Proteomic and Metabolite Analysis of Nematode Resistance

David A. Lightfoot

1 Department of Molecular Biology, Microbiology and Biochemistry, Southern Illinois University at Carbondale, IL, 62901, USA.
2 Genomics Core Facility; Department of Plant Soil and Agricultural Systems, and Center for Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale, Carbondale, IL 62901-4415, USA.
This research was funded over the past 11 years in part by grants from the NSF 04-05819 and USB 6221. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The past and continuing support of SIUC, Office of the Vice Chancellor for Research to MJI, JA and DAL is appreciated.
The Nematode Resistance Allele at the rhl1 Locus Alters the Proteome and Primary Metabolism of Soybean Roots.

Ahmed J. Afzal 1,2,5, Aparna Natarajan 1, Navinder Saini 2, M. Javed Iqbal 3, Matt Geisler 4 Hany A. El Shemy 1,2,6, Rajsree Mungur 1,7, Lothar Willmitzer 7 and David A. Lightfoot 2

1 Department of Molecular Biology, Microbiology and Biochemistry, Southern Illinois University at Carbondale, IL, 62901, USA.

2 Genomics Core Facility; Department of Plant Soil and Agricultural Systems, and Center for Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale, Carbondale, IL 62901-4415, USA.

3 Institute for Advanced Learning and Research (ISRR), Institute for Sustainable and Renewable Resources (IALR), 150 Slayton Ave, Danville, VA 24540, USA.

4 Department of Plant Biology, Southern Illinois University at Carbondale, IL, 62901, USA.

5 Present address: Department of Horticulture and Crop Science, 2021 Coffey Rd, Ohio State University, Columbus, OH, 43210

6 Present address: Faculty of Agriculture Research Park (FARP) and Department of Biochemistry, Faculty of Agriculture, Cairo University, 12613 Giza, Egypt.

7. Max Planck Institute for Molecular Plant Physiology, Am Muehlenberg 1, Potsdam 14476, Germany.
Abstract

*Heterodera glycines* I., the cyst nematode (SCN), causes the most damaging chronic disease of soybean. Host resistance requires the resistance allele at *rhl1*. Resistance destroys the giant cells created in the plants roots by the nematodes about 24-48 h after commencement of feeding. In addition, 4-8 days later a systemic acquired resistance (SAR) develops that discourages later infestations. The molecular mechanisms that control the *rhl1* mediated resistance response appear to be multigenic and complex judged by transcript abundance changes, even in near isogenic lines (NILs). This study aimed to focus on key post-transcriptional changes by identifying proteins and metabolites that were increased in abundance in both resistant and susceptible NILs. Comparisons were made among NILs 10 days after SCN infestation and without SCN infestation. Two dimensional gel electrophoresis resolved more than 1,000 protein spots on each gel. Only thirty protein spots with a significant (P<0.05) difference in abundance of 1.5 fold or more were found among the four treatments. The proteins in these spots were picked, trypsin digested and analyzed using quadruple time-of-flight (Q-TOF) tandem mass spectrometry. Protein identifications could be made for 24 out of the 30 spots. Four spots contained two proteins so that 28 distinct proteins were identified. The proteins were grouped into six functional categories. Metabolite analysis by GC-MS identified 131 metabolites among which 58 were altered by one or more treatment, 28 were involved in primary metabolisms. Taken together the data showed seventeen pathways that were altered by *rhl1* controlled metabolisms associated with SAR-like responses including xenobiotic, phytoalexin, ascorbate and inositol metabolism as well as primary metabolisms like amino acid metabolism and glycolysis. The pathways impacted by the *rhl1* allelic state and SCN infestation agreed with transcript abundance analyses but identified a smaller set of key proteins. Six of the proteins lay within the same region of the interactome identifying a key set of 159 proteins. Finally, two proteins (glucose 6 phosphate isomerase, EC 5.3.1.9; and isoflavone reductase, EC 1.3.1.45) and two metabolites (maltose and unidentified) differed in resistant and susceptible NILs without SCN infestation and may form the basis of a new assay for the selection of resistance to SCN in soybean.

**Keywords:** soybean cyst nematode (SCN), nematode resistance, proteomics, soybean roots.
Introduction

Soybean  (*Glycine max* L. Merr.) is the world’s most important legume crop, it is grown for both oil and protein co-products (Liu, 1997). Many natural by-products are also derived from soybean and several possess significant antimicrobial and pharmacological properties (Dixon and Sumner 2003). The abundance and variety of soybean’s bioactive factors suggest the species has evolved and been domesticated in the presence of many persistent pathogen populations.

After water deficit, the leading causes of seed yield losses in soybean are biotic stresses from pathogen infections (Wrather et al. 2001). Amongst soybean pathogens, the soybean cyst nematode (SCN) causes the most damage. Yield losses in the world due to SCN infection have approached $2 billion a year (Niblack et al. 2004) and in the US alone amounted to 1.5 billion for the 1996-1997 period (Chen et al. 2001). SCN preventative measures include crop rotation, the use of resistant cultivars and nematicides. The use of resistant cultivars provides the most efficient and practical control of SCN.

Soybean breeding has used heavily three plant introductions (PIs) resistant to the cyst nematode (Brucker et al. 2005b). The PIs ‘PI88788’ ‘Peking’ and ‘PI437654’ have each been shown to carry resistance loci effective against multiple nematode races. However, the resistance to SCN is quantitative and controlled by multiple loci. Previously identified were two quantitative trait loci that provided substantial resistance against the soybean cyst nematode Hg Type 0 (a biotype of race 3) in NILs derived from Peking via ‘Forrest’ (Meksem et al. 2001; Ruben et al. 2006). The resistance locus *Rhg4* on linkage group A2 provided resistance against Hg Type 0 whereas *rhg1*, the resistance locus on linkage group G provided some resistance against nematodes of all known races. The *rhg1* locus allele from Peking causes a rapid necrosis of the giant cells and the formation of thickened cell walls around the defunct feeding site (syncitium) by 24-48h after the SCN female starts to feed, some 4-5 days after infestation of the roots (Alkharouf et al., 2004; 2006; Ithal et al. 2007a; 2007b). The activity of the resistance allele is associated with global changes in transcript abundance (TA) within the syncitium and systemic changes in the root system with hundreds of transcripts altered. However, in plants the correlations between transcript, protein and metabolite abundances are low. Therefore, key proteins may have been overlooked by transcript analysis. Further, some proteins can alter the metabolome without alterations in their abundances (Edwards 1996; Ito et al. 2003; Schad et al 2005; Mungur, 2008). Therefore, key metabolites may remain undetected. Finally, protein-
protein and protein-metabolite interactions are important in many plant responses to the environment (Geisler-Lee et al. 2008). Therefore, key interactions may remain undetected by the transcript analyses reported to date.

The *rhgl* locus appears to be a multifunctional and multigeneic locus (Ruben et al. 2006; Iqbal et al. 2009; Afzal et al. 2009a). The locus contained several genes capable of altering the transcriptome, proteome, metabolome and interactome including; a receptor like kinase inferred to alter root development (RLK; Afzal and Lightfoot 2007; Afzal et al. 2008a; 2008b; 2009a); an unusual laccase involved in phenolic metabolism (Iqbal et al. 2009); and a predicted Na/H antiporter. Flanking the locus were two proteins of unknown function, one of which is transcribed in roots (Triwitayakorn et al. 2005; Ruben et al. 2006). Each of the genes was encoded different alleles with significant amino acid and nucleotide changes in resistant soybean compared to susceptible soybean (Ruben et al. 2006; Afzal et al. 2008b; 2009a). Analysis in near isogeneic lines suggests the resistance alleles of the core set of three linked genes cannot be separated by recombination without the loss of resistance (Afzal 2007; Afzal et al. 2008b; 2009a).

