

GENETIC AND MOLECULAR ANALYSIS OF SOYBEAN CYST NEMATODE
VIRULENCE

A Dissertation
Presented to
The Faculty of the Graduate School
At the University of Missouri

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy

By
MICHAEL N. GARDNER
Dr. Melissa G. Mitchum, Doctoral Supervisor

MAY 2017

The undersigned, appointed by the dean of the Graduate School,
have examined the dissertation entitled
GENETIC AND MOLECULAR ANALYSIS OF SOYBEAN CYST NEMATODE
VIRULENCE

Presented by Michael N. Gardner

A candidate for the degree of Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Melissa G. Mitchum

Dr. Walter Gassmann

Dr. Dmitry Korkin

Dr. James Schoelz

DEDICATION

This dissertation is dedicated in memory of my aunt, who was always a source of joy and encouragement to me.

Jan Arkkelin

1951 –2014

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor, Melissa Mitchum. Without her valuable assistance and advice this work would never have been accomplished. I would also like to thank my doctoral committee members Drs. Walter Gassmann, James Schoelz, and Dmitry Korkin for their valuable advice and suggestions on this project. I would like to thank all the members of the Mitchum lab both past and present for their help and comradery, in particular fellow graduate students Gregory Yeckel and Carola de la Torre. I also owe a debt of gratitude to the Life Sciences Fellowship Program for funding me during my graduate career and in particular Debbie Allen, who always had an open door. I would also like to thank my fellow Division of Plant Sciences graduate students and members of the Students for the Advancement of Plant Pathology for all their friendship and assistance. Finally, I would like to thank my family and my beautiful wife Carol for putting up with me all these years. The humor and love they bring to my life makes each day a new and wonderful surprise.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
LIST OF FIGURES.....	vii
LIST OF TABLES.....	viii
ABSTRACT.....	ix
CHAPTER	
1. INTRODUCTION.....	1
Economic importance of plant parasitic nematodes.....	1
Biology of the soybean cyst nematode.....	2
Effectors secreted by cyst nematodes.....	5
Resistance to the soybean cyst nematode.....	11
Soybean cyst nematode virulence.....	14
Rationale.....	15
References.....	18
2. INHERITENCE OF <i>HETERODERA GLYCINES</i> VIRULENCE GENES AND THE EFFECT OF SECONDARY SELECTION ON VIRULENCE.....	24
Abstract.....	24
Introduction.....	25
Results.....	28
Crosses, F ₁ and F ₂ generation host range tests.....	28
Progeny lines and parasitism gene number.....	30
Host-ranges of outselected inbred lines.....	32
Secondary selection.....	32

Discussion.....	34
Materials and Methods.....	40
Parental lines.....	40
Crosses and F ₁ generation testing.....	41
Mating and backcrossing.....	41
Cross population selections.....	42
Secondary selection.....	42
Acknowledgements.....	43
References.....	44
3. NOVEL EFFECTOR MINING FROM A <i>DE NOVO</i> TRANSCRIPTOME ASSEMBLY OF THE EARLY PARASITIC STAGES OF THE SOYBEAN CYST NEMATODE <i>HETERODERA GLYCINES</i>	47
Abstract.....	47
Introduction.....	48
Results.....	51
Transcriptome sequencing and assembly.....	51
Annotation of transcripts.....	53
Identification and GO annotation of endosymbiont-associated transcripts from the <i>H. glycines</i> transcriptome.....	55
Known SCN stylet-secreted effector analysis.....	55
Novel effector prediction.....	60
Discussion.....	64
Materials and Methods.....	72
Nematode cultivation and isolation.....	72

RNA isolation and sequencing.....	72
<i>de novo</i> transcriptome assembly.....	73
Transcriptome annotation and quantification.....	74
Identification of endosymbiont sequences within the <i>H. glycines</i> transcriptome.....	75
Gene ontology analysis of endosymbiont-associated sequences from the <i>H. glycines</i> transcriptome.....	75
Variation of known SCN effectors.....	76
Effector prediction.....	76
Acknowledgements.....	78
References.....	79
4. VIRULENT TRANSCRIPTOME ASSEMBLY AND COMPARATIVE ANALYSIS OF KNOWN <i>HETERODERA GLYCINES</i> EFFECTORS.....	85
Abstract.....	85
Introduction.....	86
Results.....	88
Transcriptome sequencing and assembly.....	88
Annotation of transcripts.....	91
Differential expression analysis.....	91
Gene ontology analysis.....	97
Discussion.....	101
Materials and Methods.....	110
Nematode cultivation and isolation.....	110
RNA isolation and sequencing.....	110
<i>de novo</i> transcriptome assembly.....	111

Transcriptome annotation and quantification.....	112
Gene ontology enrichment analysis.....	112
Acknowledgements.....	113
References.....	114
5. SUMMARY AND FUTURE DIRECTIONS.....	119
References.....	123
APPENDIX – SUPPLEMENTAL FIGURES.....	124
Appendix 1 - Summary of transcript annotation for the SCN transcriptome.....	124
Appendix 2 - Orthologs identified in sequenced nematode genomes.....	124
Appendix 3 - Predicted effector candidates from the SCN transcriptome.....	124
Appendix 4 - GO term enrichment between virulent and avirulent SCN populations.....	124
VITA.....	125

LIST OF FIGURES

Figure	Page
1.1	Distribution of the soybean cyst nematode.....3
1.2	Life cycle of the soybean cyst nematode.....3
1.3	The esophageal gland cells of the soybean cyst nematode.....6
1.4	Characterization of known cyst nematode effectors.....7
1.5	Resistance to and virulence of the soybean cyst nematode.....12
2.1	Crossing strategy for determining virulence genes.....29
2.2	Outselected inbred populations from a cross between PA3 and TN20.....33
2.3	Secondary selection of outselected inbred populations.....35
3.1	Species distribution of homologues in the SCN transcriptome.....54
3.2	Orthologs in sequenced nematode proteomes.....54
3.3	Identification of endosymbiont-associated transcripts.....56
3.4	GO annotation of endosymbiont-associated transcripts.....57
3.5	Variation of known effectors in the transcriptome.....59
3.6	Secreted effector protein prediction pipeline.....63
4.1	Species distribution of homologues in the virulent transcriptome.....92
4.2	Differential expression analysis of the virulent SCN transcriptome.....93
4.3	Conserved transcripts across virulent SCN populations.....95
4.4	Differential expression of known SCN effectors in virulent populations.....96
4.5	Gene enrichment analysis of the virulent SCN population MM8.....99
4.6	Gene enrichment analysis of the virulent SCN population MM10.....100

LIST OF TABLES

Table		Page
2.1	Host range test of PA3 x TN20 segregating progeny lines.....	31
3.1	<i>de novo</i> transcriptome assembly statistics for the SCN transcriptome.....	52
3.2	Up-regulated SCN effectors in the parasitic second-stage juvenile.....	61
3.3	Down-regulated SCN effectors in the parasitic second-stage juvenile.....	62
3.4	Predicted effector candidates of note.....	65
4.1	Description of populations used for the virulent transcriptome assembly.....	90
4.2	<i>de novo</i> transcriptome assembly statistics for the virulent transcriptome.....	90
A.1	Orthologs identified in sequenced nematode genomes.....	124
A.2	Summary of transcript annotation for the SCN transcriptome.....	124
A.3	Predicted effector candidates from the SCN transcriptome.....	124
A.4	GO term enrichment between virulent and avirulent SCN populations.....	124

ABSTRACT

The soybean cyst nematode (SCN) *Heterodera glycines* is the most economically important pathogen of soybean, capable of causing large scale yield loss on a global scale. Current management practices utilize host resistance, but commercial resistance is limited to three main sources, the soybean plant introductions (PI) 54802 (Peking), 88788, and 437654. More than 95% of soybean grown in the north central United States have resistance derived from PI 88788 and repeated use of this resistance has led to the development of highly virulent nematode populations capable of infecting these resistant plants and causing yield losses. In order to identify the mode of inheritance for virulence on the three primary sources of resistance to SCN a controlled crossing study was performed, mating a highly virulent nematode (TN20) with an avirulent nematode (PA3). By monitoring the offspring of this cross it was found that depending on the source of host resistance, virulence is inherited in a dominant or a recessive manner except for virulence on PI 437654, which is likely a multigenic trait dependent on multiple recessive genes. A *de novo* transcriptome assembly was then generated for SCN and mined for novel stylet-secreted effectors, identifying a new pool of candidates that may play a role in virulence. Finally, a comparative transcriptomic analysis was performed across multiple SCN populations to identify conserved expression patterns and genes associated with virulence. Results from these studies will be used to improve current management practices for SCN and provide new potential targets for improving SCN resistance.

CHAPTER 1

INTRODUCTION

Economic importance of plant-parasitic nematodes

Plant-parasitic nematodes are a very diverse group of obligate biotrophs capable of causing immense damage to a broad variety of host crops. To date, over 4,100 different species of plant-parasitic nematodes have been identified with additional species still being discovered (Decraemer and Hunt, 2006). Damage caused by these nematodes has been estimated to cause over \$80 billion per year in yield loss, a likely underestimate due to the non-specific symptoms they cause and the difficulty in accurately diagnosing nematode diseases (Nicol et al., 2011). Plants targeted by plant-parasitic nematodes include economically important crops such as rice, corn, soybean, sorghum, potatoes, coffee, wheat, and many other vegetable and ornamental crops (Evans et al., 1993; McDonald and Nicol, 2005; Nicol and Rivoal, 2007). While the majority of these nematodes target and infect the roots of these crops, there are some examples of plant-parasitic nematodes that infect the aboveground structures of the plant as well. The wide variety of hosts, difficulty in diagnosis, and global distribution of plant-parasitic nematodes all lead to them having such an immense economic impact.

Cyst nematodes are just one group of these pathogens, with some of the most economically important members in the genera *Heterodera* and *Globodera*. In general cyst nematodes are confined to a much smaller host range of economically important crops than other plant-parasitic nematodes, with exceptions such as the beet cyst

nematode *Heterodera schachtii* and the corn cyst nematode *Heterodera zae* with wider host ranges (Ringer et al., 1987; Evans and Rowe, 1998). The soybean cyst nematode (SCN) *Heterodera glycines* is one of the most economically important, largely due to its broad dispersal and the limited tools available for management. Since the initial discovery of SCN in North Carolina in 1954, SCN has spread to nearly all locations in the United States where soybean is grown (Figure 1.1; Tylka and Marett, 2014). Typically this dispersal occurs through infected soil on equipment or shoes moved from field to field. In particular, SCN is a devastating pathogen of soybean, causing over \$1 billion in yield loss annually in the United States alone (Koenning and Wrather, 2010). Once SCN has been introduced into a field complete eradication is nearly impossible without the use of expensive and highly toxic chemical treatments and so management practices focus on limiting the reproduction of the nematode through cultural practices and the use of host resistance.

Biology of the soybean cyst nematode

The soybean cyst nematode begins life as an egg in the soil. Before emerging from the egg the nematode will molt for the first time, progressing from the first-stage juvenile (J1) to the second-stage juvenile (J2). At this point the nematode hatches from the egg aided by a hollow, protractible stylet as well as enzymes secreted through the stylet to degrade the egg. These enzymes are part of a suite of molecules originating within the three esophageal gland cells of the nematode, one dorsal and two subventral. These gland cells produce proteins responsible for many aspects of the life cycle of the

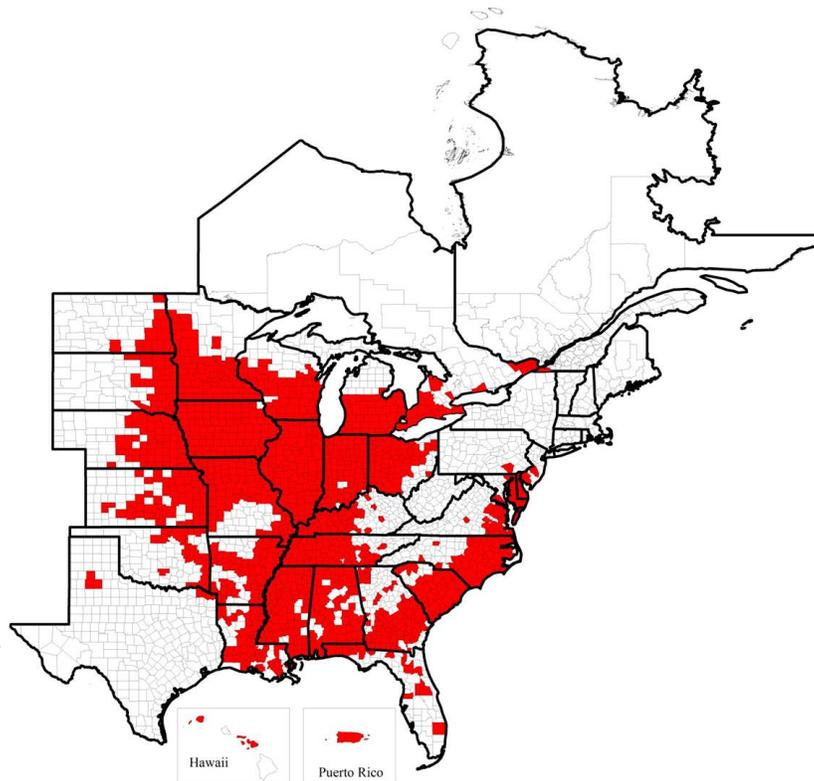


Figure 1.1. Distribution of the soybean cyst nematode *Heterodera glycines* in North America in 2014. Reproduced from Tylka and Marett, 2014.

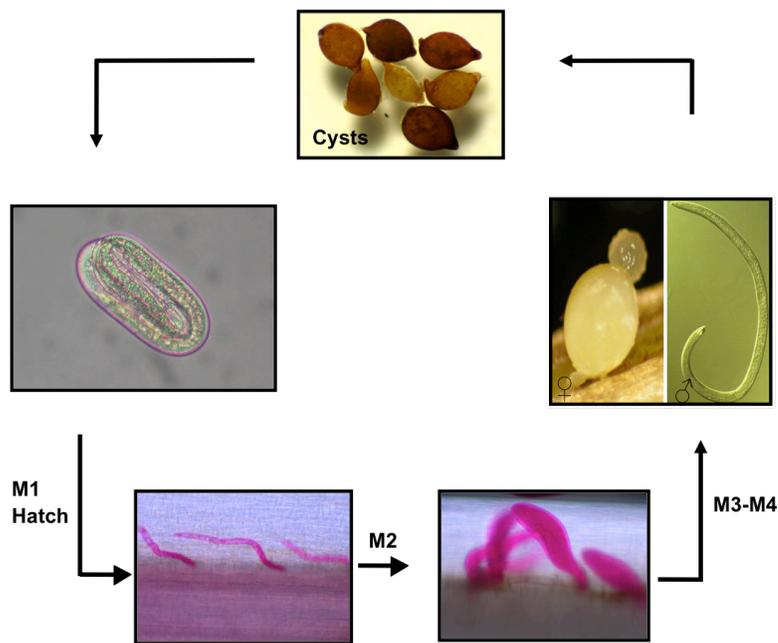


Figure 1.2. Life cycle of the soybean cyst nematode. Reproduced from Mitchum and Baum, 2008.

nematode, most importantly being the parasitism of plants. Once the nematode has hatched, it migrates towards a plant root system, drawn by an unknown signal towards the plant host (Figure 1.2). Upon reaching the root the nematode again makes use of the stylet and a set of enzymes as it moves through the root and towards the vasculature. This movement is intracellular and highly destructive, which can trigger some host defense responses (Endo and Veech, 1970). As the nematode approaches the vasculature it selects a cell from the endodermal cell layer or the pericycle and penetrates it with its stylet. At this point the nematode becomes sedentary and loses the use of its somatic musculature, relying completely on the host plant for nutrition and survival. The nematode then produces effector proteins within the esophageal gland cells and secretes these effectors into the selected root cell to trigger the formation of the syncytium, a feeding structure necessary for nutrition and survival of the nematode (Gheysen and Mitchum, 2011). The syncytium expands through the dissolution of surrounding cell walls and merging of cell contents, encompassing up to hundreds of cells (Jones et al., 1972). Once the contents of the cells have merged into the syncytium the nematode utilizes effectors again to begin reprogramming of the cell. This reprogramming results in a highly dense cytoplasm and an increase in metabolic activity within the syncytium in order to sustain the nematode (Davis and Mitchum, 2005). The sedentary nematode then continues to feed and maintain the syncytium, eventually undergoing an additional three molts before differentiating into either an adult male or female. Females are pyriform in shape and remain in place on the root, their bodies filling with up to 500 eggs. Meanwhile, adult male nematodes become vermiform again and leave the root to fertilize the females, after which their life cycle is completed. Once the fertilized female dies its

body hardens into a protective casing for the eggs called a cyst and falls off into the soil, where the eggs can lay dormant for years until a suitable host is present, upon which time the entire cycle begins again.

Effectors secreted by cyst nematodes

Effectors represent a crucial element of the life cycle of the cyst nematodes and are one of the key adaptations that allow them to infect plants. These proteins are primarily produced via genes expressed in the three esophageal gland cells of the nematode, at which point they are secreted out of the stylet and into the plant (Figure 1.3; Davis et al., 2008). Different effectors have been shown to function both in the apoplast as well as the interior of the host cell, indicating effectors could be secreted outside the cell or secreted directly into the host cell. These effectors have traditionally been identified based on the presence of an N-terminal signal peptide, lack of a transmembrane domain, and confirmed expression within the esophageal gland cells (Gao et al., 2003). To date many of these stylet-secreted effectors have been identified across multiple species of cyst nematodes and have been functionally characterized (Figure 1.4). Most of these functions fall into two overarching categories: modification of plant development and suppression of host defense responses (reviewed in Gardner et al., 2015).

Cyst nematode effectors involved in the modification of plant development are largely involved in manipulating the host cell to be a suitable syncytium for the nematode. This is done by mimicking and/or modifying existing plant proteins as well as modulating phytohormone signaling. One of the characterized effector mimics of a host

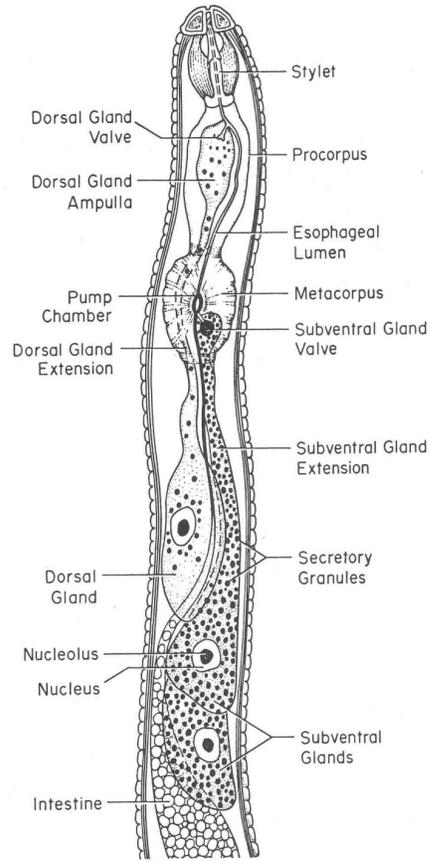


Figure 1.3. Esophageal gland cells of the soybean cyst nematode. Illustration of the esophageal gland cells and secretory machinery of the soybean cyst nematode. Effectors are produced within the subventral and dorsal gland cells and secreted out of the stylet into the outside environment. Reproduced from Hussey, 1989.

