine whether automated fluorescence-based counting of SCN females in Petri dishes is more efficient than manual counting of SCN females in Petri dishes under the microscope, we compared the speed of the two methods. Fifteen Petri dishes containing SCN females (avg 166 females) were counted in groups of three using the automated fluorescence-based imaging system or counted manually under a microscope and timed. Manual counting of SCN females in three Petri dishes under a microscope by an experienced counter took on average four minutes and forty-three seconds, whereas counts for three dishes using the imaging system took on average two minutes and fifteen

seconds (Table A1). Thus, automated fluorescence-based counting was more than 50% faster than manual counting of SCN females under a microscope.

	Dishl	Dish 2	Dish 3	Total Time (min:sec)
Manual Counting				
Set 1	133	160	164	4:45
Set 2	156	183	161	4:22
Set 3	151	198	200	4:51
Set 4	183	161	143	4:47
Set 5	182	161	162	4:48
				Avg = 4.43
Automated Counting				
Set 1	116	147	148	2:17
Set 2	127	178	122	2:18
Set 3	151	201	214	2:07
Set 4	176	181	145	2:10
Set 5	172	169	160	2:23
				Avg = 2.15

Table A1. Comparison of female counting times using automated and manual methods.

Discussion

In this paper, we report on a method for automated counting of SCN females using a fluorescence-based imaging system that is adapted for high-throughput screening of *H. glycines*-infected samples in soil. This is a significant improvement over the current method of screening *H. glycines*-infected samples in soil which requires manual counting of SCN females under a microscope. Use of the imaging system is more than 50% faster than microscope counting. In addition, this method provides a significant improvement over an existing method reported for SCN female counting using fluorescence-based imaging (Lu et al., 2005). The method reported by Lu et al (2005) was only adapted for counting SCN females grown on root explants in Petri dishes. The approach is useful for soybean hairy root assays, but has not been adapted for counting SCN females on soil-grown roots. The method is also low-throughput. Using this method, only one Petri dish can be imaged at a time and the SCN females still need to be counted manually from a fluorescent image displayed on a computer screen.

Methods to phenotype soybean varieties for resistance to SCN are currently very laborious and time consuming. This is due in part to the 30-day life cycle of SCN; an aspect of the bioassay that cannot be accelerated. Another step in the process of SCN phenotyping of soybean plants that is both laborious and time consuming is the manual counting of females. Thus, automated counting methods could greatly improve the efficiency of SCN bioassays. Here we optimized a fluorescence-based imaging system for automated counting of SCN females and compared two approaches for accuracy. One approach was to

image and count SCN females directly on soybean roots and the other approach was to image and count SCN females that had been blasted from roots in Petri dishes. In a direct comparison for accuracy, the automated counting of SCN females in Petri dishes was found to be a superior method over automated counting of SCN females directly on roots. The most likely factor contributing to the reduced accuracy of on root scanning is the different shape, size, and dimension of each root system compared to that of a Petri dish, which may contribute to the calibration being slightly off from sample to sample. Factors such as the irregular distribution of SCN females throughout the root system and the two-dimensional scanning may have also contributed to the reduced accuracy of the latter method. In some cases, we observed cultivar differences in the amount of fluorescing root tissues leading to false counts and an overestimation of female counts. In general, we noted minimal discrepancy in female counts when the same root system was flipped over and rescanned, indicating that most females are detected when scanned from one side. However, the thicker the root system, the greater the chance of underestimating female counts due to hidden females and reduced fluorescent signals from the females on the opposite side of the root system. To improve the accuracy of female counting on the roots, a visual inspection of each image can be undertaken; however, this is time consuming and dramatically decreases the efficiency of sample processing. In contrast to counting females directly on roots, automated counting of SCN females in Petri dishes is not restricted by the aforementioned limitations and was found to be highly accurate (avg $r^2 = 0.97$).

Although this approach requires the additional step of blasting the SCN females from the roots, which takes approximately one minute per root sample including washing the sieves between samples, the quality of the data is very reliable.