Recent progress in the plant disease resistance field has shed light on some R proteins and their effector molecules (Martin et al. 2003; Afzal et al. 2008a) but the full set of interacting partners implicated in resistance have not been completely decoded. A number of systematic studies have been undertaken to identify the transcriptional profile as a consequence of SCN infection (Alkharouf et al. 2004; Alkharouf et al. 2006; Ithal et al. 2007; Klink et al. 2007a; Klink et al. 2007b). However, by 2009 there was no report that had addressed the difference in protein expression profiles in SCN infested compared to non-infested samples.

Proteomics databases and reference maps for *Medicago truncatula* (Lei et al. 2005) and soybean (Xu et al. 2006) have been constructed using two dimensional (2D) gel electrophoresis. The 2D gels have also been employed in the identification of proteins expressed during seed filing (Mooney et al. 2004; Hajduch et al. 2005); symbiotic nitrogen fixation (El Yahyaoui et al. 2004); for the detection of glycoproteins using iodine labeled lectins (Basha et al. 1981); and to explore the consequences of biotic stresses (Lei et al. 2005). However, reports of proteomic analysis of the pathogen and plant relationship have been sparse. These few include infection of root hairs with *Bradyrhizobium japonicum* (Wan et al. 2005) and the analysis of pathogenesis
related proteins (PR-Ps) from tobacco necrosis virus infected soybean plants (Roggero and Pennazio 1989).

Recent proteomics projects were encouraged by marked improvements in 2D gel reproducibility. This was achieved mainly by the development of immobilized IPG strips (Gorg et al. 2000) and enhanced techniques to obtain soluble proteins (Molloy et al. 1998). Furthermore, image analysis software that automated statistical analysis eased the laborious process of spot quantification. In addition, in-gel protease digestion, refinement of ESI MS/MS, MALTI TOF MS and protein identification using various databases (NCBI, Protplot or organism specific EST databases) has improved. Accurate identification of femto-molar quantities of proteins is now routine.

The hypotheses explored in this paper were; that expression of the set of candidate genes at \textit{rhg1} might alter the abundance of other plant proteins encoded by unlinked genes; that the altered proteins might form enriched clusters within the interactome; and that protein abundance changes have effects at the metabolite layer. In order to test these hypotheses the key proteins underlying SCN resistance were identified by 2D gel electrophoresis of total proteins extracted from the roots of NILs with different alleles at \textit{rhg1} infested for 10 days with the SCN Hg type 0. Differences in protein abundances in response to both genotype and pathogen infestation were noted. TAs were determined for the genes corresponding to the differentially abundant protein spots. An interactome map generated on the basis of orthology between soybean and \textit{Arabidopsis thaliana} proteins helped identify a significant cluster of interacting proteins within the interactome.

**Results**

NILs polymorphic across the region encompassing the \textit{rhg1} locus were used for the study. The region between markers BARC-Satt214 (and Satt163) and Satt570 spans 12.7 cM (3.16 Mbp) with the candidate \textit{rhg1} gene at about 4 cM close to the center (Meksem et al. 2001; Song et al. 2004; Ruben et al. 2006). The NIL 34-3 was susceptible to the SCN Hg Type used and, in contrast, NIL 34-23 was resistant to SCN. NIL genotypes were confirmed by the TMD1 marker located in the intron of the RLK within the \textit{rhg1} locus (results not shown; Afzal et al. 2008b; 2009a). Infection of the roots with Hg Type 0 nematode at 10 days after infestation (dai) was confirmed by PCR amplification of the nematode 18S ribosomal gene from root extracted DNA (Figure 1).
Differentially abundant proteins and metabolites were identified from four two-way comparisons. The first was made between Hg Type 0 infested and non-infested SCN resistant NIL 34-23. The second comparison was the infested resistant NIL 34-23 and the infested susceptible NIL 34-3. The third comparison was made between the infested and non-infested susceptible NIL 34-3. The fourth comparison was between the non-infested resistant NIL 34-23 and non-infested susceptible NIL 34-3. These comparisons were made in an effort to find over-expressed proteins in the susceptible plants infested with SCN Hg Type 0 and proteins more abundant in the resistant NIL 34-23. Answers to three fundamental questions were sought; (a) which proteins and metabolites are differentially abundant in the resistant NIL plant roots as a consequence of SCN infestation; (b) which proteins and metabolites were at higher abundance in the infested resistant NIL plant roots and not in the infested susceptible NIL plant roots; and (c) were any proteins or metabolites at different abundance in resistant and susceptible NIL plant roots without SCN infestation?

**Proteome abundance analysis**

More than 1,000 protein spots were reproducibly resolved in each gel (Figure 2). The isoelectric point (pI) for the majority of the spots ranged from pH 4 to pH 8 whereas the molecular masses for most proteins were between 20 and 100 kDa (Figure 2). Spot comparisons for the two way analyses were subjected to the student’s t-test with a significance (P) value set at 0.05. A total of 29 spots with a 1.5 fold or higher abundance were chosen for further analysis. Nine spots were over represented in the resistant NIL-34-23 infested roots compared to non-infested roots (Table 1). Intensities for twelve spots varied between the infested resistant and susceptible NIL roots (34-23-I vs. 34-3-I comparison gels; Table 2). Four protein spots showed varied abundances in the susceptible infested roots compared to non-infested roots (34-3-I vs.34-3-NI gels; Table 3). Only 2 spots had higher abundance in the resistant non-infested compared to the susceptible non-infested roots (Table 4) showing that both NIL’s proteomes were closely isogeneic in the absence of pathogen infestation.

For 25 of the 28 spots analyzed, the theoretical and experimental pI and molecular weight deviations were within 20 % of the expected values (Tables 1 to 4). For two of the deviating spots (glucose 6-phosphate isomerase, EC 5.3.1.9; and glutathione-S-transferase, EC 2.5.1.18 ), the extent was marginally greater than 20% whereas for one spot (ascorbate oxidase; EC 1.10.3.3.) the deviation was significant (Table 2) because the predicted protein was a pre-protein.
The observed differences may be attributed to incorrect protein N or C terminus inferences, novel post translational modifications, or to length variation between the homologous proteins.

For correctly annotated proteins, in most cases, there is no more than 20 percent difference between experimental and theoretical molecular weight and pI values (Wan et al. 2005). However, setting 20% as an arbitrary cutoff may lead to the exclusion of some correctly assigned proteins that may migrate to unexpected positions on a 2D gel due to post translational modifications (Wang et al. 2005). Further, some spots yielded more than 1 protein, not an unusual occurrence for 2D analysis of total protein samples involving proteins with similar isoelectric points and molecular weights.

**Functional classification of proteins**

The soybean genome was not yet fully sequenced and annotated in 2009, about 1,200 sequence gaps remained; nevertheless most of the protein spots of different abundance were successfully identified using the DOE sequence, the soybean EST database, the NCBI or Swissprot databases. A couple of sequences had no match to an EST, nucleotide or protein sequence from soybean. These were identified on the basis of homology to proteins from the *A. thaliana* or the *M. truncatula* databases. After identification, the proteins were classified into functional categories (Bevin et al. 1998). The largest functional group was represented by proteins implicated in disease or defense responses (45%) followed by proteins involved in energy related metabolic functions (29%). About 11% of the proteins belonged to the protein destination and storage category (Figure 3).