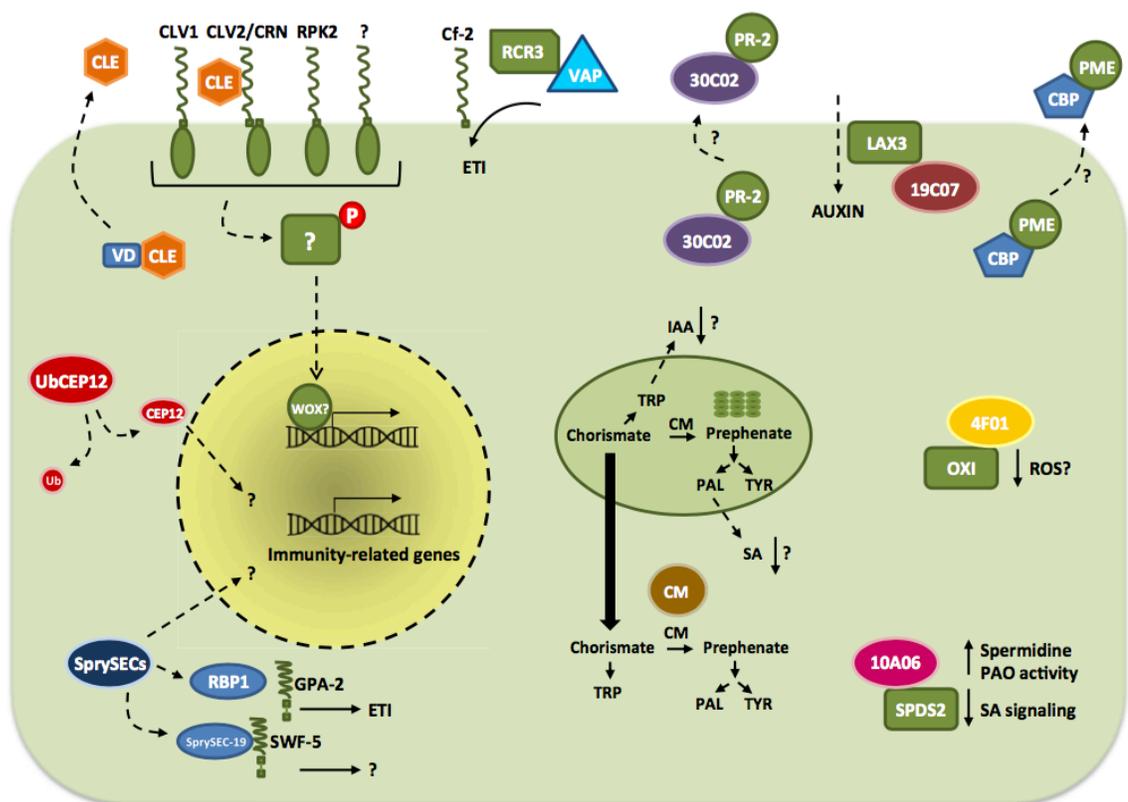


Figure 1.4. Characterization of known cyst nematode effectors. A diagram representation of known cyst nematode effectors and their targets functioning within the cell. 19C07 interacts with LAX3 to possibly modulate auxin transport. CLAVATA3/endosperm surrounding region- related (CLE) peptides are peptide hormone mimics that interact with the extracellular leucine-rich repeat domain of plant receptor-like kinases with roles in stem cell maintenance. Cellulose-binding protein (CBP) interacts with pectin methylesterase (PME3) to likely modify cell wall properties. Chorismate mutase (CM) affects the plant shikimate pathway to modulate aspects of cellular metabolism. 10A06, 4F01 and 30C02 are novel effectors that interact with spermidine synthase (SPDS2), oxidoreductase (OXI) and β -1,3-endoglucanase (PR-2), respectively and may modulate plant defense responses. GrUBCEP12 is processed in planta to release free ubiquitin and the 12 amino acid GrCEP12 peptide to possibly affect the 26S proteasome and modulate plant defense. SPRYSECs are a large class of secreted SPRY-domain containing proteins that localize to different subcellular compartments of host cells. The R protein GPA-2 mediates effector-triggered immunity (ETI) by directly or indirectly recognizing the SPRYSEC effector RBP-1. SPRYSEC-19 interacts with SWF-5; however, the biological significance of this interaction is unknown. The R protein Cf-2 causes ETI by recognizing perturbations to the active site of the apoplastic protease Rcr3 by VAP-1. Host protein targets are green; cyst effectors are coloured. VD, variable domain; VAP, venom allergen-like protein; PR, pathogenesis related; PAO, polyamine oxidase; IAA, indole-3-acetic acid; UBCEP12, ubiquitin carboxyl extension protein 12; SA, salicylic acid; ROS, reactive oxygen species; WOX, WUSCHEL-like homeobox. Reproduced from Gardner et al., 2015.

peptide is the CLAVATA3/endosperm surrounding region-related (CLE) ligand mimic family. This family of stylet-secreted effectors first identified in the cyst nematode genera *Heterodera* and *Globodera*, is the first example of a pathogen secreting a protein mimic of a plant peptide hormone (Mitchum et al., 2012). Nematode CLEs are produced in the dorsal gland cell and secreted out from the stylet to the cytoplasm of the host cell, where they are thought to be processed and post-translationally modified by the host cell machinery and eventually delivered back to the apoplast (Ohyama et al., 2009; Chen et al., 2015). These CLE peptides have been shown to bind with plant receptors and are predicted to mimic the function of endogenous plant CLEs (Guo et al., 2010; Guo et al., 2011). Plant CLEs serve as ligands for receptors that mediate signaling involved in maintenance of stem cells in the plant root, shoot, floral, and vascular structures by regulating members of the WUSCHEL-like homeobox (WOX) transcription factor family (Yamaguchi et al., 2016). By manipulating these functions the nematode CLE is able to alter host plant gene expression and promote syncytium development. This function is supported by the observation that silencing *CLE* genes in cyst nematodes or their associated plant receptors causes reduced parasitism (Bakhetia et al., 2007; Chen et al., 2015; Guo et al., 2015; Patel et al., 2008; Replogle et al., 2011; Replogle et al., 2013). Another mechanism by which cyst nematodes manipulate their host plant is through modulation of phytohormone levels at the feeding site. Previous studies have identified changes in local phytohormone levels at the feeding site for multiple cyst nematodes (Absmanner et al., 2013; Goverse et al., 2000; Grunewald et al., 2009; Hewezi et al., 2014; Karczmarek et al., 2004; Wubben et al., 2001). Now we also know these nematodes actually secrete phytohormone mimics as well (Siddique et al., 2015). It only

makes sense that stylet-secreted effectors produced by the nematode would cause these hormone changes. One specific example of this is the *Heterodera* effector 19C07, the first cyst nematode effector implicated in direct modulation of host transporter proteins to alter the balance of phytohormones in favor of syncytium development. Through protein-protein interaction studies the *H. schachtii* 19C07 was found to interact with the *Arabidopsis* LAX3 protein, a member of the AUX/LAX family of auxin influx transporters (Lee et al., 2011). LAX3 mutants have abnormal development of lateral roots caused by the inability to transport auxin to cells in the overlaying lateral root primordia. During nematode infection LAX3 is expressed in the initial syncytial cells and then shifts to adjacent cells before they are incorporated into the syncytium. This elevation of auxin leads to activation of cell wall modifying proteins that subsequently loosen the surrounding cell walls, priming these cells to merge them with the syncytium. Given the changes in phytohormone levels during nematode infection there are likely other effectors related to hormone signaling as well that still remain to be characterized. In addition to altering phytohormones, cyst nematode effectors have also been shown to directly modify cell wall architecture at the syncytium. These include not only cell wall degrading enzymes such as cellulases and pectinases but also a cellulose-binding protein (CBP) produced by *H. glycines* (reviewed in Bohlmann and Sobczak, 2014). The secreted CBP plays a role in the early stages of syncytium formation and was also shown to interact with a plant pectin methyltransferase 3 (PME3). This interaction is thought to increase the activity of PME3 during infection, leading to the demethylesterification of pectin in the cell wall to facilitate early stages of syncytium formation by allowing increased access of other cell wall modifying enzymes to the polymers of the cell wall

(Hewezi et al., 2008). The modifications caused by these effectors serve to modify the host cell and prepare it to be a suitable environment for sustaining the nematode.

In addition to manipulating the host cell into the syncytium, cyst nematodes also secrete a variety of effectors to protect the nematode from plant stress and defense responses. One of the major components of a plant defense response to pathogens, nematode or otherwise, is the production of reactive oxygen species (ROS) (Lamb and Dixon, 1997). Not only do these compounds have a direct toxic effect on the pathogen, they can also be used by the plant in order to regulate other processes such as hormone signaling and development (Mittler et al., 2004). To combat this cyst nematodes produce a variety of effectors, from those that function directly on the ROS produced such as detoxifying enzymes to those that function indirectly through interactions with host proteins (reviewed in Gardner et al., 2015). These effectors all function to reduce the levels of harmful ROS in order to protect the nematode. Other effectors target the host cell nucleus, such as the nematode secreted ubiquitin carboxyl extension proteins (UBCEPs) and the SP1a and Ryanodine receptor (SPRY) domain containing (SPRYSEC) family (Chronis et al., 2013; Cotton et al., 2014). Upon secretion of these effectors into the host cell they are imported to the nucleus, upon which time they interfere with nuclear functions, including causing the suppression of the 26S proteasome in the case of UBCEP12 (Chronis et al., 2013). By suppressing these host defense responses the nematode effectors are able to protect both the nematode and its feeding site, allowing for successful parasitism to occur. While many of the cyst nematode effectors have some level of characterization, there is likely a much larger repertoire of un-identified effectors

needed to perform all the functions required in a feeding site that can be revealed by using a comprehensive genetic analysis.

Resistance to the soybean cyst nematode

The primary management strategy against SCN is the use of resistant cultivars (Mitchum, 2016). While many resistant cultivars are available, commercial soybean cultivars typically derive their SCN resistance from one of three sources, the soybean plant introductions (PI) 54802 (Peking), 88788, and 437654. Resistance typically manifests itself at the feeding site, leading to degeneration of the syncytium and death of the nematode (Figure 1.5). In the case of Peking and PI 437654 or other cultivars utilizing these resistances the syncytium rapidly degenerates, with necrosis initiated inside the syncytium and surrounded by a necrotic layer (Kim et al., 2012). This is accompanied by dilation of the rough endoplasmic reticulum and the formation of cell wall appositions and can completely breakdown the syncytia as early as two days after infection. In contrast to this, in lines deriving their resistance from PI 88788 necrosis begins outside the cell wall of the syncytium and nuclear degeneration occurs within the syncytium cytoplasm, resulting in complete degeneration of the syncytium by eight to ten days after infection (Kim et al., 1987). Cell-specific transcriptome studies of the degenerating feeding cells comparing between resistant and susceptible lines have also revealed expression of genes associated with ER dysfunction, formation of cell wall appositions, and apoptotic and autophagic programmed cell death responses in the feeding cell formed on a resistant host (Kandoth et al., 2011; Matsye et al., 2011). The

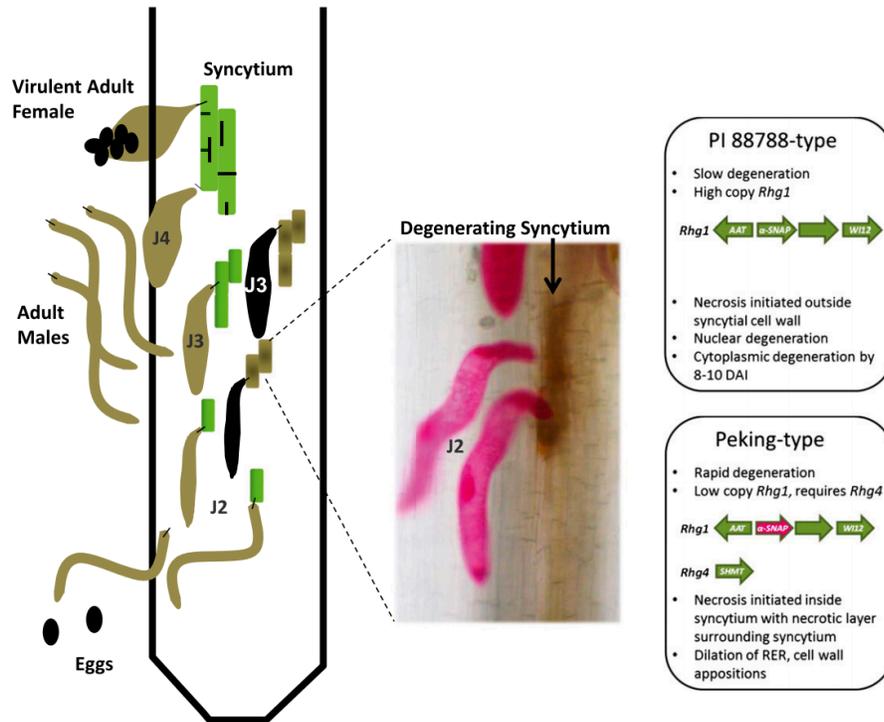


Figure 1.5. Resistance to and virulence of the soybean cyst nematode *Heterodera glycines*. Illustration representing the life cycle of a virulent nematode vs. an avirulent nematode on a resistant soybean host. Avirulent individuals trigger resistance in the second stage juvenile (J2), leading to the break down of the feeding site (enlarged image) by the early third stage juvenile (J3). Meanwhile, virulent individuals are able to successfully establish and maintain their feeding site throughout their life cycle, eventually differentiating into adult males and females and reproducing to create more virulent individuals. Reproduced from Mitchum, 2016.

differences in response to nematodes on these resistant are indicative of two distinct mechanisms of resistance.

In recent years much has been discovered about the genetics of soybean cyst nematode resistance. Resistance is primarily due to two quantitative trait loci (QTL) named *rhg1* (for resistance to *H. glycines*) and *Rhg4* (reviewed in Mitchum, 2016). The *rhg1* QTL contains three genes in tandem repeats that grant resistance to SCN (Cook et al., 2012). These three genes are an amino acid transporter, an N-ethylmaleimide-sensitive factor attachment protein (α -SNAP), and a wound-inducible protein (WI12). In soybean cultivars that derive their resistance from PI 88788 these genes are found in high copy numbers (8-10 copies), whereas cultivars that derive their resistance from Peking have lower copy numbers (1-3 copies) and susceptible soybean has a single copy. Further research identified a new allele of α -SNAP present in PI 437654 and Peking but not PI 88788 and that the α -SNAP is sufficient for resistance in these sources (Cook et al., 2014; Lee et al., 2015; Liu et al., 2017). Mapping studies served to further differentiate the genetic requirements for resistance between PI 88788, Peking, and PI 437654, establishing a requirement for only *Rhg1* in PI 88788-derived resistance while Peking-type resistance requires the *Rhg4* QTL as well (Concibido et al., 2004; Meksem et al., 2001). The gene underlying resistance at *Rhg4* was found to be a serine hydroxymethyltransferase (SHMT), an enzyme involved in the conversion of serine and tetrahydrofolate (THF) to glycine and 5,10-methylenetetrahydrofolate (MTHF) (Liu et al., 2012). Furthermore, the SHMT encoded by *Rhg4* in resistant soybean differs by two amino acids when compared to the susceptible variant. While the loci and genes

underlying resistance have been identified, little is known about the function of these genes and how they contribute to SCN resistance.

Soybean cyst nematode virulence

The limited number of resistance sources available for SCN has created a problem in the form of SCN virulence. Repeated use of the same source of resistance has led to the development of highly virulent SCN populations in the field defined as being able to infect and reproduce on resistant soybean. Statewide surveys support this observation, showing increased reproduction on PI 88788-derived resistance in particular (Niblack et al., 2006). One way to address this is to rotate resistant cultivars, however little is known about the interaction between virulence on different types of resistance and if counter-selection occurs. Without knowledge of the heritability and genes underlying virulence it is impossible to make informed decisions about rotation schemes between resistant cultivars and may actually end up producing nematodes highly virulent on multiple sources of resistance.

To better understand SCN virulence previous studies have sought to identify the genes underlying virulence. The first of these studies noted the ability to reproduce on the resistant soybean cultivars PI 88788 and Pickett was inherited independently (Triantaphyllou, A., 1987). This was followed by the identification of the *ror* (for reproduction on a resistant host) genes (Dong and Opperman, 1997). These genes were associated with reproduction on several resistant lines, including PI 88788. However, with the limited knowledge available about SCN resistance their actual identity was not found and they were never fully characterized. More recently, two candidates for SCN

virulence genes were proposed, a biotin synthase (*HgBioB*) and a putative SNARE domain (*HgSLP-1*) containing protein (Bekal et al., 2015). These two virulence candidates were identified through a genetic mapping study to identify SNPs linked to virulence genes. *HgSLP-1* is deleted in virulent SCN and may trigger resistance. The authors propose that *HgSLP-1* interacts with the *rhg1* α -SNAP during a membrane fusion event and may directly trigger resistance. However, direct evidence of role for *HgSLP-1* is lacking and the absence of *HgSLP-1* still needs to be confirmed in field populations. While progress is being made in identifying virulence genes, there is still much to be learned in this area.

Rationale

In order to better understand SCN virulence genetic tools need to be generated for *H. glycines*. Next generation sequencing technologies have resulted in the complete genome sequences for multiple species of plant-parasitic nematodes, including the cyst nematodes *Globodera rostochiensis* and *Globodera pallida* (Cotton et al., 2014; Eves-van den Akker et al., 2016). Assembling a genome, however, can be costly and represent additional challenges. Initial efforts to sequence the SCN genome have run into difficulties in the assembly process, largely due to highly repetitive sequences in the genome and the high purine content. In the absence of a genome, *de novo* transcriptome assembly can be performed to provide comprehensive genetic information of an organism. Transcriptome assemblies from the plant-parasitic nematodes *Heterodera schachtii*, *Meloidogyne graminicola*, and *Heterodera avenae* have all been used to garner

novel insight into parasitism and nematode effector biology (Fosu-Nyarko et al., 2016; Haegeman et al., 2013; Kumar et al., 2014; Petitot et al., 2016).

To investigate the mechanisms of virulence we set out to accomplish the following tasks in this dissertation. First, to identify the methods of inheritance for virulence genes on some of the common sources of resistance through controlled nematode crosses and selection experiments. Second, to generate a high quality *de novo* transcriptome assembly of the parasitic second-stage of SCN with which to identify novel effector candidates that may play a role in SCN parasitism and virulence. Finally, to sequence and assemble transcripts from SCN populations virulent on the main sources of SCN resistance and compare these populations to identify potential molecular mechanisms underlying virulence and changes in the effector repertoire that may contribute to virulence.

This information can be used not only to make better management decisions, but also to open new avenues of research into SCN virulence. Providing comprehensive transcriptome sequence for SCN will not only provide information about nematode responses to resistance but the sequence data will also be an asset in eventual assembly of the SCN genome. Novel effector candidates will provide new insight into the host-pathogen interaction and can be targeted by biotechnology in order to reduce the impact of SCN on soybean crops. Also, overlapping virulence mechanisms between SCN populations virulent on different sources of resistance can be taken advantage of to provide broad SCN resistance. Finally, with greater sequence availability for SCN comparative genetic analyses can be performed with other species of plant-parasitic

nematodes in order to discover overlapping pathways that might be targeted to provide broad-spectrum nematode resistance.

REFERENCES

- Absmanner, B., Stadler, R., and Hammes, U. Z. 2013. Phloem development in nematode-induced feeding sites: the implications of auxin and cytokinin. *Frontiers in Plant Science*, 4: 241.
- Bakhetia, M. B., Urwin, P. E., and Atkinson, H. J. 2007. qPCR analysis and RNAi define pharyngeal gland cell-expressed genes of *Heterodera glycines* required for initial interactions with the host. *Molecular Plant-Microbe Interactions* 20: 306–312.
- Bekal, S., Domier, L.L., Gonfa, B., Lakhssassi, N., Meksem, K., and Lambert, K.N. 2015. A SNARE-like protein and biotin are implicated in soybean cyst nematode virulence. *PLoS ONE*: e0145601.
- Bohlmann, H., and Sobczak, M. 2014. The plant cell wall in the feeding sites of cyst nematodes. *Frontiers in Plant Science* 5: 91-100.
- Chen, S., Lang, P., Chronis, D., Zhang, S., De Jong, W. S., Mitchum, M. G., et al. 2015. *In planta* processing and glycosylation of a nematode CLAVATA3/ENDOSPERM SURROUNDING REGION-like effector and its interaction with a host CLAVATA2-like receptor to promote parasitism. *Plant Physiology* 167: 262–272.
- Chronis, D., Chen, S., Lu, S., Hewezi, T., Carpenter, S., Loria, D., et al. 2013. A ubiquitin extension protein secreted from a plant-parasitic nematode *Globodera rostochiensis* is cleaved in planta to promote plant parasitism. *Plant Journal* 74: 185-196.
- Concibido, V., Diers, B., and Arelli, P. 2004. A decade of QTL mapping for cyst nematode resistance in soybean. *Crop Science* 44(4): 1121-1131.
- Cotton, J., Lilley, C., Jones, J., Kikuchi, T., Reid, A., Thorpe, P., et al. 2014. The genome and life-stage specific transcriptomes of *Globodera pallida* elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biology* 15: R43.
- Cook, D., Lee, T.G., Guo, X., Melito, S., Wang, K., Bayless, et al. 2012. Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338: 1206-1209.
- Davis, E., and Mitchum, M. 2005. Nematodes. Sophisticated parasites of legumes. *Plant Physiology* 137: 1182-1188.
- Davis, E., Hussey, R., Mitchum, M., and Baum, T. 2008. Parasitism proteins in nematode-plant interactions. *Current Opinion in Plant Biology* 11: 360-366.
- Decraemer, W., and Hunt, D. 2006. Structure and classification. In: Perry, R., and Moens, M. (eds) *Plant Nematology*. CABI Publishing, Wallingford, p. 3-32.