The approach used will depend on the objective of the study. For example, for early stages of germplasm screening to reduce the number of lines advanced to the next generation, faster, but less accurate SCN female counts may be acceptable. In this case, on root imaging and automated counting may be acceptable. On the other hand, SCN phenotyping of mapping populations for the identification and cloning of SCN resistance genes requires a much higher degree of accuracy so that genotypic information can be correlated to phenotype. In this case, automated counting of SCN females in Petri dishes would be required. Another application for automated counting of SCN females would be its use in HG Type testing. In this case, a combination of automated counting of SCN females in Petri dishes and manual counting may be necessary. Although the FI should always be reported with HG Type designations (Niblack et al., 2002), it would be prudent to manually count the samples yielding an FI near the threshold of 10%. For example, if the susceptible line had 100 SCN females and the average SCN female count on PI 88788 was determined to be 12 using the imaging system, this constitutes a female index of 12% and the HG Type designation would include a “2”. However, if the imager had over or undercounted by just a few SCN females the HG Type designation would change.

In summary, the automated fluorescence-based imaging system described herein for counting SCN females from infected root samples grown in soil is well-suited for high-throughput screening of soybean germplasm for SCN resistance. Application of this method will 1) improve the efficiency and accuracy of *H. glycines* resistance breeding programs to identify new sources of resistance and develop SCN-resistant cultivars, 2) streamline SCN phenotyping of mapping populations to clone SCN resistance genes in soybean, and 3) provide a more efficient process of testing transgenic plants for SCN resistance.

LITERATURE CITED

Acedo JR, Dropkin VH, Luedders VD, (1984) Nematode population attrition and histopathology of *Heterodera glycines*-soybean associations. *J Nematol* **16**, 48-56.

Alkharouf N, Khan R, Matthews B, (2004) Analysis of expressed sequence tags from roots of resistant soybean infected by the soybean cyst nematode. *Genome* **47**, 380-8.

Bendahmane A, Kanyuka K, Baulcombe DC, (1999) The *Rx* gene from potato controls separate virus resistance and cell death responses. *Plant Cell* **11**, 781-92.

Black DJ, Halling DB, Mandich DV, Pedersen SE, Altschuld RA, Hamilton SL, (2005) Calmodulin interactions with IQ peptides from voltage-dependent calcium channels. *Am J Physiol Cell Physiol* **288**, C669-76.

Briknarova K, Takayama S, Brive L*, et al.*, (2001) Structural analysis of BAG1 cochaperone and its interactions with Hsc70 heat shock protein. *Nat Struct Biol* **8**, 349-52.

Brucker E, Carlson S, Wright E, Niblack T, Diers B, (2005) *Rhg1* alleles from soybean PI 437654 and PI 88788 respond differentially to isolates of *Heterodera glycines* in the greenhouse. *Theoretical and Applied Genetics* **111**, 44-9.

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.

Clough SJ, Bent AF, (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**, 735-43.

Concibido VC, Diers BW, Arelli PR, (2004) A decade of QTL mapping for cyst nematode resistance in soybean. *Crop Science* **44**, 1121-31.

Conkling, M.A., Opperman, C.H., and Taylor, C.G. United States Patent #5,750,386. Pathogen resistant transgenic crop cultivars. Issued, 12 May 1998.

David E. Cook, Tong Geon Lee, Xiaoli Guo, et al., (2012) Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean *Science*: doi:10.1126/science.1228746.

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR, (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* **139**, 5-17.

Doong H, Vrailas A, Kohn EC, (2002) What's in the 'BAG'?--A functional domain analysis of the BAG-family proteins. *Cancer Lett* **188**, 25-32.

Doukhanina EV, Chen S, Van Der Zalm E, Godzik A, Reed J, Dickman MB, (2006) Identification and functional characterization of the BAG protein family in *Arabidopsis thaliana*. *J Biol Chem* **281**, 18793-801.

Endo BY, (1965) Histological responses of resistant and susceptible soybean varieties, and backcross progeny to entry and development *of Heterodera glycines*. *Phytopathology* **55**, 375-81.

Endo BY, (1986) Ultrastructure and function of cyst nematodes. In. *NATO ASI Series.* New York: Plenum, 47-73. (Series A: Life Sciences; vol. 121.)