**Discussion**

Among a significant sample of the high to moderate abundance proteins in soybean roots finding just 30 significant differences (3 %) was surprising. In comparison to metabolites, where 58 of 131 were altered (44 %), the number was very low. Transcript analyses in soybean genotypes that differ by both *rhg1* and *Rhg4* alleles also showed a larger proportion of differences (Alkharouf et al. 2004; Alkharouf et al. 2006; Ithal et al. 2007a; Klink et al. 2007b). Therefore, protein analysis may be focusing on key differences by eliminating TA changes that do not result in protein abundance changes. In addition the metabolite analyses may show important post-translational events adding to metabolite changes.
Proteins and metabolites altered in abundance in roots after SCN infestation in the resistant NIL

Among the 9 proteins found to be increased by more than 1.5 fold in the roots of infested compared to the non-infested SCN resistant NIL 34-23 (Table 1), were two glutathione-S-transferases (GST; EC. 2.5.1.18). GST 8 and GST 11 spots had significant difference in molecular weight and pI (Table 1). Metabolites in the glutathione pathway include the precursors glycine and alanine that were decreased in abundance (Table 5), however, glutathione was not detected. EST analysis of an 8 day SCN infected soybean root library also inferred a role for GST in resistance to SCN infection (Alkharouf et al. 2004). Similarly, comparative microarray analysis of SCN infected soybean roots showed increased GST transcripts by 8 dai (Ithal et al. 2007a; Klink et al. 2007). The GST pathway leads to the detoxification of certain compounds during periods of plant stress through conjugation with glutathione. Soybean GSTs mainly conjugate the predominant thiol, homoglutathione (Marrs 1996). Thus, soybean GTSs may be involved in detoxification reactions during nematode induced pathogenesis (Alkharouf et al. 2004). Puthoff and coworkers (2003) used an Arabidopsis Affymetrix Genechip™ to show that following sugar beet cyst nematode (BCN; H. schactii) infestation, three of the genes with altered TA encoded glutathione- S-transferases (GST). BCN and SCN are closely related species so the change in GSTs may be an example of an orthologous plant response to nematodes. However, the TA of GSTs in Arabidopsis respond to various stimuli (Chen and Singh 1999; Coleman et al. 1997; Marrs 1996) and so may be part of a general defense response.

The protein encoded by an EST (GI31468505) was increased in abundance 2.5 fold. The protein did not show similarity to any protein of known function so it was not in any pathway at Mapman (Thimm et al. 2004). The source EST library was generated after salicylic acid induction (SA sprayed daily at 2.0 mM for 4 days) in whole 14 day old seedlings from the SCN susceptible cultivar ‘Kefeng 1’ (Tian et al. 2004). Kefeng 1 was released for innate resistance to SMV. SA is a common second messenger protein used by plants to induce resistance responses that are both local and systemic to many pathogens. Here the existence of a protein product encoded by that mRNA and an association with SCN resistance were both novel.

Two spots identified as increased in abundance in SCN infested roots were identical to the soybean NADH-quinone oxidoreductase (QOR; EC 1.6.99.5) isozymes (Table 1). The enzymes were probably both involved in ubiquinone biosynthesis in roots. No metabolites in the
mevalonic acid or quinine pathways were altered in abundance (Table 5). However, a cDNA library isolated from laser micro-dissected syncytia collected 8 days after SCN infestation showed enhanced abundance of the transcript encoding a QOR in a susceptible soybean cultivar (Klink et al. 2005). Similarly, a transcript encoding a QOR was increased in abundance by 4 fold following infestation of A. thaliana roots with BCN (Puthoff et al. 2003). Increased abundance of QOR transcripts in tomato over-expressing the PTO protein has also been observed (Mysore et al. 2003). QOR was shown to act on quinines, a common class of metabolites implicated in energy production, defense against pathogens and electron transport (Matvienko et al. 2001). QOR reactions produce semi-quinones that bind and subsequently inactivate key cellular components such as lipids, proteins, nucleic acids and carbohydrates (Testa 1995). Superoxide anions and hydroxyl radicals are also produced through the reaction of semi-quinones with molecular oxygen. The superoxide anions and hydroxyl radicals cause enzyme inactivation, trigger membrane lipid peroxidation and nick DNA. Thus, QOR plays a critical role in cell toxicity associated with the HR particularly in plants combating microbial pathogens (Hammond-Kosack and Jones 1996). Plant pathogenic cyst nematode genomes like SCN encode several enzymes used in pathogenesis that are bacterial in origin (Ithal et al. 2007a). Perhaps QOR inhibits these enzymes.

One protein with a more than 3.5 fold increase in intensity showed high percentage similarity to a dirigent protein that mediates stereospecific lignin precursor couplings (Table 1) and may be involved in the process of cell wall modification/ lignification (Burlat et al. 2001). Mapman (Thimm et al. 2004) placed the enzyme in the lignin and lignan biosynthesis pathway with laccase (EC 1.11.1.7) the enzyme that works in tandem with dirigent proteins. The amino acid used as a lignin precursor, phenylalanine, was decreased in abundance in coordination with hydroxyl coumaryl alcohol derivative agmatine and cellulose breakdown product N-acetyl-galactosamine (Table 5). Together the data suggest cell walls and protective lignin derivatives are being synthesized to a greater degree in SCN infested resistant roots. The dirigent proteins have previously been implicated in insect resistance by conifers (Ralph et al. 2006) and cyst nematode resistance in soybeans (Klink et al. 2005). Interestingly, the transcript encoding the dirigent-like protein was shown to be decreased in abundance at 3 dai compatible interactions but not in 8 dai compatible interactions (Klink et al. 2007a). A gene encoding a variant laccase is found within the rhg1 locus (Ruben et al., 2006; Iqbal et al., 2009). Therefore, lignin, lignin
and/or isoflavone biosynthesis may underlie some part of resistance to SCN, perhaps the broad SAR response to secondary infestations from late hatching eggs (Iqbal et al., 2009).

An additional protein was identified within the dirigent-like protein spot. The protein corresponded to that encoded by a root EST that was predicted to encode the F1 ATPase subunit of mitochondrial ATP synthase (EC: 3.6.3.14). Combined, the two proteins appeared to be increased by 6.5 fold. Since ATP synthases are very highly conserved, the spot might also correspond to a protein from the nematode rather than plant protein, although there were many more plant cells than nematode cells in the samples. ATP synthase mediates ATP production critical to resistance reactions. In addition, a peptide derived from the proteolytic cleavage of the chloroplast ATP synthase can result in the production of the inceptin elicitor that activates plant defense responses in cowpea during insect herbivory (Schmelz et al. 2006; Schmelz et al. 2007). Nematode resistance genes can be pleiotropic to genes for resistance to other insect herbivores (Nombela et al., 2003) and there is some evidence for the phenomenon at Rhg4 in soybean (Yesudas and Lightfoot, unpublished), so this finding may yet be proven significant.

Another protein identified from the comparison of the resistant NIL infested and non-infested was identical to the Kunitz-type trypsin inhibitor, which has strong homology to the α-fucosidase protein (EC 3.2.1.51; Table 1; Augur 1995). The O-glycosyl hydrolase EC:3.2.1.-family encompasses 85 sub-families. They hydrolyse the glycosidic bond between two or more carbohydrates, or between a carbohydrate and a non-carbohydrate moiety. The protein detected here was from family 29, a lysosomal enzyme responsible for hydrolyzing the alpha-1,6-linked fucose joined to the reducing-end N-acetylglucosamine of the carbohydrate moieties of glycoproteins. Among the related metabolites glucose, sorbose, mannose, galactose, fructose and N-acetyl-galactosamine decreased in abundance 1.5-7.0 fold whilst maltose increased 13 fold in resistant NILs with SCN infestations (Table 5). Increased TA of the mRNA encoding Kunitz trypsin inhibitor transcript has been observed during plant pathogenesis within the syncitia 8 dai (Klink et al. 2007a); and in roots 12 dai (Alkharouf et al. 2006). The protein was associated with the hypersensitive response in leaves of soybeans (Park et al. 2001) and transgenic tobacco (Karrer et al. 1998). The protein was involved in regulation of nitric oxide synthase and responses to snake venom proteins (see discussion in Park et al. 2001). Therefore, the lysosomal activity of the O-glycosyl hydrolase that is a Kunitz trypsin inhibitor seems to be important in
root resistance to SCN. Determining the enzymes specific substrates and molecular basis of the

Another protein spot increased in abundance corresponded to a β-1,3 endoglucanase (EC 3.2.1.39; Table 1). The metabolite substrate glucose was decreased (Table 5). The gene family forms a well characterized group of pathogenesis related proteins that include cell wall hydrolyzing enzymes. Endoglucanases have been reported to be activated by both biotic and abiotic stresses. The pathogenesis protein, PR2 is a β-1,3 endoglucanase (Graham et al. 2003; 2007). The enzyme acts directly in plant disease resistance through digesting the cell wall of the invading pathogen (Bishop et al. 2005). The released cell wall fragments may also act as elicitors or resistance resulting in an accentuated response (Graham et al. 2007).