- Dong, K., and Opperman, C. 1997. Genetic analysis of parasitism in the soybean cyst nematode *Heterodera glycines*. *Genetics* 146: 1311-1318.
- Endo, B., and Veech, J. 1970. Morphology and histochemistry of soybean roots infected with *Heterodera glycines*. *Phytopathology* 60: 1493-1498.
- Evans, E., Trudgill, D., and Webster, J. 1993. Plant parasitic nematodes in temperate agriculture. CABI Publishing, Wallingford, p. 648.
- Evans, K., and Rowe, J. 1998. Distribution and economic importance. In: Sharma, S. (ed) *The Cyst Nematodes*. Kluwer Academic Publishers, Dordrecht, p. 1-30.
- Eves-van den Akker, S., Laetsch, D., Thorpe, P., Lilley, C., Danchin, E., Da Rocha, M., et al. 2016. The genome of the yellow potato cyst nematode, *Globodera rostochiensis*, reveals insights into the basis of parasitism and virulence. *Genome Biology* 17:124.
- Fosu-Nyarko, J., Nicol, P., Naz, F., Gill, R., and Jones, M. 2016. Analysis of the transcriptome of the infective stage of the beet cyst nematode, *H. schachtii*. *PLoS ONE* DOI: 10.1371/journal.pone.0147511.
- Gao, B., Allen, R., Maier, T., Davis, E.R., Baum, T.J., Hussey, R.S. 2003. The parasitome of the phytonematode *Heterodera glycines*. *Molecular Plant Microbe Interactions* 2003: 720-726.
- Gardner, M., Verma, A., and Mitchum, M. 2015. Emerging roles of cyst nematode effectors in exploiting plant cellular processes. Chapter 11. In: Escobar, C., and Fenoll, C. (eds) *Advances in Botanical Research Vol. 73, Plant Nematode Interactions: A View on Compatible Interrelationships*. Elsevier Publishing, Oxford, p. 259-291.
- Gheysen, G., and Mitchum, M. 2011. How nematodes manipulate plant development pathways for infection. *Current Opinion in Plant Biology* 14: 415-421.
- Goverse, A., Overmars, H., Engelbertink, J., Schots, A., Bakker, J., and Helder, J. 2000. Both induction and morphogenesis of cyst nematode feeding cells are mediated by auxin. *Molecular Plant-Microbe Interactions*, 13: 1121-1129.
- Grunewald, W., Cannoot, B., Friml, J., and Gheysen, G. 2009. Parasitic nematodes modulate PIN-mediated auxin transport to facilitate infection. *PLoS Pathogens* 5: e1000266.
- Guo, Y., Han, L., Hymes, M., Denver, R., and Clark, S. 2010. CLAVATA2 forms a distinct CLE-binding receptor complex regulating *Arabidopsis* stem cell specification. *Plant Journal* 63: 889-900.
- Guo, Y., Ni, J., Denver, R., Wang, X., and Clark, S. 2011. Mechanisms of molecular mimicry of plant CLE peptide ligands by the parasitic nematode *Globodera rostochiensis*. *Plant Physiology* 157: 476-484.

- Guo, X., Chronis, D., De la Torre, C., Smeda, J., Wang, X., and Mitchum, M. 2015. Enhanced resistance to soybean cyst nematode *Heterodera glycines* in transgenic soybean by silencing putative CLE receptors. *Plant Biotechnology Journal* 13: 801-810.
- Haegeman, A., Bauters, L., Kyndt, T., Rahman, M., and Gheysen, G. 2013. Identification of candidate effector genes in the transcriptome of the rice root knot nematode *Meloidogyne graminicola*. *Molecular Plant Pathology* 14: 379-390.
- Hewezi, T., Howe, P., Maier, T., Hussey, R., Mitchum, M., Davis, E., et al. 2008. Cellulose binding protein from the parasitic nematode *Heterodera schachtii* interacts with *Arabidopsis* pectin methylesterase: cooperative cell wall modification during parasitism. *The Plant Cell* 20: 3080-3093.
- Hewezi, T., Priya, S., Richard, G., and Rice, J. H. 2014. Spatial and temporal expression patterns of auxin response transcription factors in the syncytium induced by the beet cyst nematode *Heterodera schachtii* in *Arabidopsis*. *Molecular Plant Pathology* 15: 730–736.
- Hussey, R. 1989. Disease-inducing secretions of plant-parasitic nematodes. *Annual Review of Phytopathology*. 27: 123-141.
- Jones, M., and Northcote, D. 1972. Nematode-induced syncytium – a multinucleate transfer cell. *Journal of Cell Science* 10: 789-809.
- Kandath, P., Ithal, N., Recknor, J., Maier, T., Nettleton, D., Baum, T., and Mitchum, M. 2011. The soybean Rhg1 locus for resistance to the soybean cyst nematode *Heterodera glycines* regulate the expression of a large number of stress- and defense-related genes in degenerating feeding cells. *Plant Physiology* 155: 1960-1975.
- Karczmarek, A., Overmars, H., Helder, J., and Goverse, A. 2004. Feeding cell development by cyst and root-knot nematodes involves a similar early, local and transient activation of a specific auxin-inducible promoter element. *Molecular Plant Pathology* 5: 343–346.
- Kim, Y., Riggs, R., and Kim, K. 1987. Structural changes associated with resistance of soybean to *Heterodera glycines*. *Journal of Nematology* 19: 177-187.
- Kim, Y., Kim, K., and Riggs, R. 2012. Initial subcellular responses of susceptible and resistant soybeans infected with the soybean cyst nematode. *The Plant Pathology Journal* 28: 401-408.
- Koenning, S.R., and Wrather, J.A. 2010. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. *Plant Health Progress* 10.
- Kumar, M., Gantasala, N., Roychowdhury, T., Thakur, P., Banakar, P., Shukla, R., et al. 2014. *De novo* transcriptome sequencing and analysis of the cereal cyst nematode, *Heterodera avenae*. *PLoS ONE* 9: e96311.

- Lamb, C., and Dixon, R. 1997. The oxidative burst in plant disease resistance. *Annual Review of Plant Physiology and Molecular Biology* 48: 251-275.
- Lee, T.G., Kumar, I., Diers, B.W., and Hudson, M.E. 2015. Evolution and selection of *Rhg1*, a copy-number variant nematode-resistance locus. *Molecular Ecology* 24: 1774-1791.
- Liu, S., Kandoth, P., Warren, S., Yeckel, G., Heinz, R., Alden, J., et al. 2012. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492: 256-260.
- Liu, S., Kandoth, P., Lakhssassi, N., Kang, J., Colantonio, V., Heinz, R., et al. 2017. The soybean GmSNAP18 underlies two types of resistance to soybean cyst nematode. *Nature communications*. In press.
- Matsye, P., Kumar, R., Hosseini, P., Jones, C., Tremblay, A., Alkharouf, N., et al. 2011. Mapping cell fate decisions that occur during soybean defense responses. *Plant Molecular Biology* 77: 513-528.
- Meksem, K., Pantazopoulos, P., Njiti, V., Hyten, D., Arelli, P., and Lightfoot, D. 2001. 'Forrest' resistance to the soybean cyst nematode is bigenic: saturation mapping of the *Rhg1* and *Rhg4* loci. *Theoretical and Applied Genetics* 103: 710-717.
- Mitchum, M., and Baum, T. 2008. Genomics of the soybean cyst nematode-soybean interaction. In: Stacey, G., (ed.), *Genetics and Genomics of Soybean*. Springer, NY p. 321-341.
- Mitchum, M., Wang, X., Wang, J., and Davis, E. 2012. Role of nematode peptides and other small molecules in plant parasitism. *Annual Review of Phytopathology* 50: 175-195.
- Mitchum, M. 2016. Soybean resistance to the soybean cyst nematode *Heterodera glycines*: an update. *Phytopathology* 106: 1444-1450.
- Mittler, R., Vanderauwera, S., Gollery, M., and Van Breusegem, F. 2004. Reactive oxygen gene network of plants. *Trends in Plant Science* 9: 490-498.
- McDonald, A., and Nicol, J. 2005. Nematode parasites of cereals. In: Luc, M., Sikora, R., and Bridge, J. (eds) *Plant Parasitic Nematodes in Subtropical and Tropical Agriculture*. CABI Publishing, Wallingford, p. 131-191.
- Niblack, T., Lambert, K., and Tylka, G. 2006. A model plant pathogen from the kingdom Animalia: *Heterodera glycines*, the soybean cyst nematode. *Annual Review of Phytopathology* 44: 283-303.

Nicol, J., and Rivoal, R. 2007. Integrated management and biocontrol of vegetable and grain crops nematodes. In: Ciancio, A., and Mukerji, K. (eds) Global knowledge and its application for the integrated control and management of nematodes on wheat. Springer Publishing, Dordrecht, p. 243-287.

Nicol, J., Turner, S., Coyne, D., Den Nijs, L., Hockland, S., and Maafi, Z. 2011. Current nematode threats to world agriculture. In: Jones, J., Gheysen, G., and Fenoll, C. (eds) Genomics and Molecular Genetics of Plant-Nematode Interactions. Springer Publishing, Dordrecht, p. 21-43.

Ohyama, K., Shinohara, H., Ogawa-Ohnishi, M., and Matsubayashi, Y. 2009. A glycopeptide regulating stem cell fate in *Arabidopsis thaliana*. Nature Chemical Biology 5: 578-580.

Patel, N., Hamamouch, N., Li, C., Hewezi, T., Hussey, R., Baum, T., et al. 2010. A nematode effector protein similar to annexins in host plants. Journal of Experimental Botany 61: 235–248.

Petitot, A., Dereeper, A., Agbessi, M., Da Silva, C., Guy, J., Ardisson, M., and Fernandez, D. 2016. Dual RNA-seq reveals *Meloidogyne graminicola* transcriptome and candidate effectors during the interaction with rice plants. Molecular Plant Pathology 17: 860-874.

Replogle, A., Wang, J., Bleckmann, A., Hussey, R. S., Baum, T. J., Sawa, S., et al. 2011. Nematode CLE signaling in Arabidopsis requires CLAVATA2 and CORYNE. Plant Journal 65: 430–440.

Replogle, A., Wang, J., Paolillo, V., Smeda, J., Kinoshita, A., Durbak, A., et al. 2013. Synergistic interaction of CLAVATA1, CLAVATA2, and RECEPTOR-LIKE PROTEIN KINASE 2 in cyst nematode parasitism of Arabidopsis. Molecular Plant-Microbe Interactions 26: 87–96.

Ringer, C., Sardanelli, S., and Krusberg, L. 1987. Investigations of the host range of the corn cyst nematode, *Heterodera zaeae*, from Maryland. Journal of Nematology 19: 97-106.

Siddique, S., Radakovic, Z., De La Torre, C., Chronis, D., Holbein, J., Novak, O., et al. 2015. A plant-parasitic nematode releases cytokinins that control cell division and orchestrate feeding-site formation in host plants. Proceedings of the National Academy of Sciences 112: 12669-12674.

Triantaphyllou, A., 1987. Genetics of nematode parasitism on plants. In: Veech, J., and Dickson, D. (eds) Vistas on Nematology. Society of Nematologists, Hyattsville, p. 354-363.

Tylka, G., and Marett, C. 2014. Distribution of the soybean cyst nematode, *Heterodera glycines*, in the United States and Canada: 1954 to 2014. doi:10.1094/PHP/BR-14-0006.

Wubben, M. J., Su, H., Rodermel, S. R., and Baum, T. J. 2001. Susceptibility to the sugar beet cyst nematode is modulated by ethylene signal transduction in *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions* 14: 1206–1212.

Yamaguchi, Y., Ishida, T., and Sawa, S. 2016. CLE peptides and their signaling pathways in plant development. *Journal of Experimental Botany* 67: 4813-4826.

CHAPTER 2

INHERITENCE OF *HETERODERA GLYCINES* VIRULENCE GENES AND THE EFFECT OF SECONDARY SELECTION ON VIRULENCE

*Data presented here has been published in: Gardner, M., Heinz, R., Wang, J., and Mitchum, M.G. 2017. Genetics and adaptation of soybean cyst nematode to broad spectrum soybean resistance. *G3: Genes | Genomes | Genetics* 7:835-841. Contributions by authors other than Michael Gardner have been highlighted in the figure legends.

ABSTRACT

The soybean cyst nematode (SCN) *Heterodera glycines* is a major threat to soybean production made more challenging by the current limitations of natural resistance for managing this pathogen. The use of resistant host cultivars is effective but over time results in the generation of virulent nematode populations able to robustly parasitize the resistant host. In order to understand how virulence develops in SCN we utilized a single backcrossed BC₁F₂ strategy to mate a highly virulent inbred population (TN20) capable of reproducing on all current sources of resistance with an avirulent one (PA3) unable to reproduce on any of the resistant soybean lines. The offspring were then investigated to determine how virulence is inherited on the main sources of SCN resistance, derived from soybean lines Peking, PI 88788, PI 90763, and the broad spectrum resistance source PI 437654. Significantly, our results suggest virulence on PI 437654 is a multigenic recessive trait that allows the nematode to reproduce on all current sources of resistance. In addition, we examined how virulence on different sources of resistance interact by placing virulent SCN populations under secondary selection and identified a strong counter-selection between virulence on PI 88788- and PI 90763-derived resistances while

no such counter-selection existed between virulence on Peking and PI 88788 resistance sources. Our results suggest that the genes responsible for virulence on PI 88788 and PI 90763 may be different alleles at a common locus. If so, rotation of cultivars with resistance from these two sources may be an effective management protocol.

INTRODUCTION

The soybean cyst nematode (SCN) *Heterodera glycines* remains a major threat to soybean production and continues to spread wherever soybeans are grown worldwide (Tylka and Marett, 2014). Yield losses attributed to SCN on an annual basis are estimated at more than \$1.2 billion (Koenning and Wrather, 2010). To date, the most effective way of managing this pathogen has been through the use of host resistance. However, widespread deployment of resistant cultivars has led to the selection of virulent SCN adapted to the resistance utilized (Niblack et al., 2006). To address this, rotations of different resistant cultivars are currently recommended. However, much remains unknown about the mechanism of nematode adaptation to resistant cultivars, including the virulence genes responsible for overcoming resistance. Precise details about the heritability and independence of nematode virulence genes are needed to inform rotation schemes of resistant cultivars that minimize selection for SCN populations adapted to multiple resistant varieties.

The first study on the inheritance of SCN genes responsible for virulence noted that the ability to reproduce on the resistant soybean cultivars PI 88788 and Pickett was inherited independently (Triantaphyllou, 1985). Following this discovery, the *ror* (for

reproduction on a resistant host) genes were identified using a controlled SCN crossing strategy (Dong and Opperman, 1997). These *ror* genes allowed some populations to reproduce on PI 88788 and PI 90763 and were shown to be inherited in a Mendelian fashion independently of each other in a dominant and recessive manner. At the time of their discovery the resistance genes in soybean remained unknown, so the role these *ror* genes have in the interplay between nematode and host was not characterized.

It was also noted that SCN populations adapted to reproduce on one resistant soybean line were able to reproduce better on several other resistant lines but not others. In this way it was shown that culturing nematodes on soybean lines PI 88788, PI 209332, or PI 548316 (Cloud) resulted in SCN populations that grow well on these lines but not on PI 548402 (Peking), PI 90763, and PI 89772 (Anand and Brar, 1983; McCann et al., 1982; Young, 1994). Based on these observations, two genetic groups were proposed, with group one containing PI 88788, PI 209332, and Cloud while group two contains Peking, PI 90763, and PI 89772 (Colgrove and Niblack, 2008; Luedders and Dropkin, 1983). In recent years much has been discovered about the genes responsible for resistance in both groups, specifically underlying two quantitative trait loci named *Rhg1* (for resistance to *H. glycines*) and *Rhg4* (reviewed in Mitchum, 2016). Resistant soybean lines in group one only require *Rhg1* for resistance to SCN HG type 0 (Race 3) populations. These lines have a variable number of tandem repeats containing a set of four dissimilar genes (i.e., 7-10 copies) at *Rhg1* (Cook et al., 2012; Lee et al., 2015). Three of the genes within the repeat were shown to be required for resistance. By contrast, soybean lines in group two contain only a few repeats (i.e., 1-3 copies) at *Rhg1* (Cook et al., 2012; Lee et al., 2015) but also require a second QTL, *Rhg4*, where a single

serine hydroxymethyltransferase (SHMT) gene controls resistance to SCN HG type 0 (Race 3) populations (Liu et al., 2012). While great strides have been made to identify some of the soybean genes controlling resistance, the mechanisms by which virulent SCN overcome these two major types of resistance, often referred to as Peking-type and PI 88788-type, remain unknown.

PI 437654 is the first reported soybean line to show resistance to nematodes adapted to each of the two soybean resistance groups (Anand et al., 1985). This broad resistance is a stark contrast to previous known resistance and may provide a valuable alternative to rotating between the other two resistance groups. To date, very few field populations of SCN have been identified that are able to successfully reproduce on PI 437654, although the numbers have been increasing with the increased use of this source of resistance. However, synthetic inbred populations such as LY1 and TN20 that can reproduce on PI 437654 have been developed. LY1 was generated from a mass mating of SCN Race 2 (HG Type 1.2.-) females with SCN Race 5 (HG Type 2.-) males (Arelli et al., 2009). The TN20 inbred population was generated by single cyst-descent of LY1 on PI 437654 in 1999, and has since been mass-selected on PI 437654 (Atibalentja et al., 2005). Interestingly, populations that can reproduce on PI 437654 also show increased reproductive success on all other resistant indicator lines used in the HG type test (HG Type 1-7; Niblack et al., 2002) and so continuous usage of this cultivar is not recommended (Niblack et al., 2005). Because of the increased use of PI 437654-derived resistance in commercial soybean lines and identification of these virulent individuals in SCN field populations (Niblack et al., 2006) it is important to understand how this resistance is related to PI 88788 and Peking, the two main sources of resistance currently

in use, and subsequently how SCN virulence on PI 437654 is related to virulence on these sources.

In this study, we set forth to describe the genetic factors that allow the SCN population TN20 to parasitize PI 437654 using a controlled crossing method. Once this was understood we put it into context with SCN reproduction on other resistant soybean lines using secondary selection. Together this information provides novel insights from the nematode perspective on overcoming host resistance and the first possible evidence for multigenic virulence in SCN.

RESULTS

Crosses, F₁, and F₂ generation host range tests

Reciprocal crosses were conducted between the PA3 population which does not reproduce on resistant cultivars, and the TN20 population which can reproduce on all sources of resistance. The PA3 population had a FI of less than 10% on the resistant indicator lines, whereas TN20 had a FI of nearly 50% or more on all resistant indicator lines at the start of this study. Figure 2.1 depicts the F₁ generation host range test results from reciprocal crosses of PA3 x TN20. The F₁ generation from PA3 x TN20 and its reciprocal cross produced a FI of >10% on Peking, PI 88788, PI 90763, PI 209332, PI 89772, PI 548316 (Cloud), and Pickett indicating that at least one virulence gene for reproduction on each of these resistant hosts in TN20 is dominant or partially dominant. The relatively high FI (49-78%) of the F₁ generation on PI 88788, PI 209332, Cloud and Pickett, consistent with the TN20 parent population, suggests that virulence is dominant.

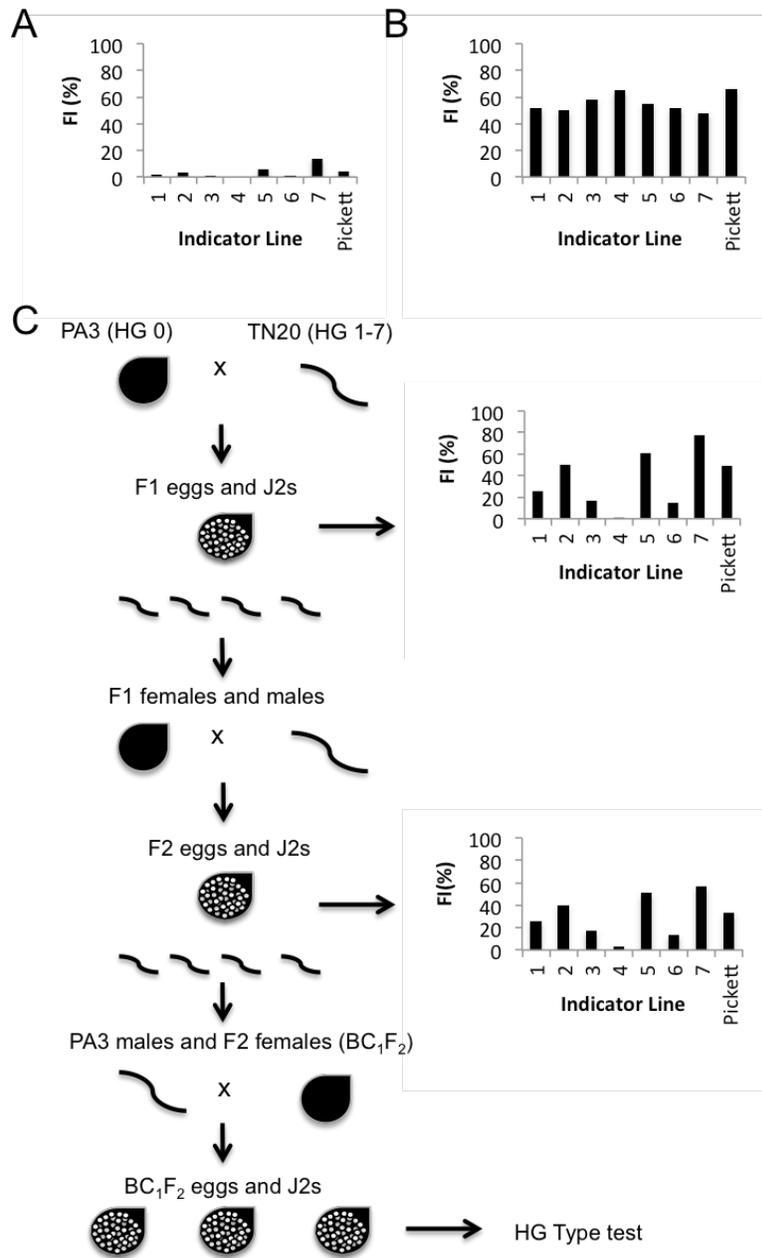


Figure 2.1. Crossing strategy for determining inheritance of virulence genes. (A) Avirulent females (PA3) were crossed with (B) highly virulent males (TN20) and the F₁ offspring were phenotyped using an HG type test. (C) F₁ males and females were then crossed and the resulting F₂ generation was backcrossed to PA3 in a BC₁F₂ strategy (drawing adapted from Dong and Opperman, 1997). The F₂ population was also outselected on eight resistant soybean indicator lines to generate inbred lines MM7-MM14. Crosses and development of out-selected populations were conducted by Robert Heinz and Jianying Wang.