Gassler CS, Wiederkehr T, Brehmer D, Bukau B, Mayer MP, (2001) Bag-1M accelerates nucleotide release for human Hsc70 and Hsp70 and can act concentration-dependent as positive and negative cofactor. *J Biol Chem* **276**, 32538-44.

Goodstein DM, Shu S, Howson R*, et al.*, (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* **40**, D1178-86.

Huang G, Allen R, Davis EL, Baum TJ, Hussey RS, (2006) Engineering broad root-knot resistance in transgenic plants by RNAi silencing of a conversed and essential root-knot nematode parasitism gene. *PNAS* **103**, 14302-6.

Hussey RS, Grundler FM, (1998) Nematode parasitism of plants. In: Perry RN, Wright J, eds. *Physiology and biochemistry of free-living and plant parasitic nematodes.* England: CAB International Press, 213-43.

Ingham RJ, Colwill K, Howard C*, et al.*, (2005) WW domains provide a platform for the assembly of multiprotein networks. *Mol Cell Biol* **25**, 7092-106.

Jefferson AR, (1987) Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Molecular Biology Reporter* **5**, 387-405.

Kandoth PK, Ithal N, Recknor J*, et al.*, (2011) The soybean *Rhg1* locus for resistance to the soybean cyst nematode *Heterodera glycines* regulates the expression of a large number of stress- and defense-related genes in degenerating feeding cells. *Plant Physiol* **155**, 1960-75.

Kang CH, Jung WY, Kang YH*, et al.*, (2006) AtBAG6, a novel calmodulin-binding protein, induces programmed cell death in yeast and plants. *Cell Death Differ* **13**, 84-95.

Kim Y, Kim K, Riggs R, (2010) Differential subcellular responses in resistance soybeans infected with soybean cyst nematode races. *Plant Pathology Journal* **26**, 154-8.

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**, 525-67.

Klink VP, Overall CC, Alkharouf NW, Macdonald MH, Matthews BF, (2007) 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**, 1389-409.

Koenning SR, Wrather JA, (2010) Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. *Plant Health Progress*.

Levy M, Wang Q, Kaspi R, Parrella MP, Abel S, (2005) *Arabidopsis* IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *Plant J* **43**, 79-96.

Li XQ, Wei JZ, Tan A, Aroian RV, (2007) Resistance to root-knot nematode in tomato roots expressing a nematicidal *Bacillus thuringiensis* crystal protein. *Plant Biotechnol J* **5**, 455-64.

Li Y, Chen S, Young ND, (2004) Effect of the *rhg1* on penetration, development and reproduction of *Heterodera glycines* race 3. *Nematology* **6**, 729-36.

Libault M, Farmer A, Joshi T*, et al.*, (2010) An integrated transcriptome atlas of the crop model Glycine max, and its use in comparative analyses in plants. *Plant J* **63**, 86-99.

Libault M, Joshi T, Takahashi K*, et al.*, (2009) Large-scale analysis of putative soybean regulatory gene expression identifies a Myb gene involved in soybean nodule development. *Plant Physiol* **151**, 1207-20.

Liu, S, Kandoth, P, Warren S, Yeckel G, et al., (2012) A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature:* doi:10.1038/nature11651.

Liu X, Liu S, Jamai A*, et al.*, (2011) Soybean cyst nematode resistance in soybean is independent of the *Rhg4* locus LRR-RLK gene. *Funct Integr Genomics* **11**, 539-49.

Lu H, Tallman J, Hu X, Anderson E, Chamberlin M, Lu G, (2005) An innovative method for counting females of soybean cyst nematode with fluorescence imaging technology. *J Nematol* **37**, 495-9.

Luders J, Demand J, Hohfeld J, (2000) The ubiquitin-related BAG-1 provides a link between the molecular chaperones Hsc70/Hsp70 and the proteasome. *J Biol Chem* **275**, 4613-7.

Mahalingam R, Skorupska HT, (1996) Cytological expression of early response to infection by *Heterodera glycines* Ichinohe in resistant PI 437654 soybean. *Genome* **39**, 986-98.

Matson A, Williams L, (1965) Evidence of a fourth gene for resistance to the soybean cyst nematode. *Crop Science* **5**.

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.