The most significant changes in protein spot intensity on the 2D gels corresponded to a vegetative storage protein (VSP) and a triose phosphate isomerase (Figure 4; Table 1; EC 5.3.1.1; Xu et al., 2006). The enzyme is at a key point in glycolysis and gluconeogenesis. Several metabolites, substrates and end products in those pathways were decreased in abundance (Table 5) only maltose and ribonic acid increased. The enzyme is a target of glutathionylation by glutathione –S-transferase in A. thaliana (Ito et al. 2003) and was increased in abundance by 2 fold in peach fruit after salicylic acid treatment (Chan et al. 2007). Therefore, this protein may be involved in the energetics of resistance. Considering the VSP, in both soybean and A. thaliana the VSP was induced as a consequence of methyl jasmonate treatment (Creelman et al. 1992), wounding and insect feeding (Liu et al. 2005). The recombinant protein added to the diets of acid gut insects can cause a significant delay in development and an increase in mortality (Liu et al. 2005). The transcript accumulated in A. thaliana plants infested with the SBN (Puthoff et al. 2003). Therefore the VSP may act as a deterrent to feeding.

Proteins and metabolites altered by SCN infestation of resistant compared to susceptible NIL roots.

Twelve spots had differential abundance (Table 2). Three of the twelve proteins were increased in abundance in the inoculated resistant NIL roots (previous section). There was a 2.7 fold increase in triose phosphate isomerase and VSP spot intensity (Figure 4) and 3.3 fold increase in a dirigent-like protein (Table 2). Therefore, these proteins were increased by SCN infestation in resistant but not susceptible NILs. Proteins increasing abundance in response to both genotype and infestation are likely to be important to resistance to SCN.
Among the proteins altered only in the resistant to susceptible infested NIL comparison (increased by the resistance allele at *rhg1* independent of SCN infestation) the enolase (EC:4.2.1.11) protein was increased 3 fold (Table 2). The enzyme converts 2-phosphoglycerate into the shikimate pathway precursor, PEP during both glycolysis and gluconeogenesis. The direct metabolites were not detected by GCMS but several sugars and Krebs cycle acids citric acid and fumaric acid were reduced in the resistant NIL (Table 5). Therefore, increased glycolysis may occurs in the SCN resistance response as it does during both tomato defense against the powdery mildew fungus (Li et al. 2006) and symbiosis between soybean root hairs infected with *Bradyrhizobium japonicum* (Wan et al. 2005).

Another protein increased in abundance was the aldehyde dehydrogenase-1A1 like enzyme (EC 1.2.1…) involved in the conjugation of isoprenoids in plants to various products, including carotenoids and terpenoids. Precursor metabolites inositol and myo-inositol were significantly decreased in the resistant roots (Table 5). Further, plants with mutations in this enzyme accumulated hydrogen peroxide, suggesting a role for this protein in plant defense against oxidative stress (Kirche et al. 2004). Conversely, elevated hydrogen peroxide can mediate the non-enzymatic condensation of lignins and lignans in association with dirigent proteins. Suppression of susceptibility to toxin inhibition of mitochondrial metabolism in T-cytoplasm maize by interference with the production of the Turf-13 protein can be mediated by aldehyde dehydrogenases. Therefore there are several mechanisms by which this enzyme might contribute to resistance to SCN through constitutively higher abundance in resistant genotypes.

A 3 fold increase in the S-adenosyl homocysteine hydrolase (SAH; EC 3.3.1.1) protein spot was detected in SCN infested resistant NILs (Table 2). This enzyme is involved in converting adenosylselenohomocysteine to adenosine and selenohomocysteine, substrates in many methyl-transferase reactions. None of the metabolites in this pathway were detected by GCMS. Methylations reactions are involved in the synthesis and activation of many antimicrobial compounds (Mitsui et al. 1997). SAH can inhibit the synthesis of ethylene, slowing senescence and the HR response. Transcripts encoding SAHs were increased during the hypersensitive response in cassava (Kemp et al. 2005) and alfalfa (*M. sativa*; Edwards 1996). Therefore, the protein seems likely to contribute to resistance by being constitutively increased in resistant NIL roots.
Three of seven enzymes in the ascorbate metabolism pathway were constitutively more abundant in resistant NIL roots. Ascorbate oxidase (EC 1.10.3.3), dehydroascorbate reductase (DHAR; EC 1.8.5.1; Figure. 5a) and mono-dehydroascorbate reductase (mDHAR; EC 1.6.5.4) were identified (Table 2). The precursor threonate was slightly decreased by about 1.4 fold in the resistant NILs (inferred by FT-ICR-MS; Afzal et al. 2009b). The oxidation of ascorbic acid (AA) to the toxic dehydro-\(L\)-ascorbic acid has roles in impeding insect growth and development (Felton and Summers, 1993). DHAR and mDHAR are involved in the detoxification of hydrogen peroxide and were increased in barley genotypes resistant to powdery mildew (El-Zahaby et al. 1995) and maize over-expressing a superoxide dismutase gene (SOD; EC 1.15.1.1; Kingston-Smith et al. 2000). Tobacco expressing a foreign mDHAR showed increased abiotic stress resistance (Eltayeb et al. 2006). Therefore, constitutively higher expression of the ascorbate pathway in resistant NILs may be involved in innate resistance to SCN.

A multi-catalytic endopeptidase protein was increased (Table 2). The protein was highly similar (89% identity, 96%, similarity) to an EST from embryos of SCN resistant soybean cv. ‘Jack’ that encoded the 20 S proteasome beta subunit A1 (EC 3.4.25.1) a central hub in the protein to protein interactome (Supplementary Table 1). The 20S proteasome, was involved in the breakdown of oxidatively modified or damaged proteins (Lee et al. 2006; Giulivi et al. 1994). Subsequently, these proteins may undergo fragmentation or aggregation (Dean et al. 1997). In rice (\textit{Oryza sativa} L.) challenged with \textit{Rhizoctonia solani}, this protein was increased in both resistant and susceptible genotypes, although the increase in abundance was higher in the resistant line. A constitutive increase in the multi-catalytic endopeptidase protein abundance in resistant soybean NILs may be involved in maintaining the resistance response by degrading nematode proteins, by reducing or repairing oxidative damages or by changing the rates of deactivation and degradation of proteins according to their contribution to defense.

Two protein spots with increased intensities were members of the chaperonin (EC 3.6.4.9) family of proteins (Table 2); a 10 KDa like chaperonin and a 23 KDa like co-chaperonin; that assist in the protein folding and are heat shock proteins (HSPs; Sangster and Queitsch 2005). Several HSPs were identified as activated upon pathogen infection (Sangster and Queitsch 2005; Wan et al. 2005; Klink et al. 2007a; 2007b). Over expression of HSP70 in the resistant plants was reported to reduce root mass in \textit{A. thaliana} (Sung and Guy. 2003). In this study, shorter roots and lower root masses was seen in plants that underwent incompatible
reactions (Afzal et al. 2008b; 2009a). Further, a different chaperonin appears to interact with the RLK at \textit{rhg1}. Therefore, the chaperonins present at higher abundance in resistant roots may underlie three phenomena; part of the SCN resistance, the reduced seedling vigor associated with resistance (Afzal et al. 2008b; 2009a) and the lower seed yield that results (Yuan et al., 2002).