The intermediate FI (15-26%) of the F₁ generation on Peking, PI 90763, and PI 89772 suggests that virulence is not controlled by a single dominant or recessive gene. In contrast, no reproduction was observed in the F₁ generation on PI 437654. Therefore, virulence in TN20 to PI 437654 is recessive. Results of reciprocal crosses indicated that there is no preferential effect of the donor parent (*i.e.*, sex linkage) on inheritance of virulence genes consistent with prior studies (Dong and Opperman, 1997). Host range test results of the F₂ generation were similar, except there was a slight increase in the FI from 0 to 3% on PI 437654.

Progeny lines and parasitism gene number

Ninety individual bulk progeny lines were established from [(TN20 X PA3) X PA3]. Results from the parasitism tests are depicted in Tables 2.1. BC₁F₂-derived single female descent lines should segregate into a 3:1 ratio for parasitism:nonparasitism for a single gene trait or 15:1 for two separate genes. The progeny lines yielded 68 parasitic and 22 non-parasitic phenotypes on PI 88788. The observed ratio was not significantly different from the 3:1 ratio ($P=0.903$) (Table 1). These results show that there is a single dominant gene in TN20 that confers the ability to parasitize PI 88788. This gene may correspond to *ror-1* described by Dong and Opperman (1997). Evaluation of the same progeny lines for the ability to parasitize PI 90763 and Peking yielded similar results. In both cases, the observed ratio was not significantly different from the 3:1 ratio ($P = 0.114$ for PI 90763; $P = 0.068$ for Peking). Prior studies reported *ror-2* and *ror-3*, recessive genes for virulence on PI 90763 and Peking, respectively (Dong and Opperman, 1997; Dong et al., 2005). Results from the

Cross ^a	PI ^b	Lines ^c	P:Np ^d	Exp-rt ^e	chi-square	P
[(PA3 x TN20) x PA3]	Peking	90	30:60	3:1	3.333	0.068
[(PA3 x TN20) x PA3]	PI 88788	90	68:22	3:1	0.015	0.903
[(PA3 x TN20) x PA3]	PI 90763	90	16:74	3:1	2.504	0.114
[(PA3 x TN20) x PA3]	PI 437654	90	0:90	3:1	270	<0.001

Table 2.1. Host range test of PA3 x TN20 segregating progeny lines on Peking, PI 88788, PI 90763, and PI 437654. ^aTN20 was the parasitic and PA3 was the nonparasitic *H. glycines* inbred line on Peking, PI 88788, PI 90763, and PI 437654. The BC₁F₂ mating strategy used PA3 as the female and recurrent parent in a single backcross. ^bPlant introduction ^cNumber of progeny lines from the mating ^dNumbers of parasitic (P) and nonparasitic (Np) lines on the indicator lines. Parasitic lines were scored based on a >10% female index. ^eExpected segregation ratio for a single gene.

progeny lines also revealed an unexpected phenomenon. Only 7 of the 90 progeny lines showed any reproduction on PI 437654 and the highest FI observed was only 2%. These data indicate that virulence on PI 437654 is unlikely to be controlled by an independent recessive gene, but rather a combination of the existing *ror* genes in the nematode.

Host-ranges of outselected inbred populations

Figure 2.2 depicts the host range test results of inbred lines MM7-MM14 developed from the PA3 x TN20 cross population outselected at least 72 generations on each of the seven HG indicator lines and Pickett. Following selection, the populations primarily fell into two groups. The first group had elevated virulence (>40% FI) on Peking, PI 90763, and PI 89772. The second group had elevated virulence on PI 88788, PI 209332, and Cloud. Cross populations MM7 (Peking), MM9 (PI 90763), and MM12 (PI 89772) fall into the first group while MM8 (PI 88788), MM11 (PI 209332), MM13 (Cloud) fall into the second group. The exceptions to this were the population selected on PI 437654 (MM10), which maintained high virulence on all indicator lines screened and the population selected on Pickett, which maintained a low level of virulence on Peking, PI 90763, and PI 89772 (Figure 2.2).

Secondary selection

We then used the MM7, MM8, and MM10 inbred populations in a secondary selection experiment. MM7, which was originally selected on Peking, was placed under secondary selection on PI 88788 and the SCN susceptible cultivar W82; MM8, originally selected on PI 88788, was placed under secondary selection on Peking and W82; MM10,

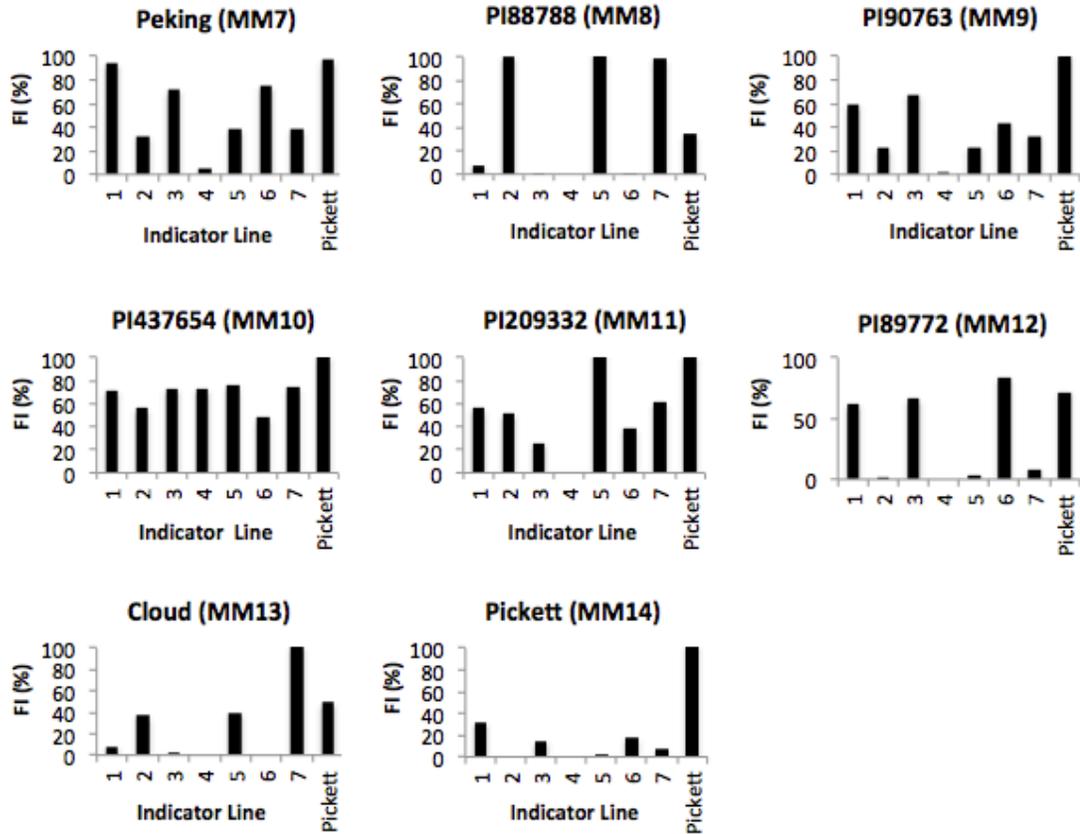


Figure 2.2. Outselected inbred populations form a cross between PA3 x TN20. Females successful on each of the lines in the F₂ SCN HG type test were used to generate SCN inbred populations MM7-MM14 through mass selection for a minimum of 72 generations on each indicator line. HG type tests conducted by Robert Heinz.

originally selected on PI 437654 was placed under secondary selection on Peking, PI 88788, and W82. At intervals of six and twelve generations the virulence of these populations was assessed using a modified HG type test restricted to four of the normal seven indicator lines. The MM7 population selected on PI 88788 gradually increased in virulence while still maintaining high virulence on Peking. In addition, MM7 virulence on PI 90763 was significantly decreased after twelve generations of selection on PI 88788 (Figure 2.3A). Importantly, no change in virulence was observed in MM7 selected on the susceptible soybean cv. W82. The MM8 population selected on Peking maintained high virulence on PI 88788 and greatly increased in virulence on Peking, but only showed a minor increase in virulence on PI 90763 (Figure 2.3B). Like MM7, no change in virulence was observed in MM8 selected on susceptible soybean W82. Finally, MM10, which was originally selected on PI 437654 maintained high virulence on all indicator lines after 12 generations of secondary selection on PI 88788, Peking, or W82 (Figure 2.3C).

DISCUSSION

Previous reports have implicated independent genes inherited in both a dominant and a recessive manner in SCN virulence on a variety of resistance sources (Dong and Opperman, 1997). The offspring of the PA3 X TN20 F₁ generation support these previous findings, with high FI on PI 88788 where the *ror-1* gene is inherited in a dominant manner and recessive genes, *ror-2* and *ror-3* for reproduction on PI 90763 and Peking, respectively (Dong et al., 2005). However, nothing was known about the inheritance of virulence genes on PI 437654-derived resistance or the impact of sex

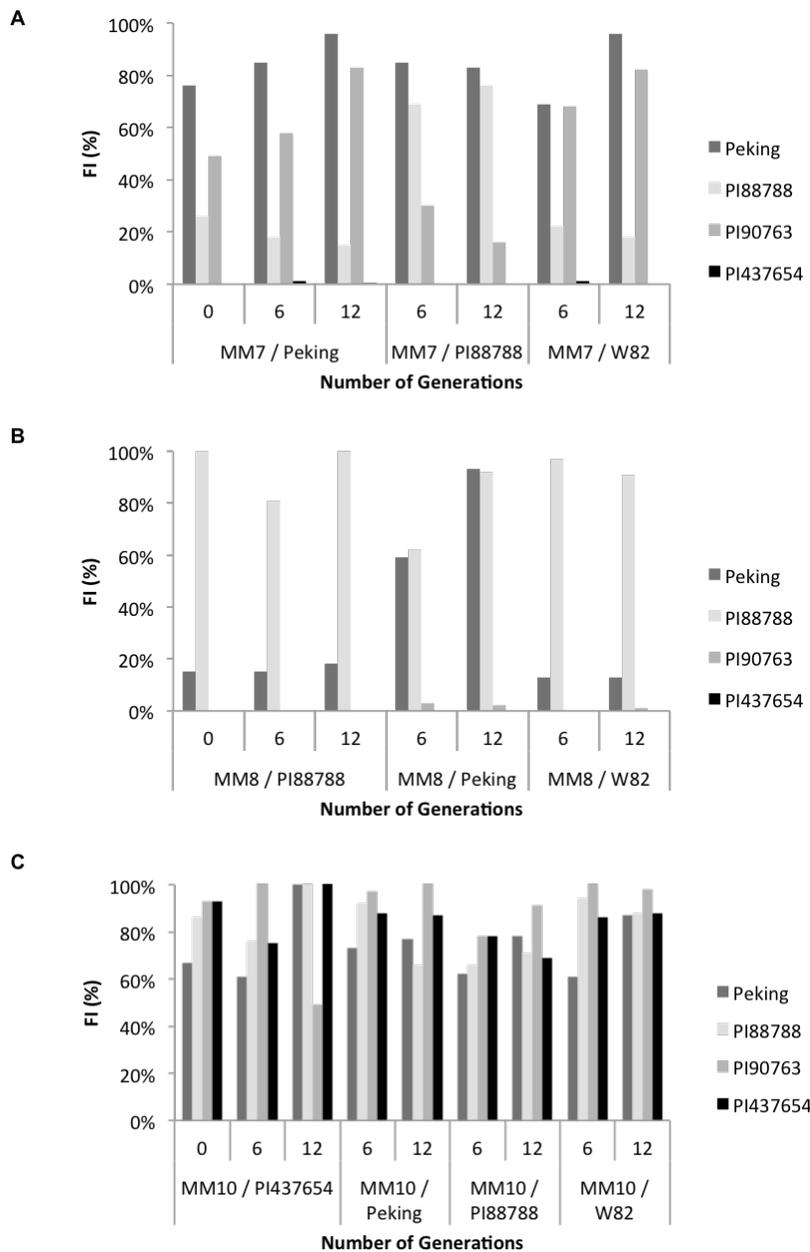


Figure 2.3. Secondary selection of outselected inbred populations. Outselected inbred populations MM7 (A), MM8 (B), and MM10 (C) were placed under secondary selection on different soybean resistance lines or on susceptible soybean. HG type tests were conducted at 0, 6, and 12 generations to assess shifts in virulence as a result of secondary selection.

linkage upon virulence gene inheritance. Our tests sought to describe genes responsible for virulence on PI 437654 and try to distinguish if these genes are independent of those required for virulence on PI 88788 and Peking resistance. To determine if there was any relationship between sex of the parent in the initial cross and the virulence profile of the offspring we performed the cross with the virulent parent as both a male and a female. The resulting F₁ and F₂ offspring had a nearly identical HG type, indicating that the virulence genes were likely inherited independently of sex.

Following the initial cross between TN20 and PA3 almost no females were observed within the F₁ generation on PI 437654. Within the F₂ generation there were slightly more females on PI 437654, but still many fewer than on other resistant soybean lines. This indicated that the virulence gene (s) responsible for reproduction on PI 437654 is inherited in a recessive manner. By limiting the population to those individuals successfully able to infect PI 437654 we were able to generate a robust HG type 1-7 population. The fact that these individuals were able to reproduce on PI 88788 and Peking despite never having been exposed to these sources of resistance was an indication that the gene (s) responsible for virulence on PI 437654 also conferred virulence on these other sources. This is also consistent with the mass-mating studies of Race 2 and Race 5 SCN populations to generate the LY1 inbred population (Arelli et al., 2009). Race 2 can reproduce on Peking and PI 88788, but not PI 90763, whereas Race 5 can reproduce on PI 88788, but not Peking and PI 90763. Following selection of the cross population on Hartwig, which derives resistance from PI 437654, a robust HG Type 1-7 population was generated. The striking absence of virulent individuals on PI 437654 in host range tests of TN20 X PA3 segregating progeny lines suggests that virulence on PI

437654 may be multigenic and therefore specified by a combination of *ror* genes inherited in a recessive fashion allowing it to reproduce on all current sources of resistance.

The inbred SCN populations outselected from the F₂ population on each of the indicator lines (MM7-MM14) normally used in a HG type test allowed us to follow nematode adaptation to each resistance source. In the outselected SCN populations, we observed distinct 1.3.6 or 2.5.7 HG types for those populations selected on PI 88788, PI 89772, Cloud, and Pickett following closely with patterns observed in previous research (Dong et al., 2005; Colgrove and Niblack, 2008). By contrast, populations outselected on Peking, PI 90763, and PI 209332 showed a 1.2.3.5.6.7 HG type with elevated FI on indicator lines 1, 3, and 6 for Peking and PI 90763 outselected populations and elevated FI on indicator lines 5 and 7 for the PI 209332 outselected population. SCN HG type 1.2.3.5.6.7 populations contain a combination of the genes required for virulence on Peking-type and PI 88788-type resistances yet still lack virulence on PI 437654. Also of note are the results observed for the outselected population on Pickett. Pickett was used in the original race determination scheme for SCN populations, but was dropped with the adoption of the HG type system because it derives resistance from Peking and was thus deemed redundant (Niblack et al., 2002). Based on the results of this study, Pickett has likely inherited genes for resistance to HG Type 0 (Race 3), but not other resistance genes in the Peking background (Rao-Arelli et al., 1992). This likely explains why field populations are commonly identified that can reproduce on both Peking and Pickett (Races 2, 4, 9, 14), Pickett and not Peking (Races 5, 6, 10, 15), but not the opposite combination (Peking and not Pickett; Races 11, 12, 13, 16).

Secondary selection of the three virulent inbred populations MM7, MM8, and MM10 also revealed interesting connections between mechanisms of virulence on the PI 88788, Peking, and PI 90763 resistance sources. Following secondary selection on PI 88788, MM7 shifted from its initial strong virulence on Peking and PI 90763 to be highly virulent on Peking and PI 88788. Meanwhile, MM7 lost virulence on PI 90763, indicating reciprocal secondary selection on this source of resistance. This was again visible in the MM8 population as secondary selection on Peking increased virulence on Peking, but not PI 90763. Both of these observations support the conclusion that there is an antagonistic relationship between virulence on PI 88788 and PI 90763. Our results support earlier work of Luedders and Dropkin (1983) who reported the same phenomena for different SCN populations on Cloud and PI 89772 or PI 88788 and PI 90763. The authors suggested that some SCN genes for reproduction on genetically different soybeans may be alleles, although linkage is possible. Further evidence exists when examining selection of a Race 5 SCN isolate on Bedford (PI 88788-derived resistance) and Cordell (Peking and PI 90763-derived resistances) soybean lines (Young, 1994). SCN selected on Cordell for 14 generations lost virulence on PI 88788-derived resistant soybean while gaining virulence on Peking and PI 90763-derived resistant soybean. Here the authors again suggest that the genes responsible for virulence on PI 88788 and PI 90763 may be different alleles at a common locus. If so, rotation of cultivars with resistance from these two sources may be an effective management protocol. The secondary selection studies also show a lack of reciprocal secondary selection between the PI 88788 and Peking resistance sources. Following secondary selection, both MM7 and MM8 maintained high FI on PI 88788 and Peking, but not PI 90763. This is

congruent with observations of secondary selection on PI 89772 using SCN populations initially selected on PI 209332, where secondary selection did not completely eliminate virulence on the initial soybean host while gaining virulence on the secondary host (Luedders, 1985).

In summary, this study illustrates how nematodes that have adapted to PI 437654 are able to overcome all sources of resistance and highlights the potential threat of this type of population to soybean production. PI 437654 was the first source of resistance identified to have resistance to races 3, 4, and 5. Anand et al (1985) proposed that the use of this line in cultivar development could eliminate the need to combine resistance genes from PI 88788 and PI 90763 to achieve multi-race resistance (Anand et al., 1985) in contrast to generating cultivars with resistance from PI 88788 and PI 90763 and using these in rotation. Our results suggest that HG type 1-7 SCN populations are likely to develop through a combination of *ror* genes and therefore generating cultivars with resistance from PI 88788 and PI 90763 and using these in rotation may be a better strategy. In addition, all virulent (MM7, MM8, MM10) populations retained their high FI on resistant soybean even following secondary selection on the susceptible soybean W82 for 12 generations indicating that virulence genes remain fixed in a population in the absence of selection pressure.

Together with previous studies, the research presented here has implications on the management of SCN using natural resistance. Currently, extension agencies generally advise to rotate between resistant soybean lines with different sources of resistance when possible to slow the build-up of virulent SCN populations (Niblack et al., 2012). This may be advisable in some situations and not others as effective management is likely to

depend on the resistant cultivar-SCN population combination. As the molecular details of the soybean-SCN interaction are elucidated through the cloning of soybean resistance genes (*Rhg*) and SCN virulence genes (*ror*) genes, management recommendations can be tailored for more strategic management of field populations.

MATERIALS AND METHODS

Parental lines

Inbred soybean cyst nematode (*Heterodera glycines* Ichinohe) lines PA3 and TN20 were used in these experiments. PA3 and TN20 were maintained on the SCN-susceptible cultivar Williams 82 (W82) and PI 437654, respectively. The female index (FI), where $FI = (\text{mean number of females on a test soybean line}) / (\text{mean number of females on the standard susceptible}) \times 100$, is used to describe the percentage of an SCN population that can reproduce on a resistant line. The *ror* genes are present in extremely low frequency in the PA3 population that results in a female index (FI) of less than 10 on any one of the seven plant introductions (PI) used in the HG type test (Niblack et al., 2002). These include PI 548402 or Peking (1), PI 88788 (2), PI 90763 (3), PI 437654 (4), PI 209332 (5), PI 89772 (6), and PI 548316 or Cloud (7). The *ror* genes are present in high frequency in the TN20 population that results in a FI of greater than 50 on all seven indicator lines. The host ranges of these populations were tested in the greenhouse with four replications and 1000 eggs per plant on SCN-susceptible cultivar Lee74, the seven HG test indicator lines, and Pickett (formerly used in the race designation scheme). PA3

is a HG Type 0 (Race 3) population. TN20 is a HG Type 1-7 (Race 4) population (Figure 2.1).