Mitchum MG, Sukno S, Wang XH*, et al.*, (2004) The promoter of the *Arabidopsis thaliana Cel1* endo-1,4-beta glucanase gene is differentially expressed in plant feeding cells induced by root-knot and cyst nematodes. *Molecular Plant Pathology* **5**, 175-81.

Mitchum MG, Wrather JA, Heinz RD, Shannon JG, Danekas G, (2007) Variability in distribution and virulence phenotypes of *Heterodera glycines* in Missouri during 2005. *Plant Disease* **91**, 1473-6.

Moribe Y, Niimi T, Yamashita O, Yaginuma T, (2001) Samui, a novel cold-inducible gene, encoding a protein with a BAG domain similar to silencer of death domains (SODD/BAG-4), isolated from *Bombyx diapause* eggs. *Eur J Biochem* **268**, 3432-42.

Mudge **J**, (1999) Marker assisted selection for soybean cyst nematode resistance and accompanying agronomic traits. *PHD Thesis* **University of Minnesota, St, Paul**.

Niblack TL, Arelli PR, Noel GR*, et al.*, (2002) A revised classification scheme for genetically diverse populations of *Heterodera glycines*. *Journal of Nematology* **34**, 279-88.

Niblack, T. L., Colgrove, A. L., Colgrove, K., and Bond, J. P. (2008) Shift in virulence of soybean cyst nematode is associated with use of resistance from PI 88788. *Plant Health Progress*: doi:10.1094/PHP-2008-0118-01-RS.

Niblack TL, Heinz RD, Smith GS, Donald PA, (1993) Distribution, density, and diversity of *Heterodera glycines* in Missouri. *J Nematol* **25**, 880-6.

Opperman CH, Taylor CG, Conkling MA, (1994) Root-knot nematode--directed expression of a plant root--specific gene. *Science* **263**, 221-3.

Pascale M, Rosati A, Festa M, Basile A, D'avenia M. , Falco A, Torino G. , Turco M, (2010) BAG3 protein: role in some neoplastic cell types and Identification as a candidate target for therapy. *Apoptosome*, 137-46.

Petrezselyova S, Zahradka J, Sychrova H, (2010) *Saccharomyces cerevisiae* BY4741 and W303-1A laboratory strains differ in salt tolerance. *Fungal Biol* **114**, 144-50.

Prior TI, Kunwar S, Pastan I, (1996) Studies on the activity of barnase toxins in vitro and in vivo. *Bioconjug Chem* **7**, 23-9.

Rana R, Dong S, Khan A, Zhang H, (2012) Identification and characterization of the Bcl-2-associated anthanogene (BAG) protein family in rice. *African Journal of Biotechnology* **11**, 88-98.

Reed JC, (1994) Bcl-2 and the regulation of programmed cell death. *J Cell Biol* **124**, 1-6.

Replogle AJ, Wang J, Paolillo V*, et al.*, (2012) Synergistic Interaction of CLAVATA1, CLAVATA2, and RECEPTOR-LIKE PROTEIN KINASE 2 in cyst nematode parasitism of *Arabidopsis*. *Mol Plant Microbe Interact*.

Rhoads AR, Friedberg F, (1997) Sequence motifs for calmodulin recognition. *FASEB J* **11**, 331-40.

Sacco MA, Mansoor S, Moffett P, (2007) A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant J* **52**, 82-93.

Sawitzki B, Lehmann M, Vogt K*, et al.*, (2002) Bag-1 up-regulation in anti-CD4 mAb treated allo-activated T cells confers resistance to apoptosis. *Eur J Immunol* **32**, 800-9.

Schmutz J, Cannon SB, Schlueter J*, et al.*, (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**, 178-83.

Sijmons PC, Grundler FMW, Vonmende N, Burrows PR, Wyss U, (1991) *Arabidopsis thaliana* as a new model host for plant parasitic nematodes. *Plant Journal* **1**, 245-54.

Sondermann H, Scheufler C, Schneider C, Hohfeld J, Hartl FU, Moarefi I, (2001) Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. *Science* **291**, 1553-7.

Takayama S, Reed JC, (2001) Molecular chaperone targeting and regulation by BAG family proteins. *Nat Cell Biol* **3**, E237-41.