**Proteins induced by SCN in the infested susceptible NIL roots**

Proteins in this class might be part of a defeated defense response or proteins induced by signals from SCN to root cells. Four proteins were identified (Table 3). One, mDHAR, had been detected as constitutively increased in the resistant genotype. The mDHAR spot was, therefore, increased in both the resistant infested and susceptible infested NILs but to a lesser degree. Therefore the protein increase was a general response to SCN infestation enhanced by resistance to SCN directed by \textit{rhg1}.

A spot with a 6 fold increased abundance was identified as a glycine-rich RNA binding protein (RBP; Table 3). RBP TAs were increased by a number of external stimuli, including wounding, abscisic acid (ABA), dehydration, stress, mercuric chloride treatment and infection with TMV (Brady et al. 1993; Naqvi et al. 1998). The role of nuclear localized RBP in adjustments to changing internal and external stimuli may be based upon interactions with several RNAs, proteins, and kinases. Therefore, the change in protein abundance during a compatible reaction with SCN may be part of a general response to infection or be directed by the nematode.

Two spots corresponding to peroxidase (EC 1.11.1.7) proteins were increased in abundance (Table 3; Figure 5b). The substrates are inorganic chemicals not detected by GC MS. Peroxidases TAs have dual roles in plant defense, sometimes decreased in abundance to channel resources to more urgent regulatory needs (Moy et al. 2004) or increased, as in 8 day post SCN infestation syncitial cells compared to roots cells (Klink et al. 2007b). Peroxidases may have an active role in wheat stem rust disease (Seevers and Daly. 1971; Tyagi et al. 2000). The increase in peroxidase isozymes during pathogen infection may restrict pathogen growth. Alternately, peroxidases could be involved in the synthesis of plant phenolics or detoxification of peroxide substrates (Seevers and Daly. 1971).

**Protein comparisons between non infested resistant and non infested susceptible NILs**

Proteins in this class are expected to be constitutively increased by the presence of the resistance allele at \textit{rhg1}. Only glucose-6 phosphate isomerase and isoflavone reductase showed
increased abundance in the comparison (Table 4). Glucose-6 phosphate isomerase (E.C 5.3.1.9), is a glycolytic enzyme that converts glucose 6-phosphate into fructose 6-phosphate. The metabolite fructose and product sorbose were increased about 2 fold while the substrate precursor maltose decreased 13 fold in resistant NILs (Table 5). Previously, glucose-6 phosphate isomerase was amongst a number of glycolytic enzymes with increased abundance in cereals in response to the sheath blight fungus (Danson et al. 2000). Therefore, increased glycolysis in roots might contribute to greater readiness to mount a resistance response to SCN.

Isoflavone reductase (EC 1.3.1.45) was also implicated in readiness against SCN infestation by a 3 fold increase (Table 4). Precursor phenylalanine was slightly decrease in abundance in resistant genotypes (Table 5) possibly as a result of the more active isoflavone synthesis. Isoflavone reductase accumulation in plants increases with increased resistance toward pathogen infection (Kuc 1995; Ithal et al. 2007) by increased production of phytoalexins and deposition of lignin (Keen 1992). Interactions with the laccases and dirigent proteins are expected. Therefore, the isoflavonoid phytoalexin synthesis pathways are strongly inferred to be under the control of the rhg1 locus.

**Interactome analysis**

Ortholog analysis was conducted on 31,921 proteins from *A. thaliana* (from TAIR) and 62,199 proteins from *G. max* (from the DOE at www.phytozome.net). Shown at SoyGD (http://soybeangenome.siu.edu) was the % score that is the likelihood for orthology, while the other members of the cluster were in-paralogs (multiple orthologs in the same species). If a single protein could not be assigned as the best ortholog, INPARANOID assigned it to clusters as one to many (one protein in species having many orthologs in the other), and many to many (many proteins in one species having many other orthologs in the other). Here only one to one orthologs were considered. This, secondary, predicted soybean interactome was compared with the proteins altered in abundance by SCN infestation (Figure 6). Amongst them proteins were spread into many clusters as expected from pathway analysis (nineteen enzymes in 17 pathways). However, four proteins were clustered in the interactome map and their interactions were expanded by another node. After two iterations two more proteins were added. Thus, from the huge soybean interactome, a small portion with significance to resistance to SCN was identified with an enrichment for four thousand fold suggesting significance to the cluster.
The predicted interactome showed the chaperonin, F1-ATPase, multi catalytic endopeptidase, thaumatin-like protein, cytosolic heat shock protein and triose phosphate isomerase were interacting indirectly through 2 to 7 intermediates (Figure 6). The cluster of all the potential interacting neighbors showed 159 proteins in total, with 194 interactions. The protein functions included transcription factors, chaperonins, signal transduction factors and metabolisms involved in energy generation (Supplementary Table 1). The four proteins in the cluster with most interacting partners were a GTP binding-family protein (28 partners), importin alpha-2 subunit (64 partners), importin beta-2 subunit (22 partners) and a 26S proteasome non-ATPase regulatory subunit 3a (21 partners). Therefore, the regulation of protein degradation and transport of proteins across the nuclear membranes may be important aspects of resistance to SCN.

Amongst the 159 proteins predicted to interact, were many proteins previously associated with disease resistance. One was the thaumatin-like (TL) protein induced by SCN as well as response to pathogens (displayed in blue). TL proteins were previously shown to be defense proteins that reduced the permeability of the cell walls to pathogenic fungi (Abad et al. 1996). The TL protein in soybean roots was predicted to interact with the importin alpha-2 subunit-like node (Figure 6; Supplemental Table 1) a nodal protein. The chaperonin detected changed in protein abundance by SCN was also predicted to interact through at this node. Therefore, the regulation of transport of both plant and nematode secreted proteins across the nuclear membranes during resistance to SCN will be a focus of future studies.

The second protein previously linked to disease resistance was a cytosolic heat shock protein (AT4G11260.1; displayed in green; Figure 6). This protein interacted with the 26S proteasome non-ATPase regulatory subunit 3a (Table 2). In Arabidopsis, HSP90 was required for complete RPS2 mediated signaling in response to pathogens (Takahashi et al. 2003) and was shown to interact with the disease resistance signaling component SGT1b (Azevedo et al. 2002). Therefore during resistance to SCN protein-protein interactions were predicted to play an important role in cellular functions. Eventually, the interactome map for soybean may be used to predict which proteins in a network respond to various external stimuli and to design new mechanisms of resistance.

Comparison of TA between infested and non-infested resistant NILS
Seven out of the 19 genes that encoded enzymes in pathways were chosen to compare protein with TAs because mRNA specific primers that produced single amplicons could be designed. Glutathione-S-transferase (1.57±0.38), quinine oxidoreductase (1.38±0.27), F1-ATPase (1.87±0.53), triose phosphate isomerase (1.25±0.34) and glutathione S transferase (1.31±0.24) were increased in TA when the comparison was between the infested and non-infested NIL. There was agreement between the 2D gel and real-time quantitative PCR results for these genes. However, the gene transcripts encoding trypsin inhibitor (-2.43±0.34), the EST induced by salicylic acid (-2.41±0.39) and quinone oxidoreductase (-1.2±0.20) were decreased in TA but increased in protein abundance (Figure. 7; Table 6). The protein abundance increase in response to SCN was inferred to be a post-transcriptional event for these three proteins that may reflect differences between in protein and RNA half lives (Yu et al. 2007). Correlations between transcript and protein abundances have mostly been poor (Greenbaum et al. 2003; Yu et al. 2007). Therefore, root resistance to SCN was shown to be derived from changes in both TA and protein abundance.