Crosses and F₁ generation testing

Reciprocal cross matings between PA3 and TN20 were conducted using a bulk sample of females from one line and a bulk sample of males from another line in soil. Juveniles were hatched from eggs of freshly harvested females of each inbred line and inoculated onto 7-day-old W82 seedlings in 400-ml tri-cornered plastic beakers. After 7-10 days, the roots were carefully washed free of soil, and the plants were suspended in half-strength Hoagland's solution and continuously aerated by bubbling air from aquarium pumps at 27C. Males began to emerge from the roots 14 days after inoculation and sunk to the bottom of the tube. They were collected on a 400-mesh sieve every day between 15-20 days post-inoculation. Plants with virgin females were transferred to pots and inoculated with males. Fifteen days after mating, the females were harvested, bulked and crushed to release F₁ eggs. Eggs of the F₁ generation from reciprocal crosses were advanced to the F₂ generation on Lee74 and used for host range tests on Lee74, Peking, PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, Cloud, and Pickett. Tests were conducted according to Brown et al (2010) with four replications and 1,000 eggs per plant. F₁ generation females from each host were counted.

Mating and backcrossing

We used a backcross 1 F₂ (BC₁F₂)-derived mating strategy (Dong and Opperman, 1997). The female and recurrent parent PA3 was nonparasitic (avirulent) on resistant

hosts, and the male parent, TN20, was parasitic (virulent). In the BC₁F₂ strategy, the individual eggs, juveniles and virgin females segregate for parasitism:nonparasitism at the F₂ generation. The backcross was made at this stage, and individual fertilized females were selected, disrupted to release eggs and individual batches of eggs were inoculated to the susceptible soybean Lee74 to generate 90 progeny lines. After several generations of amplification, individual plants were harvested and cysts were isolated. These individual bulk progeny lines were used for host range testing.

Cross population selections

Eggs of the F₂ generation from reciprocal crosses were used for host range tests as described above. F₂ generation females from each host were counted, crushed to release eggs, and inoculated back to Peking, PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, Cloud, and Pickett. These outselected populations, named MM7-MM14 were inbred for at least 72 generations following which host-range tests were performed according to above.

Secondary selection

Eggs from the outselected populations MM7, MM8, and MM10 were used to inoculate the susceptible soybean variety W82 and a different resistant soybean line. MM7 was inoculated to PI 88788, MM8 to Peking, and MM10 to both PI 88788 and Peking. The resulting offspring were propagated under the same conditions on the same host for a total of twelve generations. At zero, six, and twelve generations host range tests were conducted on Lee, Peking, PI 88788, PI 90763, and PI 437654.

ACKNOWLEDGEMENTS

This research was made possible through support from the Missouri Soybean Merchandising Council, United Soybean Board, North Central Soybean Research Program, and a USDA Special Grant to MU for SCN research. We thank Prakash Arelli and Terry Niblack for the PA3 and TN20 SCN populations, respectively, and Charlie Opperman for useful discussions and peer review of the manuscript. MNG was supported in part by a MU Life Sciences Graduate Research Fellowship.

REFERENCES

- Anand, S.C., and Brar, G.S. 1983. Response of soybean lines to differentially selected cultures of soybean cyst nematode *Heterodera glycines* Ichinohe. *Journal of Nematology* 15: 120-123.
- Anand, S.C., Wrather, J.A., and Shumway, C.R. 1985. Soybean genotypes with resistance to races of soybean cyst nematode. *Crop Science* 25: 1073-1075.
- Arelli, P.R., Young, L.D., and Concibido, V.C. 2009. Inheritance of resistance in soybean PI 567516C to LY1 nematode population infecting cv. Hartwig. *Euphytica* 165:1-4.
- Atibalentja, N., Bekal, S., Domier, L.L., Niblack, T.L., Noel, G.R., and Lambert, K.N. 2005. A genetic linkage map of the soybean cyst nematode *Heterodera glycines*. *Molecular Genetics and Genomics* 273: 273-281.
- Brown, S., Yeckel, G., Heinz, R., Clark, K., Sleper, D., and Mitchum, M.G. 2010. A high-throughput automated technique for counting females of *Heterodera glycines* using a fluorescence-based imaging system. *Journal of Nematology* 42: 201-206.
- Colgrove, A.L., and Niblack, T. 2008. Correlation of female indices from virulence assays on inbred lines and field populations of *Heterodera glycines*. *Journal of Nematology* 40: 39-45.
- Cook, D., Lee, T.G., Guo, X., Melito, S., Wang, K., Bayless, A., et al. 2012. Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338: 1206-1209.
- Dong, K., and Opperman, C. 1997. Genetic analysis of parasitism in the soybean cyst nematode *Heterodera glycines*. *Genetics* 146: 1311-1318.
- Dong, K., Barker, K.R., and Opperman, C. 2005. Virulence genes in *Heterodera glycines*: allele frequencies and *Ror* gene groups among field isolates and inbred lines. *Phytopathology* 95: 186-191.
- Koenning, S.R., and Wrather, J.A. 2010. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. *Plant Health Progress* doi:10.1094/PHP-2010-1122-01-RS.
- Lee, T.G., Kumar, I., Diers, B.W., and Hudson, M.E. 2015. Evolution and selection of *Rhg1*, a copy-number variant nematode-resistance locus. *Molecular Ecology* 24: 1774-1791.

- Liu, S., Kandoth, P.K., Warren, S.D., Yeckel, G., Heinz, R., Alden, J., et al. 2012. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492: 256-260.
- Luedders, V.D., and Dropkin, V.H. 1983. Effect of secondary selection on cyst nematode reproduction on soybeans. *Crop Science* 23: 263-264.
- Luedders, V.D. 1985. Selection and inbreeding of *Heterodera glycines* on *Glycine max*. *Journal of Nematology* 17:400-404.
- McCann, J., Luedders, V.D., and Dropkin, V.H. 1982. Selection and reproduction of soybean cyst nematodes on resistant soybeans. *Crop Science* 22.1: 78-80.
- Mitchum, M.G. 2016. Soybean resistance to the soybean cyst nematode *Heterodera glycines*: an update. *Phytopathology* 106: 1444-1450.
- Niblack, T.L., Arelli, P.R., Noel, G.R., Opperman, C.H., Orf, J.H., Schmitt, D.P., et al. 2002. A revised classification scheme for genetically diverse populations of *Heterodera glycines*. *Journal of Nematology* 34: 279-288.
- Niblack, T.L. 2005. Soybean cyst nematode management reconsidered. *Plant Disease* 89: 1020-1026.
- Niblack, T.L., Lambert, K.N., and Tylka, G.L. 2006. A model plant pathogen from the kingdom Animalia: *Heterodera glycines*, the soybean cyst nematode. *Annual Review of Phytopathology* 44: 283-303.
- Niblack, T.L., and Tylka, G.L. 2012. Soybean cyst nematode management guide: 5th edition. University of Illinois. http://www.ncsrp.com/pdf_doc/SCN_Management.pdf
- Rao-Areli, A.P., Anand, S.C., and Walker, J.A. 1992. Soybean resistance to soybean cyst nematode race 3 is conditioned by an additional dominant gene. *Crop Science* 32:862-864.
- Triantaphyllou, A.C., 1975. Genetic structure of races of *Heterodera glycines* and inheritance of ability to reproduce on resistant soybeans. *Journal of Nematology* 7:356-364.
- Triantaphyllou, A C., 1987. Genetics of nematode parasitism on plants. In: Veech, J., and Dickson, D. (eds) *Vistas on Nematology*. Society of Nematologists, Maryland p. 354-363.
- Tylka, G.L., and Marett, C.C. 2014. Distribution of the soybean cyst nematode, *Heterodera glycines*, in the United States and Canada: 1954-2014. *Plant Health Progress* 15:13-15.

Young, L.D. 1982. Reproduction of differentially selected soybean cyst nematode populations on soybeans. *Crop Science* 22: 385-388.

Young, L.D. 1984. Changes in reproduction of *Heterodera glycines* on different lines of *Glycine max.* *Journal of Nematology* 16:304-309.

Young, L.D. 1994. Changes in reproduction of a *Heterodera glycines* race 5 isolate cultured on 'Cordell' and 'Bedford' soybean. *Journal of Nematology* 26: 653-655.

CHAPTER 3

NOVEL EFFECTOR MINING FROM A *DE NOVO* TRANSCRIPTOME ASSEMBLY OF THE EARLY PARASITIC STAGES OF THE SOYBEAN CYST NEMATODE *HETERODERA GLYCINES*

*Data described in this chapter is being prepared for publication: Gardner, M., Dhroso, A., Davis, E.L., Baum, T.J., Korkin, D., and Mitchum, M.G. Novel effector mining from a *de novo* transcriptome assembly of the early parasitic stages of the soybean cyst nematode *Heterodera glycines*. Contributions from authors other than Michael Gardner are noted in the appropriate figure legends.

ABSTRACT

Despite its economic significance, the soybean cyst nematode (SCN) *Heterodera glycines* lacks a complete reference genome or transcriptome. Molecular studies to identify effectors and putative virulence genes are greatly hindered without these genetic resources. In the absence of an available annotated SCN genome, we generated a *de novo* transcriptome assembly representing the early stages of SCN in both a compatible and an incompatible host interaction. Previously identified SCN effectors were mined from the transcriptome to determine the level of variation existing within a population of SCN. From this analysis, we were able to distinguish between effectors that co-evolve with the host genotype and those conserved by the pathogen to maintain a core function in parasitism. A dual effector prediction strategy including a traditional secreted protein prediction pipeline and a novel nematode effector prediction tool, N-Preffector, was used to identify putative effector candidates for further research. Using this approach, we

discovered several homologues of effectors from other plant-parasitic nematodes and uncovered a suite of new effector candidates. This transcriptome represents the most comprehensive whole nematode sequence currently available for SCN and can be used as a tool for annotation of the eventual SCN genome assembly. In addition to serving as a mapping resource, the transcriptome also serves as a source of genetic information to mine for further insights into cyst nematode biology.

INTRODUCTION

The soybean cyst nematode (SCN) *Heterodera glycines* is the most economically important pathogen of soybean, causing over one billion dollars in yield loss annually (Koenning and Wrather, 2010). This microscopic roundworm begins its life as an egg in the soil, undergoing one molt before hatching as a second-stage juvenile (J2). Once the nematode has hatched, it migrates through the soil towards a host plant where it invades the root tissue and migrates towards the vasculature, selecting a single cell to establish a feeding site called a syncytium. At this point, the nematode penetrates the cell wall using its stylet and releases a set of secretions into the host cell, including effector proteins. Stylet-secreted effector proteins share many characteristics including the presence of a signal peptide, lack of a transmembrane domain, and expression in the esophageal gland cells (Mitchum et al., 2013). These effector proteins manipulate the host cell by modulating a variety of cellular processes to make it more suitable for the nematode, including suppression of host defense and stress responses and causing significant transcriptional re-programming in the host cell nucleus (Gheysen and Mitchum, 2011).

Effectors harboring nuclear localization signals are recognized by host cellular machinery for targeting to the nucleus where they modulate host nuclear functions (Quentin et al., 2013). Similar to effectors delivered by the type III secretion system of bacteria or the haustorium of pathogenic fungi, these effectors represent an interface between the nematode pathogen and host (Toruño et al., 2016). Once the feeding site is established, the nematode becomes sedentary and relies entirely on the host for nutrition for the remainder of its life cycle. The nematode slowly swells up as it undergoes a series of molts and differentiates into either a male or a female. Females protrude from the roots while the males regain mobility and exit the root to fertilize females, following which the males die. The females eventually die after fertilization, their bodies hardening into a protective casing for the eggs called a cyst that breaks off into the soil and begins the process anew. The early stages of the nematode infection cycle represent a key point in determining the fate of a cyst nematode. Whether or not the nematode will survive long enough to complete its life cycle depends on the ability of the nematode to survive and circumvent the hostile environment presented by the plant host.

In recent years, next generation sequencing technologies have been applied to several phytoparasitic nematode species, resulting in the assembly of complete genomes for *Meloidogyne hapla*, *Meloidogyne incognita*, *Globodera pallida*, *Globodera rostochiensis*, *Ditylenchus destructor*, and *Bursaphelenchus xylophilus* (Abad et al., 2008; Cotton et al., 2014; Eves-van den Akker et al., 2016; Kikuchi et al., 2011; Zheng et al., 2016). Despite the enormous economic importance of *H. glycines*, no genome is currently available. In the absence of a sequenced genome, several other plant-parasitic nematode systems have turned to *de novo* transcriptome-level studies instead (Fosu-

Nyarko et al., 2016; Haegeman et al., 2013; Kumar et al., 2014; Petitot et al., 2016).

These studies were able to identify key features of the interaction of plant host and nematode pathogen, including the discovery of suites of new effectors.

In the SCN system, studies have primarily focused on identifying and characterizing stylet-secreted effectors produced in the esophageal gland cells, which has resulted in the identification of 72 SCN effectors (Bekal et al., 2015; Gao et al, 2003; Noon et al, 2015). These studies based their identification of SCN effectors on the presence of a signal peptide as well as expression in the esophageal gland cells confirmed by *in situ* hybridization. Multiple functional studies have since been performed using these effectors, identifying host targets and characterizing their role in cyst nematode parasitism (reviewed in Gardner et al., 2015). Though the approach focused on the gland cells has been highly successful in identifying stylet-secreted effector proteins, low abundance transcripts, those harboring non-canonical secretion signals, and those encoding secreted proteins originating in other structures of the nematode are lacking. In addition, a global analysis would not only provide a comprehensive assessment of effectors, but would enable studies to assess effector variation within and across populations to identify highly variable effectors potentially correlated with virulence, as well as those effectors highly conserved across the population that may be key components of the SCN infection process. Effector variation has been shown to be important in other plant pathogens such as bacteria and fungi as a tactic for evading host recognition and resistance (Jones and Dangl, 2006; Na and Gijzen, 2016).

To provide comprehensive biological insight and a tool for comparative analyses between different nematode species and populations of *H. glycines* in the absence of a

reference genome, a *de novo* transcriptome assembly of early parasitic stages was generated. An analysis of the transcriptome confirmed previous reports of microorganisms present within the nematode and parallels to other plant-parasitic nematode species. We then performed multiple analyses focused on effectors; both predicting novel effectors using a newly developed bioinformatics tool called N-Preffector that is not reliant on the presence of a signal peptide and investigating variation of previously identified stylet-secreted effector protein sequences. This allowed for the identification of an additional suite of novel effectors that may play a pivotal role in SCN infection and could serve as potential targets for development of novel SCN control strategies.

RESULTS

Transcriptome sequencing and assembly

mRNA sequencing of pre-parasitic second-stage juvenile (ppJ2) and parasitic second-stage juvenile (pJ2) life-stages yielded a total of 603.6 million paired 100 base reads. Following initial filtering steps and removal of reads mapping to the soybean genome, the final input for transcriptome assembly was 430 million reads. Trinity *de novo* transcriptome assembly resulted in a final assembly of 147,910 transcripts with a total assembly length of 46.7 Mb and estimated 23-fold transcriptome coverage. The average length of these transcripts was 658 basepairs with a N50 of 1,085 bp (Table 3.1). When translated, 78,625 resulting proteins were predicted. This transcriptome assembly was then assessed using BUSCO (Benchmarking Universal Single-Copy Orthologs,

	<i>Heterodera glycines</i> de novo assembly
Number of transcripts	147,910
Total Assembly length (Mb)	46.7
Number of Trinity 'genes'	71,093
N50 (bp)	1,085
Maximum contig size (bp)	11,112 bp
Minimum contig size (bp)	201 bp
Avg. contig length (bp)	658 bp
Predicted proteins	78,625
BUSCO score	C: 68%, F: 14%, M: 18%

Table 3.1. *de novo* transcriptome assembly statistics for the SCN early parasitic stage assembly. The assembly was generated from *H. glycines* pre-parasitic second-stage juvenile samples as well as parasitic second stage juvenile samples from susceptible and resistant host interactions using the Trinity *de novo* transcriptome assembly tool. The transcriptome was assessed for completeness using the tool BUSCO (benchmarking universal single-copy orthologs) to identify complete (C), fragmented (F), and missing (M) sequences representing conserved orthologs found in all eukaryotes.

Simão et al., 2015). Based on the 429 single copy orthologs for eukaryotes, the SCN assembly is 68% complete, with an additional 14% of the orthologs represented in fragmented transcripts and the remaining 18% missing from the transcriptome.

Annotation of transcripts

Transcripts from the *H. glycines* transcriptome were annotated following the Trinotate pipeline (Haas et al., 2013). Transcripts were first compared to GenBank, Swissprot, and TrEMBL databases using BLASTX, resulting in a total of 66,601 (45.03%) out of the 147,910 transcripts annotated using an e-value cutoff of 1e-5 (Appendix 1). When examining the species distribution of these significant hits, most transcripts hit to prior *H. glycines* database entries followed by animal-parasitic nematode species such as *Ascaris suum* and *Strongyloides ratti* (Figure 3.1). In total, 1315 species are represented in the BLASTX results. Other species of note in the annotated transcripts include a *Cardinium* endosymbiont of *Encarsia pergandiella* and several soybean cyst nematode associated viruses (Bekal et al., 2011; Bekal et al., 2014; Noel and Atibalentja, 2006; Penz et al., 2012). These transcripts were characterized in more detail as described below.

Transcripts were further compared to several nematode species with sequenced and annotated genomes representing free-living, animal-parasitic, and plant-parasitic trophic groups to identify potential overlap and genes that are uniquely shared. The *H. glycines* transcriptome uniquely shares 76 potential homologs with *Bursaphelenchus xylophilus*, 313 homologs with *Meloidogyne hapla*, 200 with *Meloidogyne incognita*, and

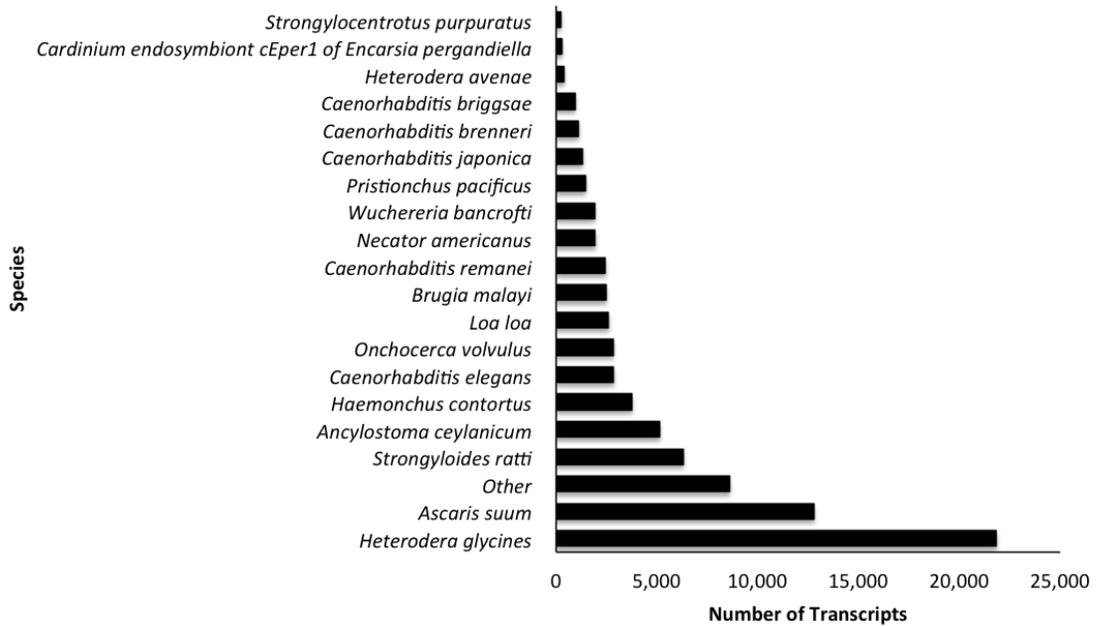


Figure 3.1. Species distribution of predicted homologues to *H. glycines*. Homologues were predicted using a BLASTX search against several protein databases at an e-value cutoff of 1e-5. The top 20 species with the most homologues are shown here.