Takayama S, Sato T, Krajewski S*, et al.*, (1995) Cloning and functional analysis of BAG-1: a novel Bcl-2-binding protein with anti-cell death activity. *Cell* **80**, 279-84.

Takayama S, Xie Z, Reed JC, (1999) An evolutionarily conserved family of Hsp70/Hsc70 molecular chaperone regulators. *J Biol Chem* **274**, 781-6.

Thress K, Henzel W, Shillinglaw W, Kornbluth S, (1998) Scythe: a novel reaper-binding apoptotic regulator. *EMBO J* **17**, 6135-43.

Townsend PA, Stephanou A, Packham G, Latchman DS, (2005) BAG-1: a multi-functional pro-survival molecule. *Int J Biochem Cell Biol* **37**, 251-9.

Urwin PE, Lilley CJ, Mcpherson MJ, Atkinson HJ, (1997a) Characterization of two cDNAs encoding cysteine proteinases from the soybean cyst nematode *Heterodera glycines. Parasitology* **114**, 605-13.

Urwin PE, Lilley CJ, Mcpherson MJ, Atkinson HJ, (1997b) Resistance to both cyst and root-knot nematodes conferred by transgenic *Arabidopsis* expressing a modified plant cystatin. *Plant Journal* **12**, 455-61.

Urwin PE, Mcpherson MJ, Atkinson HJ, (1998) Enhanced transgenic plant resistance to nematodes by dual proteinase inhibitor constructs. *Planta* **204**, 472-9.

Van Petegem F, Chatelain FC, Minor DL, Jr., (2005) Insights into voltage-gated calcium channel regulation from the structure of the CaV1.2 IQ domain-Ca2+/calmodulin complex. *Nat Struct Mol Biol* **12**, 1108-15.

Veal EA, Ross SJ, Malakasi P, Peacock E, Morgan BA, (2003) Ybp1 is required for the hydrogen peroxide-induced oxidation of the Yap1 transcription factor. *J Biol Chem* **278**, 30896-904.

Vrain TC, (1999) Engineering natural and synthetic resistance for nematode management. *J Nematol* **31**, 424-36.

Wang J, Replogle A, Hussey R*, et al.*, (2011) Identification of potential host plant mimics of CLAVATA3/ESR (CLE)-like peptides from the plant-parasitic nematode *Heterodera schachtii*. *Mol Plant Pathol* **12**, 177-86.

Wang X, Replogle A, Davis EL, Mitchum MG, (2007) The tobacco *Cel7* gene promoter is auxin-responsive and locally induced in nematode feeding sites of heterologous plants. *Mol Plant Pathol* **8**, 423-36.

Wei JZ, Hale K, Carta L*, et al.*, (2003) *Bacillus thuringiensis* crystal proteins that target nematodes. *Proc Natl Acad Sci U S A* **100**, 2760-5.

Williams B, Kabbage M, Britt R, Dickman MB, (2010) *AtBAG7,* an *Arabidopsis* Bcl-2-associated athanogene, resides in the endoplasmic reticulum and is involved in the unfolded protein response. *Proc Natl Acad Sci U S A* **107**, 6088-93.

Wrather J, Anand S, Dopkin V, (1984) Soybean cyst nematode control. *Plant Disease* **68**.

Wrather JA, Koenning SR, (2006) Estimates of disease effects on soybean yields in the United States 2003 to 2005. *Journal of Nematology* **38**, 173-80.

Yadav BC, Veluthambi K, Subramaniam K, (2006) Host-generated double stranded RNA induces RNAi in plant-parasitic nematodes and protects the host from infection. *Molecular and Biochemical Parasitology* **148**, 219-22.

Yan J, He C, Zhang H, (2003) The BAG-family proteins in *Arabidopsis thaliana*. *Plant Science* **165**.

Youle RJ, Strasser A, (2008) The BCL-2 protein family: opposing activities that mediate cell death. *Nat Rev Mol Cell Biol* **9**, 47-59.

Zhang C, Bradshaw JD, Whitham SA, Hill JH, (2010) The development of an efficient multipurpose bean pod mottle virus viral vector set for foreign gene expression and RNA silencing. *Plant Physiol* **153**, 52-65.