**TA differences between resistant and susceptible NILs infested with SCN.**

A similar uncoupling between TA and protein abundance was observed when the mRNA abundance in infested resistant NIL 34-23 and infested susceptible NIL 34-3 was compared. Out of seven genes encoding proteins increased in abundance, four; aldehyde dehydrogenase (16.32±0.76); glutathione dehydrogenase (ascorbate) (1.33±0.45); multicatalytic endopeptidase (1.75±0.56); and monodehydroascorbate reductase (1.16±0.41) were increased in TA. However, three genes encoding S-adenosyl homocystine hydrolase (-13.5±0.25); hypothetical disease responsive protein(-1.65±0.40) and chaperonin (-1.11±0.42) were found decreased in TA (Table 7; Figure 8). Therefore, both root resistance and susceptibility to SCN were shown to be derived from changes jointly in TA and protein abundance.

**Conclusion**

Numerous attempts to determine the key transcripts involved in SCN resistance in soybean have been made (Alkharouf et al., 2004; 2006; Klink 2005, 2007a; 2007b). Variation in experimental design included analysis of compatible and incompatible interactions; plant responses to resistant and susceptible SCN biotypes; and analysis of transcript variation at
different days post inoculation. None of the studies thus far sought to determine changes in actual protein abundance as a function of SCN infection. Previous studies relied heavily on genomic tools; mainly microarray analysis using soybean Affymetrix chips to measure TA. One weakness of TA analysis is that the abundances of transcripts and those of the respective proteins may not correlate well. In this study there was limited correlation between some transcripts and their encoded protein but more between proteins and their substrates or metabolites. Therefore, results obtained from transcriptomic, proteomic and metabolomics must be integrated and compared with the interactomes to elucidate the nature of resistance to SCN and other phenomena. This study showed that differentially abundant proteins exist in the context of genotype and reaction following SCN infection and lays the groundwork for a systems biology approach to SCN resistance. In addition, proteins and metabolites in roots were discovered that may be used to predict the resistance of seedlings without SCN infestation at high-throughput. Future research will explore the usefulness of these biomarkers further.

Materials and Methods

Materials

Seeds of soybean were obtained from SIUC. Seeds of NIL 34-23 (resistant haplotype between markers Satt 214 to Satt 570) and NIL 34-3 (susceptible haplotype from Satt 214 to Satt 570) were obtained at the F5:13 generation. Genotypes were $rhg1rhg1Rhg4Rhg4$ for NIL 34-23 and $Rhg1Rhg1Rhg4Rhg4$ for NIL 34-3.

Growth conditions

Soybean ($Glycine\ max$) plants were grown tubes 5 cm by 30 cm placed in buckets of sand with 20 tubes in a completely randomized design. Each tube contained a 1:1 ratio of sand:soil mix. The buckets were placed in a water bath to maintain root zone temperature at 24 C and grown at 14h light cycle, with aerial day time temperature of 30°C and a night-time temperature of 22°C. The humidity was maintained at approximately 40-50% (v/v).

Infection with Hg Type 0 SCN populations consisted of infestation of the root zone of 14 day old seedlings with 2,000 eggs in 2 ml water. Non-infested plants were provided the same volume of water. For this Hg Type female indices (FI) were ‘PI54840’ (FI 7%), PI 88788 (FI 2%), PI90763 (FI 1%), PI437654 (FI 0%), ‘PI 209332’ (FI 1%), ‘PI89772’ (FI 2%) ‘PI548316’
and ‘PI548402’ (FI 3%). Therefore the standard differentials showed this HG Type to be 0 (Niblack et al. 2003) corresponding to a biotype of race 3 (Riggs and Schmitt, 1988)

Soybean plants (infested and non-infested) were removed from the cones, 10 dai. The cones were transferred to a cold water (4 C) container and briefly soaked until the sand soil mixture became loose. The plant roots were subsequently cleaned with distilled water, blotted dry, cut from the shoots and stored in a -80°C freezer.

DNA extraction

DNA was isolated following a modified protocol by Saghai-Maroof et al. (1984). Briefly, 100 mg of plant tissue was homogenized frozen and 600 µl preheated (65°C) extraction buffer (Tris HCl (pH 8.0) (100 mM), EDTA (pH 8.0) (20 mM), NaCl (1.4 M), CTAB (2%; w/v), 2-mercaptoethanol (0.4%; v/v)) was added. The samples were incubated at 65°C for 1 hour, cooled and centrifuged at 10,000 g for 15 minutes. The supernatant was decanted and 5 µl of RNase (5 mg/ml) was added and incubated at 37°C for one hour. The DNA in the aqueous phase was purified by phenol: chloroform: isoamyl alcohol (25:24:1), extraction followed by chloroform: isoamyl alcohol (24:1). DNA was precipitated by the addition of iso-propanol and centrifugation at 12,000 g for 5 min. The DNA pellets were washed twice with 70 % (v/v) ethanol, dried and finally dissolved in 30 µl Tris buffer. DNA concentrations were calculated by measuring absorbance at 260 nm.

NIL genotype and infection confirmation

About 50 ng of DNA was used for microsatellite analysis according to Yuan et al. (2002). The micro-satellite marker, TMD1 was used to differentiate between the resistant and susceptible NILs. To confirm infection in SCN infested plants at 10 dai and lack thereof in the non-infested NILs, PCR amplification used Heterodera glycine 18S ribosomal gene primer pair (Forward-HG-18S; 5′-TATTCACCAC CTACCTG CTGTCCT-3′ and Reverse-HG 18S; 5′-GCAGCCATA CACAGGCTTTCACA T -3′).

Total protein extraction and quantification

Total protein extraction followed Hurkman et al (1986) and Hajduch et al (2005). Briefly, frozen root samples from 10 plants were pooled together, ground to a fine powder and 2 g of
tissue was resuspended in 5 mL of Tris buffered phenol (pH 8.8) and 5 mL of extraction buffer (0.1 M Tris-HCl pH 8.8, 0.4 % (v/v) β-mercaptoethanol, 10 mM EDTA and 0.9 M sucrose). The solution was vortexed vigorously and centrifuged at 5,000 g for 15 minutes. After removal of the top phase (phenol) the bottom phase was back extracted with Tris buffered phenol (5ml) and an equal volume of the extraction buffer. Proteins were pelleted by centrifugation at 20,000 g for 20 minutes and washed according to (Hajduch et al, 2005). The pellet was dried and resuspended in the DeStreak™ rehydration agent (GE Healthcare, Piscataway, NJ). Protein concentration was determined using a non-interfering protein assay (Sheffield et al. 2006).

2D gel electrophoresis and Image analysis

Protein extracts (275 µg) from infested NILs (34-3 or 34-23) and non-infested NILs 34-23 and 34-3 was used for the 2D electrophoretic analysis. Briefly, samples were initially hydrated overnight on a 17 cm BioRad IPG gel strips (pH 3-10) and IEF (iso-electric focusing) performed using the Protean IEF Cell (BioRad, Hercules, CA). Equilibration of the strips was according to the manufacturer’s instructions. Linear SDS PAGE gradient gels (8 % to 16 % (w/v)) were used to resolve proteins in the second dimension. A BioRad Protean II apparatus was used for gel electrophoresis at 15 mA/cm for 30 min, followed by 25 mA/cm for approximately 5 hrs at ambient temperature (20±2 °C). Gels were washed with distilled water and stained with SYPRO Ruby. SYPRO Ruby fluorescent dye used for gel staining provided both higher sensitivity and a broad linear range for accurate protein quantification and detection (Sheffield et al. 2006). Protein spots on duplicate gels were compared using Imagemaster 2D software (GE Healthcare).