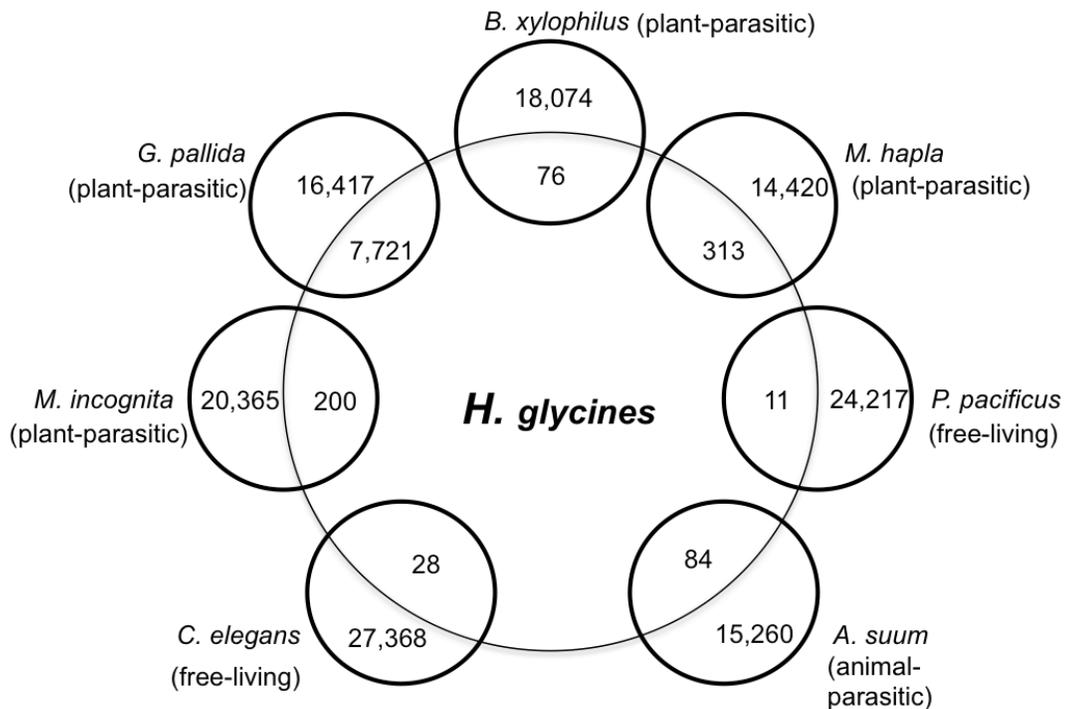


Figure 3.2. *H. glycines* orthologs in proteomes from completely sequenced nematodes with diverse feeding behaviors. Interior numbers represent *H. glycines* predicted proteins with unique orthologs to one of the other seven nematode species examined. Exterior numbers represent predicted proteins in the proteomes with no predicted ortholog in the early parasitic *H. glycines* transcriptome.

7,721 with *Globodera pallida*. In addition, the transcriptome shares 11 homologs with the free-living nematode *Pristionchus pacificus*, 28 with the free-living nematode *Caenorhabditis elegans*, and 84 homologs with the animal-parasitic nematode *Ascaris suum* (Figure 3.2; Appendix 2).

Identification and GO annotation of endosymbiont-associated transcripts from the *H. glycines* transcriptome

Within the early parasitic SCN transcriptome 468 transcripts were annotated as endosymbiont-associated transcripts, all of which were confirmed by BLASTX mapping to the *Cardinium hertigii* proteome (Figure 3.3). GO terms were assigned to the 468 endosymbiont-associated transcripts using BLAST2GO, resulting in GO annotation of 328 of the 468 transcripts (Figure 3.4). Of those sequences with GO annotation, the majority were involved in ATP binding, with 24.5% of the annotated transcripts in this category. Other significant GO categories include DNA binding (17.4%), translation (20.1%), transport (14.6%), and RNA binding (14.6%).

Known SCN stylet-secreted effector protein analysis

We next examined the 72 previously predicted stylet-secreted *H. glycines* effectors within the transcriptome (Bekal et al., 2015; Gao et al., 2003; Noon et al., 2015). Of these, transcripts corresponding to each effector were identified using a BLASTN search, indicating that the transcriptome contained sufficient depth to detect expression of the complete gland cell effector repertoire of SCN. An analysis of effector variation within the population was then performed. We first grouped the known

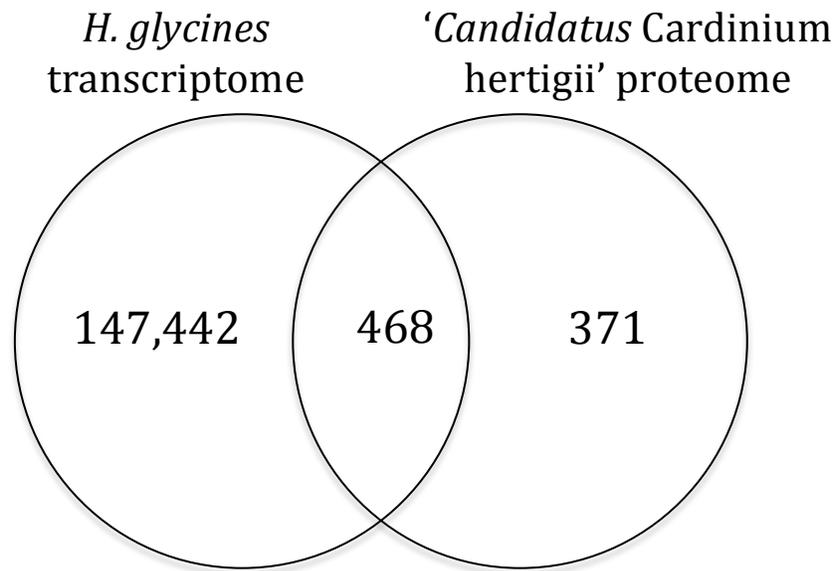


Figure 3.3. Identification of '*Candidatus Cardinium hertigii*'-associated transcripts within the SCN transcriptome. Transcripts from the *H. glycines* transcriptome were extracted and mapped against the proteome for '*Candidatus Cardinium hertigii*' to identify potential endosymbiont-associated transcripts, resulting in the identification of 468 of the 839 described proteins for this endosymbiont within the SCN transcriptome.

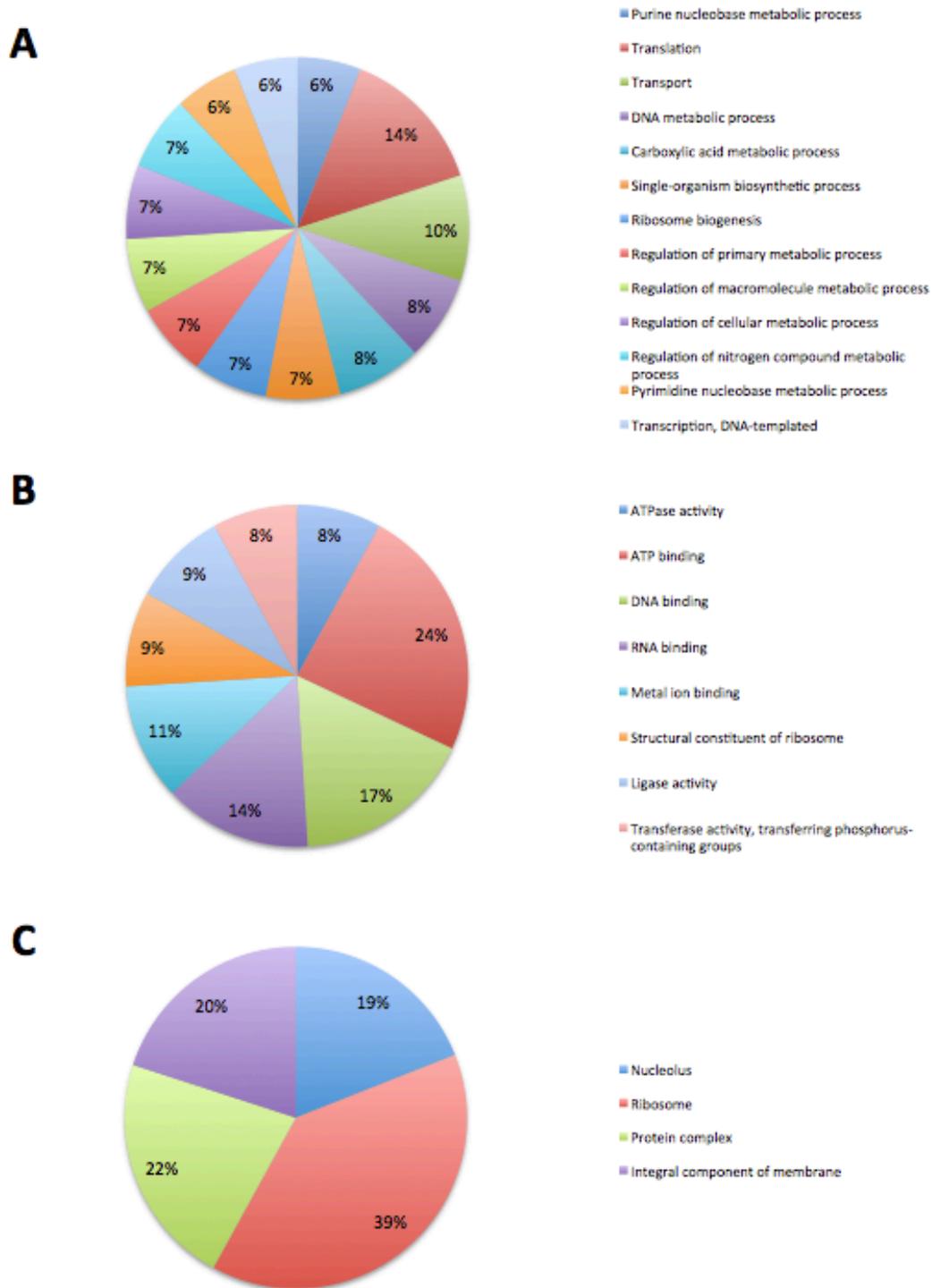


Figure 3.4. GO annotation of endosymbiont-associated transcripts identified within the SCN transcriptome. Available gene ontology annotation was added to the endosymbiont-associated transcripts by BLAST2GO and grouped by the parent terms (A) biological process, (B) molecular function, and (C) cellular component.

effectors into stylet-secreted effector families (SSEFs) with greater than 70% sequence identity. To assess the level of variation of these known effectors within the sequenced *H. glycines* population the predicted peptide sequences were mined for protein variants using BLASTP at a 1e-5 cutoff. Protein variants were identified for 69 of the 72 known effectors (Figure 3.5). The remaining three (17G06, 30C02, and GLAND9) were found to have single nucleotide insertions and/or deletions leading to a frame shift in the predicted peptide, resulting in a completely different peptide compared to the reference sequence, and consequently were not examined for sequence variation. A wide scope of variation was identified, with some effectors having over 70 predicted protein variants across the population (e.g., 4F01), while others were limited to a single, highly conserved protein sequence (e.g., 7E05).

We then examined the expression of known SCN effectors during a compatible and an incompatible interaction to determine if the host environment influences the expression of any of these effectors. The effectors were split into two different groups (upregulated or downregulated) based on their expression pattern from the ppJ2-pJ2 life stage and then compared across the two conditions (Tables 3.2 and 3.3). Most of the known SCN effectors followed the same pattern of expression across both comparisons, but the level of expression change was slightly reduced in the incompatible interaction. However, a subset of effectors exhibited an opposite trend of increased expression in the incompatible interaction, including members of SSEFs 1, 9, 17, 22, 54, 48, and 11.

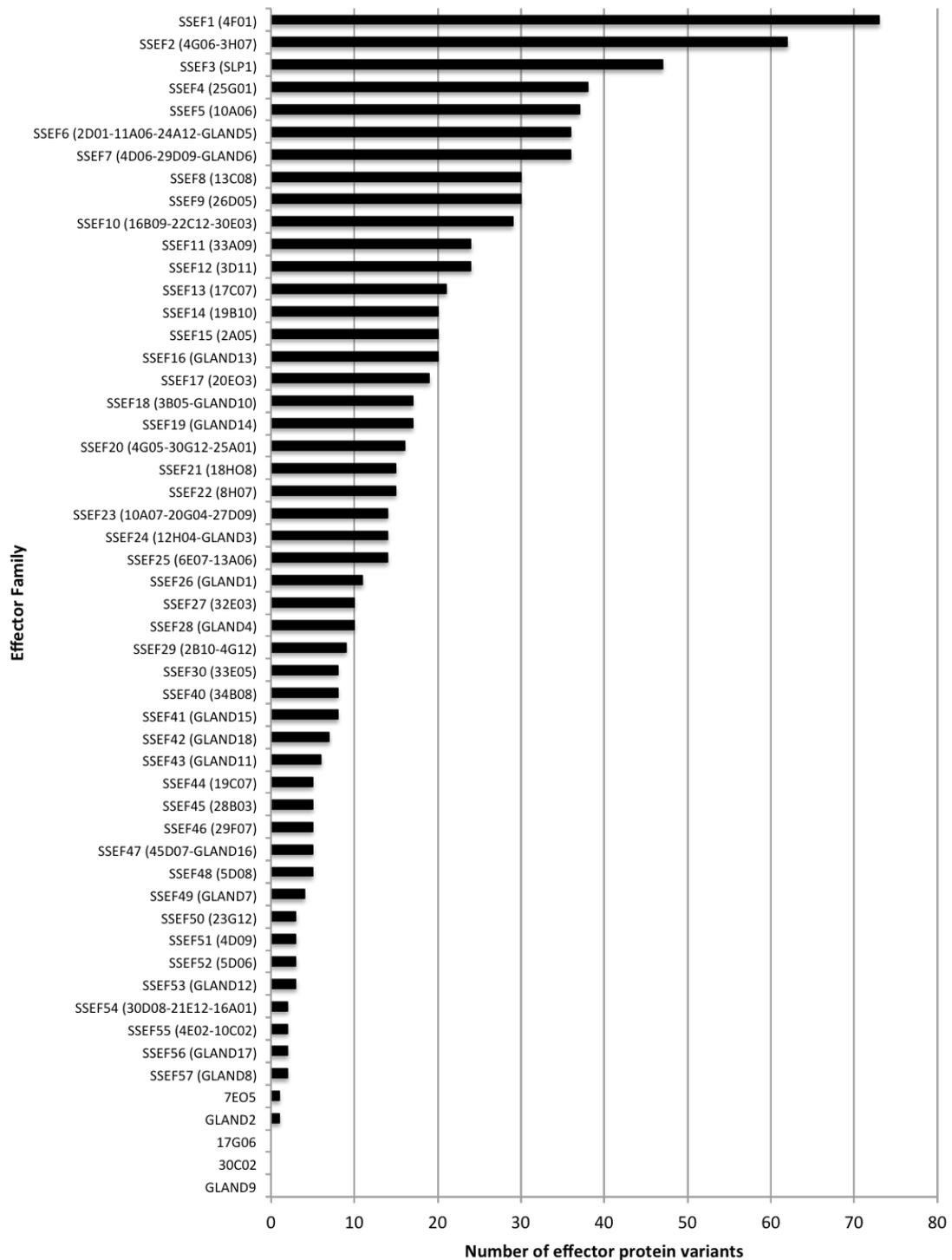


Figure 3.5. Variation of known effectors in the *H. glycines* early parasitic transcriptome. Protein variants of previously published *H. glycines* effectors (Bekal et al., 2015; Gao et al., 2003; Noon et al., 2015) were identified using a BLASTP against predicted proteins from the transcriptome at a $1e-5$ cutoff and quantified. Known effector sequences with $>70\%$ amino acid identity conserved between them were grouped into stylet-secreted effector families (SSEF) to facilitate the analysis.

Novel effector prediction

We then performed a comprehensive effector analysis on the SCN transcriptome. Effectors were predicted using two separate pipelines, and then the results were compared to determine the overlap of each pipeline (Figure 3.6). The first of these pipelines relies on the presence of a signal peptide and follows the method used in previous studies for the prediction of putative stylet-secreted effectors (Gao et al., 2003; Noon et al., 2015). This pipeline predicted 4,846 putative effectors. To identify putative new effectors with higher confidence, we focused on genes upregulated from the pre-parasitic J2 to parasitic J2 life stage and analyzed the sequences for the presence of a nuclear localization signal. Following these filtering steps, this pipeline predicted 734 effector candidates, including 139 nuclear localization signal (NLS) positive effector candidates up-regulated from the pre-parasitic J2 to the parasitic J2 life stage (Figure 3.4; Appendix 3). The 72 known SCN effector proteins, known to contain signal peptides, were re-discovered at a rate of 74% using this pipeline. This pipeline is reliant upon the presence of a N-terminal signal peptide, which may not be present if the N-terminus is absent from the transcript. This is reflected in the fact that several known SCN effectors were not recovered despite the nucleotide sequences being present within the transcriptome. A second pipeline independent of the presence of a signal peptide was performed using N-Preffector, a machine learning algorithm trained on known nematode and bacterial effectors. The N-Preffector-based pipeline predicted 1,251 putative effectors, including 338 NLS positive effector candidates up-regulated from the pre-parasitic J2 to the parasitic J2 life stage (Figure 3.4; Appendix 3). In this pipeline, 67% of the known SCN effectors were re-discovered. When the two pipelines were compared,

Up-regulated transcripts corresponding to known effectors					
Effector Name	Top BLASTN Transcript ID	ppJ2-pJ2 Compatible		ppJ2-pJ2 Incompatible	
		logFC	p-value	logFC	p-value
ZD01	c30906_g1_i6	11.95066	7.38E-19	10.74789156	3.45E-15
24A12	c30906_g1_i6	11.95066	7.38E-19	10.74789156	3.45E-15
GLAND6	c30021_g1_i5	11.21349181	7.60E-26	10.59596635	4.00E-24
2B10(SYV46)	c29696_g3_i4	11.07316385	4.24E-28	10.39148844	4.46E-26
4G12	c29696_g3_i4	11.07316385	4.24E-28	10.39148844	4.46E-26
4G05	c30424_g1_i7	9.676374354	7.68E-23	9.378915652	8.81E-22
30G12	c30424_g1_i7	9.676374354	7.68E-23	9.378915652	8.81E-22
GLAND13	c25123_g1_i1	9.193821155	4.70E-22	8.780188552	6.39E-21
16B09	c26955_g1_i1	9.138673725	2.09E-22	9.021326651	4.95E-22
22C12	c26955_g1_i1	9.138673725	2.09E-22	9.021326651	4.95E-22
30E03	c26955_g1_i1	9.138673725	2.09E-22	9.021326651	4.95E-22
GLAND4	c26387_g1_i2	9.020885275	9.16E-22	8.360785745	8.20E-20
GLAND1	c23695_g1_i6	8.688379898	1.86E-20	7.8877792	4.73E-18
10A07	c28700_g2_i12	8.605585789	2.61E-20	8.011601611	1.30E-18
20G04	c28700_g2_i12	8.605585789	2.61E-20	8.011601611	1.30E-18
27D09	c28700_g2_i12	8.605585789	2.61E-20	8.011601611	1.30E-18
3B05	c21711_g1_i2	8.131187058	1.23E-16	7.465795874	8.35E-15
32E03	c30645_g1_i2	7.589623968	7.70E-18	7.384439228	3.40E-17
11A06	c30906_g1_i5	7.488642215	1.33E-17	6.227490926	6.97E-16
6E07	c28700_g2_i14	7.343832367	5.81E-16	7.040162927	7.12E-15
13A06	c28700_g2_i14	7.343832367	5.81E-16	7.040162927	7.12E-15
GLAND12	c24211_g1_i1	6.521582447	3.73E-15	6.14953692	4.66E-14
19B10	c29810_g1_i4	6.303591759	2.93E-14	5.488786868	7.02E-12
GLAND17	c20480_g1_i1	6.225211081	3.38E-14	5.655306885	1.65E-12
GLAND5	c30906_g1_i1	5.682919123	2.67E-12	3.899841408	2.78E-07
4D06	c30021_g1_i1	5.408743446	7.82E-12	5.081153925	6.74E-11
29D09	c30021_g1_i1	5.408743446	7.82E-12	5.081153925	6.74E-11
18H08	c31157_g1_i1	4.410816182	5.69E-09	2.197474269	0.001562634
GLAND11	c24868_g1_i2	4.398087138	4.65E-09	3.885285871	1.15E-07
4F01	c29598_g1_i7	4.358867525	7.36E-09	4.578297168	1.88E-09
GLAND2	c22416_g1_i1	4.357098189	7.29E-09	4.240042733	1.55E-08
4E02	c22252_g1_i2	4.008921773	5.54E-08	3.748867886	2.74E-07
10C02	c22252_g1_i2	4.008921773	5.54E-08	3.748867886	2.74E-07
10A06	c30256_g4_i4	3.835724138	1.84E-07	3.410985388	2.39E-06
GLAND16	c16189_g1_i1	3.743120485	2.73E-07	3.385876919	2.30E-06
GLAND8	c26584_g5_i1	3.444604828	5.52E-06	1.911891167	0.010972282
7E05	c20676_g1_i1	3.388418122	1.00E-05	1.716586957	0.024732707
26D05	c27827_g2_i15	2.913589598	4.23E-05	3.509853	1.48E-06
GLAND14	c30543_g2_i10	2.791463157	9.08E-05	2.512918667	0.000381914
20E03	c25984_g1_i10	2.565313858	0.000219099	2.904484999	3.63E-05
5D06	c28325_g3_i2	2.536369336	0.000254098	1.790033676	0.008094752
8H07	c27687_g6_i1	1.98511251	0.004515331	2.487751068	0.000465562
3D11	c25690_g1_i6	1.972889683	0.003580946	1.921653429	0.004485267
GLAND7	c30220_g2_i9	1.934650864	0.004253197	1.420878165	0.0331378
30D08	c30942_g5_i1	1.517825759	0.029759309	2.397033056	0.000741241
GLAND15	c25766_g1_i6	1.268845482	0.056291715	0.84215756	0.201826305
16A01	c30942_g2_i1	0.837527119	0.208546118	1.373214663	0.040903086
21E12	c30942_g2_i1	0.837527119	0.208546118	1.373214663	0.040903086
5D08	c29247_g3_i3	0.70070194	0.308181925	2.066825767	0.002899175

Table 3.2. Up-regulated effectors from the pre-parasitic second-stage juvenile to the parasitic second-stage juvenile in both a compatible and an incompatible interaction.

Down-regulated transcripts corresponding to known effectors					
Effector Name	Top BLASTN Transcript ID	ppI2-ppI2 Compatible		ppI2-ppI2 Incompatible	
		logFC	p-value	logFC	p-value
GLAND9	c29829_g1_i4	-0.078646873	0.905553682	-0.48793881	0.458126102
17G06	c30557_g1_i1	-0.114331242	0.866057779	-0.165496811	0.803841129
GLAND18	c26584_g4_i17	-0.136662174	0.841709694	1.049450732	0.116629336
GLAND10	c26675_g1_i10	-0.257455469	0.696725637	-0.186579295	0.779841563
4D09	c26533_g5_i1	-0.286114895	0.661773357	-0.214616167	0.742740131
30C02	c25605_g1_i2	-0.413915471	0.52742605	-0.451561013	0.490756531
19C07	c28029_g1_i4	-0.525898278	0.445424257	-2.391062968	0.001212882
45D07	c23887_g1_i3	-0.602412606	0.358285797	-0.494842557	0.449987489
12H04	c29466_g3_i7	-0.660554232	0.320415347	-0.320506109	0.630704588
28B03	c30256_g4_i16	-0.744283971	0.257274898	-0.511443682	0.435021613
29F07	c30256_g4_i16	-0.744283971	0.257274898	-0.511443682	0.435021613
3H07	c30309_g5_i3	-0.798383513	0.224625693	-0.539822256	0.41006245
4G06	c30309_g5_i3	-0.798383513	0.224625693	-0.539822256	0.41006245
33E05	c27363_g1_i8	-0.981531762	0.136719108	-0.865718965	0.188457621
34B08	c27363_g1_i8	-0.981531762	0.136719108	-0.865718965	0.188457621
GLAND3	c29466_g4_i2	-1.17845815	0.075115339	0.656721076	0.31689761
25G01	c29413_g2_i8	-1.528723502	0.022086616	-1.5380044	0.021327647
6F06	c29413_g2_i11	-1.723101586	0.01031014	-1.653856432	0.013620506
13C08	c29995_g4_i4	-1.964059592	0.003727324	-2.26060053	0.000955661
17C07	c26763_g1_i3	-3.158903943	0.00000906	-2.933981217	0.0000316
23G12	c31084_g1_i19	-3.349233178	0.00000343	-3.448176267	0.00000198
2A05	c24990_g3_i2	-3.925589956	0.00000009	-4.038485845	4.51E-08
33A09	c28580_g2_i13	-5.815461954	0.001996045	-0.84421285	0.5148133

Table 3.3. Down-regulated effectors from the pre-parasitic second-stage juvenile to the parasitic second-stage juvenile in both a compatible and an incompatible interaction.

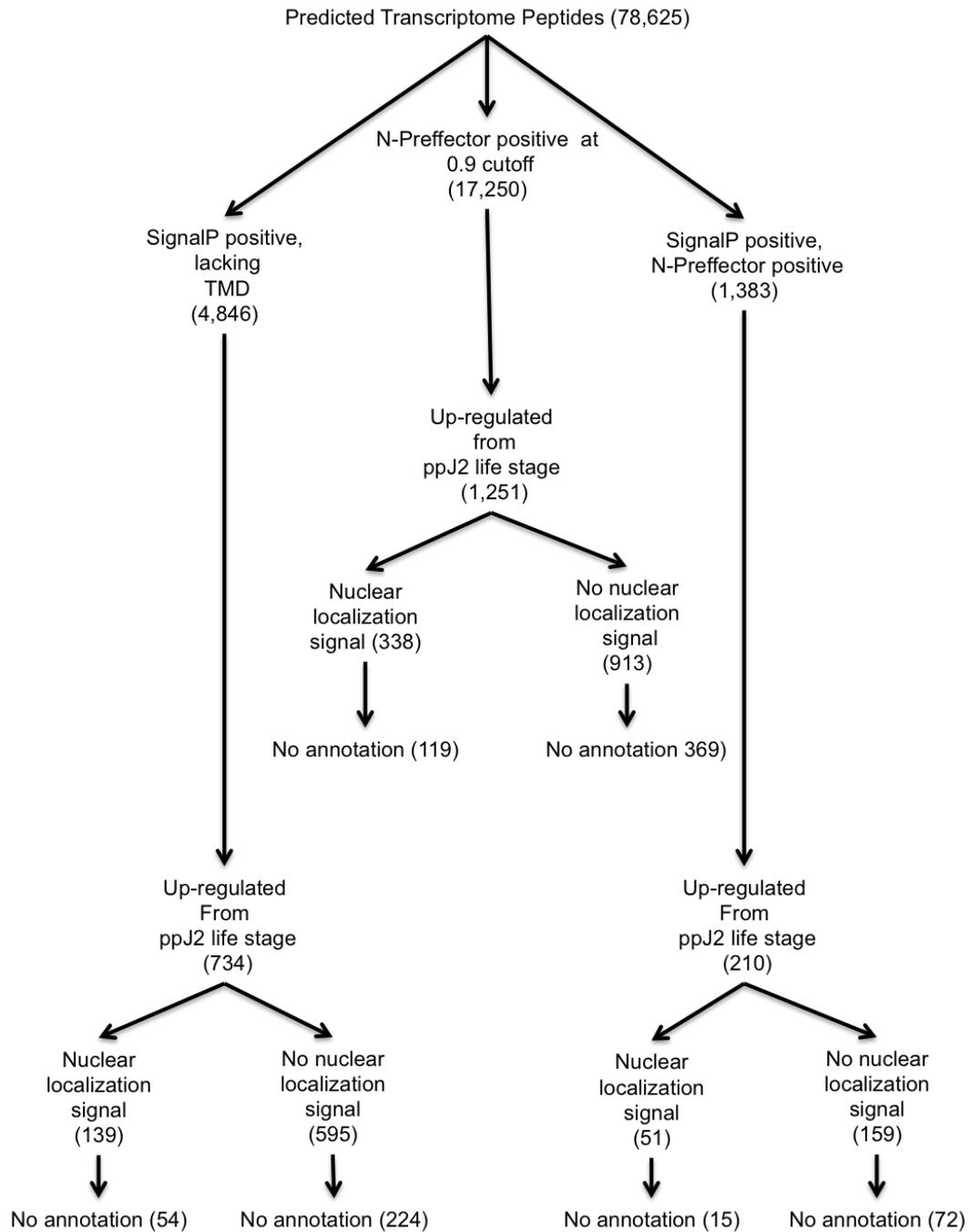


Figure 3.6. Secreted effector protein prediction in the early parasitic transcriptome of *H. glycines*. Predicted peptides from the transcriptome were put through two separate pipelines to identify effector candidates. The pipelines were then compared to identify overlapping peptides found in both pipelines. Numbers shown indicate predicted peptides remaining after each step in the pipeline. N-prefector was trained by Andi Dhroso to identify effector candidates.

210 effector candidates were found to be overlapping, including 51 NLS positive candidates (Figure 3.4; Appendix 3). Many of these sequences have little or no annotation available. Among those sequences with available annotation are many homologs of effectors from other plant-parasitic nematodes that were not identified or characterized in *H. glycines* previously (Table 3.4).

DISCUSSION

In this study, we sequenced the early stages of the plant-parasitic nematode *Heterodera glycines*, including the infective (pre-parasitic) second-stage juvenile life stage and parasitic second-stage juvenile life stage in two different host conditions, resistant and susceptible. We then carried out a *de novo* transcriptome assembly with an emphasis on assessing the level of variation of known effectors within a single population and identifying novel secreted effectors within *H. glycines* that may play important roles in establishing a parasitic interaction with its host, soybean. The resulting transcriptome from these samples consisted of nearly 150,000 transcripts encoding 78,625 predicted proteins. By comparison, a transcriptome containing two life stages of the closely related parasitic nematode *H. avenae* contained 27,765 transcripts (Kumar et al., 2014). However, the *H. avenae* transcriptome was assembled using a different assembly tool and only represented pre- parasitic J2 and adult female life stages, which could account for some of the variation in transcript number between the two transcriptomes. Of the 147,910 transcripts contained within the *H. glycines* transcriptome, 66,601 (48%) were annotated based on BLAST homology, whereas only 37% of the *H. avenae*

Differential Expression				BLASTX Annotation				SP ¹	NS ²	
Transcript ID	Protein ID	logFC p2/p2 C	p-value	logFC p2/p2 C	p-value	ID	e-value	Define	SP ¹	NS ²
c25351_r1.10	c25351_r1.i0.m.26334	Not D/E	N/A	7.513959037	7.05E-06	[053747297] [0A1C81913.1]		Q:Cairectin, partial [Heterodera avenae]	Y	N
c25351_r1.18	c25351_r1.i8.m.26332	3.077137185	2.82E-05	Not D/E	N/A	[A0A0688A57] [A0A0688A57_HETAV]	1.00E-94	C:Cairectin [Fragment] OS=Heterodera avenae PE=2 SV=1	Y	N
c26095_r1.1	c26095_r1.i1.m.33139	6.902359072	2.42E-05	Not D/E	N/A	[F1KWN0] [F1KWN0_ASCSU]	2.00E-07	G:Glutathione synthetase OS=Ascaris suum GN=ASU_12312 PE=2 SV=1	Y	N
c26095_r2.1	c26095_r2.i1.m.33143	9.303958616	3.43E-11	8.693956688	2.26E-09	[A8NM47] [A8NM47_BRUMA]	6.00E-30	G:Glutathione synthetase OS=Buglia malyi GN=Bm-ess-1 PE=3 SV=2	Y	N
c26095_r2.2	c26095_r2.i2.m.33146	7.020320783	6.20E-15	5.521469013	1.90E-10	[F1KWN0] [F1KWN0_ASCSU]	9.00E-33	G:Glutathione synthetase OS=Ascaris suum GN=ASU_12312 PE=2 SV=1	Y	N
c26095_r2.3	c26095_r2.i3.m.33148	7.108119971	2.00E-16	6.33705449	4.68E-14	[F1KWN0] [F1KWN0_ASCSU]	5.00E-6	G:Glutathione synthetase OS=Ascaris suum GN=ASU_12312 PE=2 SV=1	Y	N
c26889_r1.1	c26889_r1.i2.m.41895	3.708271333	2.04E-07	3.306467493	8.82E-06	[F1KWN0] [F1KWN0_ASCSU]	2.00E-58	G:Glutathione synthetase OS=Ascaris suum GN=ASU_12312 PE=2 SV=1	Y	Y
c26349_r1.1	c26349_r1.i2.m.35883	7.018462968	1.75E-16	6.295558182	2.46E-14	[F1KWN0] [F1KWN0_ASCSU]	4.00E-98	G:Glutathione synthetase OS=Ascaris suum GN=ASU_12312 PE=2 SV=1	Y	Y
c26349_r1.3	c26349_r1.i3.m.35884	7.575901268	1.03E-06	7.057484298	1.80E-05	[F1KWN0] [F1KWN0_ASCSU]	1.00E-39	G:Glutathione synthetase OS=Ascaris suum GN=ASU_12312 PE=2 SV=1	Y	N
c27439_r4.1	c27439_r4.i1.m.49354	Not D/E	N/A	7.556795591	1.45E-06	[I6R429] [I6R429_ROTRE]	4.00E-08	C-type lectin domain protein OS=Heterodera reniformis GN=CTL8 PE=2 SV=1	Y	N
c27491_r6.1	c27491_r6.i1.m.50307	9.02016562	1.25E-22	6.643524047	1.55E-15	[Q8T506] [Q8T506_HETGL]	4.00E-08	C-type lectin domain protein OS=Heterodera glycomes GN=ctl-1 PE=2 SV=1	Y	N
c27491_r6.2	c27491_r6.i2.m.50309	7.706506999	9.16E-13	5.278968598	5.19E-06	[Q8T506] [Q8T506_HETGL]	4.00E-08	C-type lectin domain protein OS=Heterodera glycomes GN=ctl-1 PE=2 SV=1	Y	N
c27491_r6.3	c27491_r6.i3.m.50311	12.23963859	1.02E-19	9.68257352	4.12E-12	[Q8T506] [Q8T506_HETGL]	4.00E-08	C-type lectin domain protein OS=Heterodera glycomes GN=ctl-1 PE=2 SV=1	Y	N
c27491_r6.4	c27491_r6.i4.m.50313	9.173867646	3.18E-19	7.494721024	8.75E-14	[Q8MU03] [Q8MU03_HETGL]	6.00E-15	P:Putative C-type lectin [Fragment] OS=Heterodera glycomes PE=2 SV=1	Y	N
c27491_r6.7	c27491_r6.i7.m.50320	7.125879925	7.51E-16	5.299244923	1.73E-10	[Q8T506] [Q8T506_HETGL]	7.00E-77	A:Arabian endo-1.5.alpha-L-arabinosidase OS=Streptomyces fradiae GN=SFRA_02705 PE=4 SV=1	Y	Y
c26005_r1.1	c26005_r1.i1.m.32160	4.591061117	2.85E-09	4.202830621	3.46E-08	[A0A081EQM5] [A0A081EQM5_STFRF]	7.00E-77	A:Arabian endo-1.5.alpha-L-arabinosidase OS=Streptomyces fradiae GN=SFRA_02705 PE=4 SV=1	Y	Y
c25123_r1.1	c25123_r1.i2.m.24638	7.583787987	2.06E-12	8.252331567	6.65E-14	[A0A060F48] [A0A060F48_RHIET]	3.00E-09	P:Putative glycoside hydrolase family 32 protein OS=Rhizobium cell bv. mimosae str. I47771 GN=I47771_P000040	Y	N
c16214_r1.1	c16214_r1.i1.m.3704	6.940327862	5.59E-16	5.600953207	4.46E-12	[WZDZL2] [WZDZL2_9P5ED]	7.00E-16	G:Glycosyl hydrolase family protein OS=Pseudomonas sp. FHI GN=H096_03325 PE=4 SV=1	Y	N
c26100_r1.1	c26100_r1.i2.m.33205	3.150568712	9.20E-06	Not D/E	N/A	[I3HP94] [I3HP94_RHRXZ]	1.00E-144	B:Beta-fructofuranosidase OS=Rhizobium grahamii CGE 502 GN=RCCGE502_24898 PE=4 SV=1	Y	N
c23521_r1.1	c23521_r1.i1.m.15080	5.513826106	4.81E-08	4.64980961	1.59E-05	[G0WX17] [G0WX17_HETGL]	5.00E-41	E:Expansion OS=Heterodera glycomes GN=expb1.1 PE=2 SV=1	Y	N
c23521_r1.2	c23521_r1.i2.m.15081	7.991882961	1.84E-18	7.200136917	3.26E-16	[G0WX17] [G0WX17_HETGL]	3.00E-45	E:Expansion OS=Heterodera glycomes GN=expb1.1 PE=2 SV=1	Y	N
c28740_r8.5	c28740_r8.i5.m.69168	Not D/E	N/A	7.638832011	1.03E-06	[I380447937] [0A1FD54151.1]	6.00E-32	P:Pan-binding protein 4 [Heterodera glycomes]	Y	N
c24113_r1.3	c24113_r1.i3.m.17940	3.846768968	1.17E-06	3.386697069	1.89E-05	[I7ANYW0] [I7ANYW0_GLORO]	6.00E-20	S:Secreted SPRY domain-containing protein.18 OS=Globobera rostochiensis GN=SPRYSEC-18 PE=2 SV=1	Y	N
c29474_r1.3	c29474_r1.i3.m.81483	6.116874597	3.76E-13	5.294803822	8.26E-11	[I7AL59] [I7AL59_GLORO]	5.00E-09	S:Secreted SPRY domain-containing protein.10 OS=Globobera rostochiensis GN=SPRYSEC-10 PE=2 SV=1	Y	N
c29474_r1.2	c29474_r1.i2.m.81482	5.200527433	3.88E-11	4.076150523	4.91E-08	[C1KBN7] [C1KBN7_GLOPA]	1.00E-28	R:RBP-4 protein OS=Globobera pallida GN=rpb-4 PE=2 SV=1	Y	N
c29474_r1.4	c29474_r1.i4.m.81484	3.792815666	3.43E-07	3.411680378	4.14E-06	[C1KBN7] [C1KBN7_GLOPA]	1.00E-28	R:RBP-4 protein OS=Globobera pallida GN=rpb-4 PE=2 SV=1	Y	N
c31053_r11.1	c31053_r11.i1.m.108435	3.72462795	2.69E-05	Not D/E	N/A	[Q5K405] [Q5K405_GLORO]	4.00E-11	E:Expansion protein OS=Globobera rostochiensis GN=exp PE=2 SV=1	N	N
c28296_r1.4	c28296_r1.i4.m.62404	5.242592723	6.66E-11	5.01272283	3.23E-10	[Q5K405] [Q5K405_GLORO]	7.00E-06	E:Expansion protein OS=Globobera rostochiensis GN=exp PE=2 SV=1	Y	N
c28989_r2.3	c28989_r2.i3.m.72038	3.291385821	6.02E-06	Not D/E	N/A	[Q7YW95] [Q7YW95_MEUC]	2.00E-08	P:Putative esophageal gland cell secretory protein.2 OS=Meloidogyne incognita GN=msp2 PE=2 SV=1	N	Y
c25519_r1.5	c25519_r1.i5.m.27699	5.80724819	1.08E-12	4.800801191	7.71E-10	[O9S1C3] [O9S1C3_1ATL41_ARATH]	4.00E-11	E3:ubiquitin-protein ligase AT1L41 OS=Arabidopsis thaliana GN=AT1L41 PE=1 SV=1	Y	N
c25519_r1.8	c25519_r1.i8.m.27702	3.46731683	1.18E-05	Not D/E	N/A	[O9S1C3] [O9S1C3_1ATL41_ARATH]	3.00E-11	E3:ubiquitin-protein ligase AT1L41 OS=Arabidopsis thaliana GN=AT1L41 PE=1 SV=1	Y	N
c22414_r1.1	c22414_r1.i1.m.10953	3.818429607	3.89E-07	3.025351434	4.03E-05	[L5080821] [L5080821.AC007258_19]	9.00E-10	H:Hydrophobic protein [Arabidopsis thaliana]	Y	Y
c30336_r2.12	c30336_r2.i12.m.96934	4.966248846	5.12E-10	4.43622147	1.40E-08	[A0A078D2H0] [A0A078D2H0_BRANA]	4.00E-09	B:Arabidopsis thaliana GN=BraA09a24390D PE=4 SV=1	Y	N
c27986_r1.5	c27986_r1.i5.m.52433	9.056459001	1.69E-10	7.513959037	7.05E-06	[I278SK7] [I278SK7_PHAVU]	1.00E-09	U:Uncultured protein [Fragment] OS=Phaeoelus vulgarius GN=PHAVU_0056028100g PE=4 SV=1	Y	N
c23341_r2.4	c23341_r2.i4.m.14226	10.19589858	1.05E-13	9.4970701043	1.39E-11	[O1LWV2] [O1LWV2_ARATH]	7.00E-08	A:Arabidopsis thaliana PE=2 SV=1	Y	N

Table 3.4. Predicted effector candidates of interest from the novel effector prediction pipeline. ³log FC ppJ2-pJ2 C = log fold-change from the ppJ2 to the pJ2 compatible. ⁴log FC ppJ2-pJ2 IC = log fold-change from the ppJ2 to the pJ2 stage on the incompatible. ⁵ID = database entry. ⁶N-terminal signal peptide. ⁷Nuclear localization signal

transcripts had homology to previously annotated genes (Kumar et al., 2014). Many of these potential homologues existed in other nematode species, including plant- and animal-parasitic nematodes. In addition, some transcripts showed homology to a bacterial endosymbiont, from the genus *Cardinium*, of *Encarsia pergandiella*, a parasitic wasp. Previous work identified this endosymbiont and characterized it as *Candidatus Paenicardinium endonii*, later renamed to *Candidatus Cardinium hertigii* (Nakamura et al., 2009; Noel and Atibalentja, 2006). However, little is known about the function of this endosymbiont and what role it may play, if any, in plant parasitism. Related endosymbionts found in insects and arachnids have been shown to have prominent impacts on their hosts, leading to changes in host reproductive capacity and also modulating host immunity (Eleftherianos et al., 2013). To better understand the function of a putative endosymbiont in *H. glycines* all transcripts associated with this endosymbiont were identified and extracted from the transcriptome, representing a majority of the characterized sequences for this endosymbiont. Those transcripts identified were primarily associated with metabolic processes, which may contribute to both nematode and endosymbiont metabolism. Further studies into the function of this endosymbiont and the effect on parasitism removing the endosymbiont has will be vital in elucidating what role it plays inside the nematode. In addition to a bacterial endosymbiont, several putative homologs from viruses were also identified. Previously, researchers found representative viruses from the *Bornaviridae*, *Rhabdoviridae*, and *Bunyaviridae* families contained within *H. glycines* (Bekal et al., 2011; Ruark et al., 2017). Thus, there appears to be a significant microbial community active within *H. glycines* that has until now remained largely unexplored. Further examination of these

organisms could reveal vital connections that can be exploited for improving resistance against the soybean cyst nematode.