Image analysis and spot picking

Acquisition of gel images used a high resolution laser scanner (Typhoon 9410 from GE Healthcare) and a CCD camera linked to a GelPix protein spot excision system (Genetix, New Milton, UK). Images were analyzed with the Imagemaster 2D software (GE Healthcare). The analysis was fully automated unless there was a need to edit unresolved spots. Identical spots from the compared gels were chosen by assigning landmarks to each gel. The volume for each protein spot was normalized against the total spot intensity. Students T-test (p, 0.05) was used to
determine whether spot intensities between treatments were significantly different. Differentially abundant spots were excised manually or with the GelPix system.

**Protein digestion and clean-up**

Proteins were digested in-gel as described previously (Gabelica et al. 2002), with the exception that digestion was carried out at 37°C overnight with 6 mg/ml trypsin (Promega; Madison, WI) in 50mM NH₄HCO₃. The samples were initially extracted with 30 µl of 1 % (v/v) formic acid, 2 % (v/v) methyl-cyanide followed by incubation at 30°C for 30 min on a shaking platform. For the second extraction, 60 % (v/v) acetonitrile was used. The pooled extractions were separated into aliquots, lyophilized and stored. Aliquots were resuspended in 2 % (v/v) acetonitrile, 1 % (v/v) formic acid solution for use. Proteins were further purified cleaned using a 10 µl C18 ZipTip™ (Millipore, Bedford MA) according to manufacturer.

**Database searching**

Proteins were identified via peptide sequencing using ESI MS/MS as described previously (Chen et al. 2005). Analyst QS software (Applied Biosystems) was used for spectral processing. The peptides were searched against the soybean and *M. truncatulata* EST databases (downloaded from NCBI, January 2006), the non-redundant NCBI database, and the Swissprot database using MASCOT version 1.9. Input parameters for variable and fixed modifications were specified as “oxidation of methionine and carbamidomethylation of cysteine” respectively. Positive identification was based on; (a) number of peptide sequences identified in a protein; (b) calculated and theoretical PI /molecular weight; and (c) total MASCOT and MOWSE scores ([http://www.matrixscience.com/help/scoring_help.html](http://www.matrixscience.com/help/scoring_help.html)) at a P value below 0.05. Protein sequences derived from EST databases or without significant matches were searched against the NCBI blastx, tblastx databases for evidence of orthology and DOE soybean peptide library (www.phytozome.net) for evidence of identity.

**Q-PCR analysis**

Total RNA was isolated using Trizol (Invitrogen, Carlsbad CA) from both mock and SCN infected root samples. Total RNA was treated with DNaseI and the absence of DNA was confirmed by PCR amplification using intron flanking primers before reverse transcription.
Reverse-transcriptase Super script II (Invitrogen) was used to generate first strand cDNA following manufacturer’s protocol. Primer pairs for the genes corresponding to differentially expressed proteins were designed using the IDT primer design tool (www.IDTDNA.com). SYBR Green (iQ SYBR green Supermix, Bio Rad) was used to measure the number of amplicons after the Q-PCR reaction. The experiments were performed in triplicate. The soybean tubulin gene was used as an internal control (Iqbal et al., 2002; 2005). At the end of each experiment, the melting curve for each amplicon was determined at 62 °C to 95 °C, with readings every 1 °C. The 2$^{-\Delta\DeltaCT}$ method was used to analyze the relative changes in TA as described previously (Livak and Schmittgen 2001).

**Metabolite extraction, identification and pathway analysis.**

Metabolites were extracted and identified from nine replicated plant samples using adapted standard operating practices for GC-TOF-MS analysis of plant extracts as outlined in Mungur (2008). Briefly, 15 mg of freeze dried root tissue was ground in 2 ml extraction mix (methanol:water, 1:1, v/v) on ice, shaken for 4-6 min at 4°C and centrifuged for 2 min at 14,000 g. About 1 mL of supernatant was collected as two aliquots, which were dried in a vacuum centrifuge. All samples and replicates were stored at –20°C under argon or nitrogen until further processing. The dried extracts were dissolved in 5 µL of methoxamine hydrochloride (20 mg/ml pyridine) and incubated at 30°C for 90 min with continuous shaking. Then 45 µL of N-methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) was added to derivatize polar functional groups at 37°C for 30 min.

GC-TOF MS (Leco Pegasus II GC-TOF mass spectrometer; Leco, St. Joseph, MI, USA) analyses were performed on an HP 5890 gas chromatograph with tapered, deactivated split/splitless liners containing glasswool (Agilent, Böblingen, Germany) and 1.5 µL splitless injection at 230°C injector temperature. Before each injection, the liner was rinsed with a pure MSTFA injection (1 µl). Sample injection was carried out without sample wash steps due to the limited amount of total sample volume. The GC was operated at constant flow of 2 ml/min helium and a 30 m 0.32 mm id 0.25 µm MDN35 column (Macherey-Nagel, Düren, Germany). The temperature gradient started at 80°C, was held isocratic for 2 min, and subsequently ramped at 15°/min to a final temperature of 330°C which was held for 6 min. Twenty spectra per second were recorded between m/z 85–500. Peak identification and quantification were performed using
the Pegasus software package ChromaTOF 1.61 (Leco, St Joseph, MI, USA). Retention time shifts were corrected by linear interpolation using known metabolites as reference markers. All files were subsequently processed against a reference which was generated using a signal/noise threshold of 10 with automated peak identification based on mass spectral comparison to a standard NIST 98 library and available in house customized mass spectral libraries. GC-MS identified 131 metabolites among which 58 were significantly altered by one or more treatment. Of the 58, 27 could not be unequivocally identified but 31 were named. A full description of the metabolite analysis will be published elsewhere (Afzal et al., 2009b). Here metabolite data was examined for concordance with the pathways in which the protein enzymes altered in abundance. The analysis used MAPMAN tuned for the soybean proteins that were known enzymes (Goffard and Weiler, 2006) or their Arabidopsis thaliana orthologs (Thimm et al., 2004). Mapman showed the enzymes in pathways with their metabolites.

**Protein-Protein Interactome analysis**

A Linux based software InParanoid (Remm et al. 2001) was used to detect orthologs between A. thaliana and G.max. Inparanoid works on the NCBI based BLAST server where 100% orthology is assigned for an e-value of 0.01 or lower. The first round of the program used BLOSSUM45 as the scoring matrix. The second and third rounds used BLOSSUM62 and BLOSSUM80 respectively. Proteins with 100% orthology between the two species were used for the analysis. Protein annotations for the orthologs of soybean were imported from (http://www.biomart.org) which contained curated information for A. thaliana proteins from TAIR and Ensemble. Proteins of soybean which showed 100% orthology to A. thaliana were superimposed on to the Arabidopsis interactome kindly obtained from Dr. Matt Geisler. Cytoscape 2.5.1 was used to build the interactome.

**ACKNOWLEDGEMENTS**

The authors thank Dr. P. Gibson, O Myers Jr. and M. Schmidt for assistance with germplasm development and maintenance from 1991-2000 and Gary and Mary Kinsel for assistance with protein analysis and interpretation. The support of SIUC, College of Agriculture
and Office of the Vice Chancellor for Research to DAL, JA and AS was appreciated. We thank the DOE CJSP for release of the WGS reads and scaffolds. We thank the Government of India for support of the Fellowship for NS. We thank the Government of Egypt for support HAE via the Egypt-USA junior scientist visit grants.
References


Chan ZL, Qin GZ, Xu XB, Li BQ, Tian, SP (2007) Proteome approach to characterize proteins induced by antagonist yeast and salicylic acid in peach fruit. *J Proteome Res* **6**: 1677-1688


drawbacks of nanospray for studying noncovalent protein-DNA complexes by mass


Giulivi C, Pacifici RE, Davies KJA (1994) Exposure of hydrophobic moieties promotes the
selective degradation of hydrogen peroxide-modified hemoglobin by the multicatalytic
proteinase complex, proteasome. *Arch Biochem Biophys* **311**: 329-341

*Bioinformatics* **22**: 2958-2959

current state of two-dimensional electrophoresis with immobilized pH gradients.
*Electrophoresis* **21**:1037-1053

pathogenesis-related protein genes in soybean by wounding and the *Phytophthora sojae*
cell wall glucan elicitor. *Physiol Molec Plant Pathol* **63**:141-149

elicititation or biosynthesis of 5-deoxyisoflavonoids suppresses race specific resistance and
hypersensitive cell death in *Phytophthora sojae* infected tissues. *Plant Physiol* **144**: 728-
740

and mRNA expression levels on a genomic scale. *Genome Biol* **4**:117

filling in soybean: Establishment of high-resolution two-dimensional reference maps,
expression profiles, and an interactive proteome database. *Plant Physiol*, **137**:1397-1419


Hurkman, WJ and Tanaka, CK (1986) Solubilization of plant membrane proteins for analysis
by two-dimensional gel electrophoresis. *Plant Physiol* **81**:802-806

Ithal N, Recknor J, Nettleton D, Hearne L, Maier T, Baum TJ, Mitchum MG (2007a) Parallel genome-wide expression profiling of host and pathogen during soybean cyst nematode infection of soybean. Mol Plant Microbe Interact 20:293-305


Klink VP, Alkharouf N, MacDonald M, Matthews B (2005) Laser capture microdissection (LCM) and expression analyses of Glycine max (soybean) syncytium containing root regions formed by the plant pathogen Heterodera glycines (soybean cyst nematode). Plant Mol Biol 59: 965-979

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

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


Figures

Figure 1: Shown is a 2 % (w/v) agarose gel after electrophoresis showing presence (lanes 1 and 2) and absence (lanes 3 and 4) of SCN DNA in 10 day after infestation (lanes 1 and 2) or non-infested plants (lanes 3 and 4). The SCN 18S ribosomal gene was successfully amplified in infested NILS 34-23 and 34-3 (lanes 1 and 2; upper band). No amplification was obtained from DNA extracted from non-infested NILs (lanes 3 and 4). Lane 5 contained the negative control with no template. The lower band in each lane was the primers.

Figure 2: The 2D electropherograms obtained from the infested resistant NIL 34-23 (top) and the non-infested resistant NIL 34-23 (bottom). Most protein spots were present under both conditions. The pI for majority of the spots was in the pH 4 to 8 range whereas the molecular mass for most proteins was between 20 and 100 kDa.

Figure 3: Pie chart showing functional distribution of the proteins identified from 2D-PAGE. About 45% of the identified proteins we involved in disease or defense responses. The classification was based on the functional catalog of plant genes by Bevan et al. (1998).

Figure 4: The regions of three 2D electropherograms displaying differences in protein expression between the infested resistant NIL (a); infested susceptible NIL (b); and the non-infested resistant NIL (c). The encircled spot showed quantitative differences in all three gels. Two proteins were identified within the spot. These were the vegetative storage protein and triose phosphate isomerase.

Figure 5: The regions of three 2D electropherograms displaying differences in protein expression between the infested resistant NIL and the infested susceptible NIL. The protein circled in panel (a) was glutathione dehydrogenase (Table 2). Panel (b) showed differentially abundant proteins in the susceptible infested compared to the susceptible non-infested NIL. The protein circled was a peroxidase (Table 3).
**Figure 6:** A part of the predicted interactome for soybean showing six proteins from the 2D gel analysis (chaperonin, F1ATPase, multi catalytic endopeptidase, thaumatin like protein, cytosolic heat shock protein and triose phosphate isomerase). The interactome was extended by one node to include 159 nodes (proteins) and 194 edges (interactions; Supplemental Table 1). Two proteins with a previously characterized role in disease resistance were identified within this part of the interactome; the thaumatin like protein induced in response to fungal pathogens (blue) and a cytosolic heat shock protein in disease resistance signaling (green). All six proteins interacted through 2-7 intermediate partners (see Supplemental Table 1 for pink spot identification numbers). Autoregulation was predicted for the triose phosphate isomerase and the chaperonin.

**Figure 7:** Relative TA ratios measured by quantitative RT-PCR. The fold change in relative TA in the resistant SCN infested NIL compared to the resistant non-infested NIL is shown for 10 plant replicates pooled after tissue grinding for RNA extraction. Date shows mean values ± standard deviations from three technical replicates.

**Figure 8:** Relative TA ratios measured by quantitative RT-PCR. The fold change in relative TA in the resistant infested NIL compared to the susceptible infested NIL is shown for 10 plant replicates pooled after tissue grinding for RNA extraction. Date shows mean values ± standard deviations from three technical replicates.
Figure 1: Shown is a 2 % (w/v) agarose gel after electrophoresis showing presence (Lanes 1 and 2) and absence (Lanes 3 and 4) of SCN DNA in 10 day post infested (lanes 1 and 2) or mock infested (lanes 3 and 4) plants. The SCN 18S ribosomal gene was successfully amplified in infested NILS 34-23 and 34-3 (lanes 1 and 2; upper band). No amplification was obtained from DNA extracted from non-infested NILs (lanes 3 and 4). Lane 5 shows the negative control with no DNA template. The lower band in each lane is the primer dimer.
Figure 2: The 2D electropherograms obtained from the infested resistant NIL 34-23 (top) and the non-infested resistant NIL 34-23 (bottom). Most protein spots were present under both conditions. The pI for majority of the spots was in the pH 4 to 8 range whereas the molecular mass for most proteins was between 20 and 100 kDa.
Figure 3: Pie chart showing functional distribution of proteins identified from 2D-PAGE. About 45% of the identified proteins were involved in disease or defense responses. The classification was based on the functional catalog of plant genes (Bevan et al. 1998).
Figure 4: The regions of three 2D electropherograms displaying differences in protein expression between the infested resistant NIL (a); infested susceptible NIL (b); and the non-infested resistant NIL (c). The encircled spot showed quantitative differences in all three gels. Two proteins were identified within the spot. These were the vegetative storage protein and triose phosphate isomerase.
Figure 5: The regions of three 2D electropherograms displaying differences in protein expression between the infested resistant NIL and the infested susceptible NIL. The protein circled in panel (a) was glutathione dehydrogenase (Table 2). Panel (b) shows differentially abundant proteins in the susceptible infested compared to the susceptible non-infested NIL. The protein circled was a peroxidase (Table 3)
Figure 6: A part of the predicted interactome for soybean showing six proteins from the 2D gel analysis (chaperonin, F1ATPase, multi catalytic endopeptidase, thaumatin like protein, cytosolic heat shock protein and triose phosphate isomerase). The interactome was extended by one node to include 159 nodes (proteins) and 194 edges (interactions; Supplemental Table 1). Two proteins with a previously characterized role in disease resistance were identified within this part of the interactome; the thaumatin like protein induced in response to fungal pathogens (blue) and a cytosolic heat shock protein in disease resistance signaling (green). All six proteins interacted through 2-7 intermediate partners (see Supplemental Table 1 for pink spot identification numbers). Auto-regulation was predicted for the triose phosphate isomerase and the chaperonin.
Figure 7: Relative TA ratios measured by quantitative RT-PCR. The fold change in relative TA in the resistant SCN infested NIL compared to the resistant non-infested NIL is shown for 10 plant replicates pooled after tissue grinding for RNA extraction. Date shows mean values ± standard deviations from three technical replicates.
Figure 8: Relative TA ratios measured by quantitative RT-PCR. The fold change in relative TA in the resistant infested NIL compared to the susceptible infested NIL is shown for 10 plant replicates pooled after tissue grinding for RNA extraction. Date shows mean values ± standard deviations from three technical replicates.