Stylet-secreted effectors represent a key component of the plant-nematode interaction, serving a wide variety of functions required for successful invasion and establishment of the nematode feeding site. Previous studies have identified a suite of these effectors using microaspiration techniques to isolate the contents of the esophageal gland cells where these genes are expressed (Gao et al., 2003; Noon et al., 2015). These studies then prioritized potential effectors based on those sequences possessing a signal peptide and lacking a transmembrane domain. Despite previous knowledge about the effector repertoire of SCN, very little is known about the structure of these sequences within a population, specifically how these sequences vary from one individual to another. To address this question we undertook an effector variation analysis within the transcriptome, identifying putative variants of known effectors and examining their level of variation within the population. Within the SCN transcriptome, predicted sequence variants of known effectors ranged from over 70 (effector 4F01) down to one (effectors 7E05 and GLAND2). This effector variation may be a result of different alleles being present in the population and/or reflect variation in the copy numbers of genes encoding these effectors. The level of variation of these effectors is likely related to the function of the effector in question. For example, a highly variable effector such as 4F01 may be under constant selection pressure to avoid host recognition, resulting in a wide level of variation across the gene pool. A prior study demonstrated that 4F01 might function as a mimic of host plant annexins to promote successful plant infection (Patel et al., 2010). By contrast, effectors with limited variation across the population are likely constrained by

their function. It would be interesting to see how a highly virulent population or a field population compares to the highly inbred population used here for sequencing. Certain effectors may be expanded or reduced depending on the population and host selection pressure. Effectors with a very low number of variants across populations may represent key elements of infection that could be targeted for further study in the attempt at identifying a novel source of broad spectrum SCN resistance.

We also mined the early parasitic transcriptome to identify additional effectors expressed within *H. glycines* using the SignalP predictive tool, as well as a novel pipeline called N-preffector. The use of N-preffector allowed for the identification of an entirely new class of effectors not necessarily containing a signal peptide. Examples of secreted effectors lacking a signal peptide have been shown in other plant-parasitic nematode species such as *Globodera rostochiensis*, where they have been shown to play a role in disrupting host reactive oxygen species production (Robertson et al., 1999; Robertson et al., 2000). These effectors may contain a previously unknown secretion signal or utilize a novel secretion pathway in order to be secreted. Between the two pathways utilized for effector discovery, 86% of known SCN effectors were re-discovered within the early parasitic SCN transcriptome. The remaining 14% were not re-discovered either due to truncated sequences relative to the reference sequence or a change in the predicted protein sequence between the transcriptome and reference sequence. It is interesting to note that 47% of these effectors were identified by both pipelines, illustrating a need for both pipelines to accurately detect all possible effectors. The signal peptide dependent method is excellent at predicting putative effectors, but misses out on transcripts that may be truncated or simply lack the signal peptide, which can be complemented using the N-

preffector pipeline. Effector candidates predicted here by both SignalP- and N-preffector-based pipelines provide new candidates for downstream expression and functional analysis.

The novel effector candidates identified by these two pipelines represent a set of genes for investigating the interaction between SCN and soybean. Many of these sequences have little or no annotation available, much like the original gland isolated effector sequences obtained for *H. glycines* (Bekal et al., 2015; Gao et al., 2003; Noon et al., 2015). These novel effector sequences may play pivotal roles in nematode parasitism and will require more in depth functional studies to determine their function. Among those sequences with available annotation are many homologs of effectors from other plant-parasitic nematodes that were not identified or characterized in *H. glycines* previously. Included in this category are genes such as the glutathione synthetase family, the novel *G. rostochiensis* effector E9, and candidates showing homology to the SPRYSEC family of effectors from *Globodera rostochiensis*. Glutathione synthetases have many potential roles in the interaction between the nematode and host plant. In the interaction of the root-knot nematode *Meloidogyne incognita* it was found that glutathione is needed for successful infection of the host plant *Medicago truncatula* (Baldacci-Cresp et al., 2012). In addition, glutathione synthetase genes were found to be greatly expanded in the genome of the potato cyst nematode *Globodera pallida*, where these genes are theorized to be involved in protection of the nematode from antioxidant proteins as well as potentially in nematode nutrition (Cotton et al., 2014). Several transcripts annotated as glutathione synthetase also contained a secretion signal, something that differentiates them from glutathione synthetases found in animal parasites

that may function within the nematode. The putative effector E9 has been identified in both *Globodera rostochiensis* and *Globodera pallida* and was confirmed to be expressed in gland cells via *in situ* hybridization (Ali et al., 2015; Thorpe et al., 2014). Thus far little is known about the function of the E9 effector, other than it being expressed in the gland cells of *Globodera* species. The SPRYSEC effectors on the other hand have been heavily investigated in the *Globodera*-tomato pathosystem, with demonstrated roles in the suppression of plant immune responses (Ali et al., 2015; Rehman et al., 2009). Currently, however, the SPRYSEC family has been thought to be exclusive to the potato cyst nematodes of the genus *Globodera*. Thus, these could be very interesting candidates to pursue and contrast in virulent populations of SCN to determine whether or not they play the same role as in *Globodera*. Another effector candidate of note is a putative secreted calreticulin. A calreticulin secreted by *Meloidogyne incognita* has been previously shown to be necessary for successful infection and may play a role in suppression of plant defenses, functions that may be retained in *H. glycines* (Jaouannet et al., 2013). Another nematode effector homolog group identified in the transcriptome involved in suppression of host defenses are the C-type lectins (CTLs) from *Rotylenchus reniformis*. These effectors were identified in the *R. reniformis* transcriptome and subsequently shown to be expressed in the hypodermis of parasitic stages of the nematode (Ganji et al., 2014). It is hypothesized that these effectors are involved in protecting the nematode from environmental stress. While these homologs are all predicted to have the same function in *H. glycines* as their originating species, further functional characterization is necessary to confirm this.

Interestingly, we identified several effector candidates with sequence similarity to proteins originating in plants and other organisms. These included multiple effector candidates with homology to members of the plant RING/U-box superfamily of proteins. These proteins are typically involved in protein modification and regulation of plant pathways, including defense responses and regulation of cell death (Guzman, 2012). Nematode mimics of these proteins may be involved in manipulation or suppression of host defense pathways in order to allow successful establishment of the feeding site. Among the identified effector candidates are also several homologs related to plant metabolism and cell wall degradation. These included arabinosidase, fructosidase, glycoside hydrolase, and expansin. These cell wall modifying proteins have been shown to aid in the loosening and degradation of polysaccharides present in the plant cell wall (Cosgrove, 2000; Minic and Jouanin, 2006; Weinstein and Albersheim, 1979) and have been identified from other plant-parasitic nematodes where they play a crucial function in migration and establishment of the nematode feeding site (Qin et al., 2004; Wieczorek et al., 2006). Therefore, these plant mimics all represent avenues of study to be pursued in order to better understand the interplay between SCN and its plant host, soybean.

In conclusion, a *de novo* transcriptome of the pre-parasitic and parasitic second-stage juvenile life stages of *H. glycines* has been generated and annotated. Within this transcriptome novel effector candidates were identified utilizing a new prediction tool not reliant on sequences possessing a signal peptide, N-preffector. In addition, the level of variation of previously identified *H. glycines* effectors was examined for the first time at the population level and identified highly conserved and highly variable effectors. Finally, this transcriptome provides a useful genetic resource that will aid in annotation of

the SCN genome. Combining these data will provide insights into the biology of SCN with the hopes of discovering new ways to combat this pathogen.

MATERIALS & METHODS

Nematode cultivation and isolation

The SCN inbred population PA3 (HG Type 0) was propagated under greenhouse conditions on soybean cultivar Williams 82. Freshly hatched pre-parasitic second-stage juveniles (ppJ2) were inoculated onto 10-day old seedlings of the susceptible host (cv. Williams 82) or the resistant host (cv. Forrest) and the inoculated plants were placed in the greenhouse. The remaining ppJ2 nematodes were pelleted by centrifugation and flash-frozen in liquid nitrogen and stored at -80°C prior to RNA isolation. Five days post-inoculation, parasitic second-stage juveniles (pJ2) nematodes were isolated from the roots by blending the roots for 30s in a kitchen blender. Following this, the root homogenate was poured over a nested stack of sieves with pore sizes of 850µm, 250µm, and 25µm before purifying the nematodes from the sample using sucrose centrifugal flotation (de Boer et al., 1999). Samples were frozen in liquid nitrogen and stored at -80°C prior to RNA isolation.

RNA Isolation and Sequencing

RNA was isolated from frozen nematode pellets using the PerfectPure Fibrous Tissue Kit (5Prime) and a modified version of the manufacturer's extraction protocol. Tissue was homogenized in 30 second intervals in the provided lysis solution containing

0.5 μ M TCEP using a bead beater and 1.0 mm zirconia beads, followed by a 30 second incubation on ice. This was repeated three times. The sample was centrifuged briefly at room temperature before transferring the supernatant to a fresh tube. Following lysis and homogenization, 10 μ l of the provided Proteinase K was added and the sample was allowed to incubate on ice for 10 minutes, after which the manufacturer's protocol for RNA purification was followed. RNA quality was determined using a Fragment Analyzer (Advanced Analytical) and quantified using a Qubit Fluorometer prior to library preparation. RNA-seq libraries (ppJ2, pJ2 infecting susceptible host, pJ2 infecting resistant host) were constructed using the TruSeq mRNA Stranded Library Prep Kit (Illumina) and sequenced on the Illumina HiSeq 2500 platform in a paired-end manner (2x100 for ppJ2 and pJ2-Compatible samples and 2x50 for pJ2-Incompatible sample.). Library preparation and high-throughput sequencing services were performed at the University of Missouri DNA Core Facility. Three biological replicates of each sample were sequenced.

De novo transcriptome assembly

Prior to assembly, raw reads from these libraries were filtered using Trimmomatic (Bolger et al., 2014) to remove low quality reads. The remaining reads were paired and orphan reads discarded. High quality paired-end reads were used as input for transcriptome assembly. *De novo* transcriptome assembly was completed using the de Bruijn graph-based tool Trinity (Grabherr et al., 2011). As part of the assembly process, an *in silico* read normalization step was performed. Assembly quality was then assessed

by mapping raw reads back to transcripts using Bowtie2 (Langmead and Salzberg, 2012) at default parameters.

Transcriptome annotation and quantification

The transcriptome was annotated following the established Trinotate pipeline (Grabherr et al., 2011). Homology searches were performed against the protein sequences contained in Genbank (Benson et al., 2005) and UniProt (Magrane et al., 2011) databases using BLASTX at an evalue cutoff of $1e-5$ (Altschul et al., 1997). Transcripts were translated into protein using TransDecoder, a component of Trinity (Grabherr et al., 2011). HMMER and Pfam databases were used to predict protein domains contained within each transcript (Finn et al., 2011; Punta et al., 2012). Presence of a signal peptide was determined using SignalP version 4.0 and TMHMM version 2.0 was utilized to identify predicted transmembrane domains (Petersen et al., 2011; Sonnhammer et al., 1998). The resulting annotation information was then combined and pooled into a SQLite database. In addition, sequenced nematode genomes were leveraged to identify potential homologs within the transcriptome. For this, predicted protein datasets from the genomes of *Bursaphelenchus xylophilus*, *Meloidogyne hapla*, *Meloidogyne incognita*, *Globodera pallida*, *Pristionchus pacificus*, *Ascaris suum*, and *Caenorhabditis elegans* were downloaded from WormBase (<http://ws204.wormbase.org/>) and used. BLASTP hits from the *H. glycines* transcriptome with e-values less than $1e-5$ were considered potential homologs. Lists of potential homologs from each of the seven species examined were then compared and contrasted to determine uniquely shared homologs between the sequenced nematode and *H. glycines*.

For quantification and differential expression analysis, reads from the libraries used for assembly were mapped and quantified using RSEM (Li and Dewy, 2011) to determine transcript abundance. RSEM was utilized as it has been shown to correlate well with RT-qPCR measurements and produce expression values with high accuracy (Chandramohan et al., 2016). Following quantification, differential expression analysis was conducted using edgeR (Robinson et al., 2010), identifying all genes with a minimum 4-fold expression difference and under a p-value cutoff of 0.001 between any of the samples.

Identification of endosymbiont sequences within the *H. glycines* transcriptome

The entire *de novo* early parasitic transcriptome for *H. glycines* was mined for transcripts related to the endosymbiont “*Candidatus Cardinium hertigii*”. All transcripts annotated with the species designation ‘Cardinium endosymbiont’ were extracted from the transcriptome and combined into a file. A database was then constructed from the complete proteome of the closest available sequenced bacterial isolate, *Cardinium hertigii* cEper1 isolated from *Encarsia pergandiella* (Penz et al., 2012). Then all putative *Cardinium*-associated sequences were mapped against the proteome database using BLASTX at an e-value cutoff of 1e-5 to confirm their identity as putative endosymbiont-associated transcripts. The resulting transcripts were then used for gene ontology analysis.

Gene ontology analysis of endosymbiont-associated sequences from the *H. glycines* transcriptome

Gene ontology (GO) analysis was performed to identify the putative function of endosymbiont-associated sequences within SCN. To do this endosymbiont-associated sequences from the SCN transcriptome were used in the research tool BLAST2GO (Conesa et al., 2005). This tool uses a similarity searches to assign GO annotation to sequence data lacking well-characterized GO annotation. In BLAST2GO BLASTX was performed at an e-value cutoff of $1e-5$ and the top available BLAST hit used to pull available GO annotation. Once available GO annotation was assigned to the 468 endosymbiont-associated transcripts the results were examined for their potential role in SCN biology.

Variation of known SCN effectors

The protein sequences for the 72 known SCN effector sequences (Bekal et al., 2015; Gao et al., 2003; Noon et al., 2015) were aligned using MUSCLE (Edgar, 2004) and then a maximum likelihood tree was constructed based on sequence homology in MEGA7 (Kumar et al., 2016). MUSCLE (multiple sequences comparison by log-expectation) is a high accuracy tool for protein alignment. Effectors with bootstrap values greater than 50 were grouped into stylet-secreted effector families (SSEFs). Predicted transcript peptide sequences from the SCN transcriptome were then mapped to known SCN effector protein sequences using BLASTP at an e-value cutoff of $1e-5$ and quantified for each known effector. Variants of known SCN effectors in a SSEF were pooled for quantification.

Effector prediction

The effector prediction pipeline started with all predicted peptides from the SCN transcriptome. First, sequences represented in the gland cell transcriptome were subjected to two different prediction tools: SignalP (Petersen et al., 2011) and N-Preffector, developed in this study. For the SignalP-based prediction, peptides were run through SignalP 4.0 and TMHMM (Sonnhammer et al., 1998) to predict signal peptides and transmembrane helices, respectively. Predicted peptides containing a signal peptide and lacking a transmembrane domain were then filtered based on their expression between the ppJ2 and pJ2 life stages of the nematode, with those peptides showing a minimum 4-fold up-regulation into the pJ2 lifestage retained. Finally, nuclear localization signals were predicted using NLStradamus (Nguyen et al., 2009). For N-Preffector based prediction, predicted peptides were run through a machine-learning algorithm trained on 72 known *H. glycines* effector sequences and 150 known non-effector sequences from *H. glycines* in addition to the original sequences (gram negative bacteria) in which the Prefector model was trained (Dhroso et al., in preparation). For each protein sequence, N-Preffector calculates a vector of length-invariant features; the feature vector is then used as an input for the classification model. Feature categories that were considered are: residue composition, sequence/structure information, and physico-chemical properties of proteins. To select highly correlated features with the class and not correlated with each other, Prefector utilizes the correlation-based feature selection (CFS) method (Hall, 1999). Our goal was to minimize the number of proteins erroneously misclassified as effectors, *i.e.* false positives, while trying to maximize the number of predicted real effectors, using the same exact protocol utilized in Prefector. N-Preffector achieves this through a more stringent classification criterion. Given an

SVM model M and a training data of size n , for each training example x_k let $f_k \in [+1, -1]$ be its decision value predicted by the SVM model, and $y_k \in \{+1, -1\}$ to be its true annotation of being an effector or non-effector. Given the SVM model M , the prediction probability for a training example x_k is defined as $p_k^{(i)} = \frac{1}{(1 + \exp(A^{(i)} f_k^{(i)} + B^{(i)}))}$. The coefficients $A^{(i)}$ and $B^{(i)}$ are estimated during the SVM training process by minimizing the log-likelihood function. Those peptides predicted by N-preffector at or above a 0.9 confidence score cutoff were then filtered based on expression, retaining peptides with a minimum 4-fold up-regulation from the ppJ2 to the pJ2 lifestage. Nuclear localization signals were then predicted using NLStradamus for the remaining peptides (Nguyen et al., 2009).

ACKNOWLEDGEMENTS

We would like to acknowledge William Spollen and Christopher Bottoms of the University of Missouri Informatics Research Core Facility for their expertise with the bioinformatics analysis. We thank summer NSF-REU interns Ying Lin, Eric Penton, Harrison Hennessey for assistance with annotation and Robert Heinz for maintenance of nematode cultures. This work was supported by USDA-NIFA-AFRI Awards 2012-67013-19345 (to TJB, MGM, and ELD) and 2015-67013-23511 (to TJB, MGM, and DK). MNG was supported in part by a MU Life Sciences Graduate Research Fellowship.

REFERENCES

- Abad, P., Guozy, J., Aury, J., Castagnone-Sereno, P., Danchin, E., Deleury, E., et al. 2008. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nature Biotechnology* 26: 909-915
- Ali, S., Magne, M., Chen, S., Obradovic, N., Jamshaid, L., Wang, X., Bélair, G., and Moffett, P. 2015. Analysis of *Globodera rostochiensis* effectors reveals conserved functions of SPRYSEC proteins in suppressing and eliciting plant immune responses. *Frontiers in Plant Science* 6: 623. doi: 10.3389/fpls.2015.00623.
- Baldacci-Cresp, F., Chang, C., Maucourt, M., Deborde, C., Hopkins, J., Lecomte, P., et al. 2012. (Homo)glutathione deficiency impairs root-knot nematode development in *Medicago truncatula*. *PLoS Pathogens* 8: e1002471.
- Bekal, S., Domier, L.L., Niblack, T.L., and Lambert, K.N. 2011. Discovery and initial analysis of novel viral genomes in the soybean cyst nematode. *Journal of General Virology* 92: 1870-1879.
- Bekal, S., Domier, L.L., Gonfa, B., McCoppin, N.K., Lambert, K.N., and Bhalerao, K. 2014. A novel flavivirus in the soybean cyst nematode. *Journal of General Virology* 95: 1272-1280.
- Bekal, S., Domier, L.L., Gonfa, B., Lakhssassi, N., Meksem, K., Lambert, K.N. 2015. A SNARE-like protein and biotin are implicated in soybean cyst nematode virulence. *PLoS ONE*: e0145601.
- Benson, D.A., Karsh-Mizrachi, I., Lipman, D.J., Ostell, J., and Wheeler, D.L. 2005. GenBank. *Nucleic Acids Research* 1: D34-8.
- Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, btu170.
- Chandramohan, R., Wu, P., Phan, J., and Wang, M. 2016. Benchmarking RNA-Seq quantification tools. *Conference Proceedings of the IEEE Engineering in Medicine and Biology Society*. doi: 10.1109/EMBC.2013.6609583
- Cotton, J.A. et al. 2014. The genome and life-stage specific transcriptomes of *Globodera pallida* elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biology* 15: R43.
- Cosgrove, D. 2000. Loosening of plant cell walls by expansins. *Nature* 407: 321-326.
- De Boer, J.M., Yan, Y., Wang, X., Smant, G., Hussey, R.S., Davis, E.L., and Baum, T.J. 1999. Developmental expression of secretory β -1,4-endoglucanases in the subventral

esophageal glands of *Heterodera glycines*. *Molecular Plant Microbe Interactions* 12: 663-669.

Dhroso, A., Warren, S., and Korkin, D. 2017. Genome-wide prediction of bacterial effectors across six secretion system types using a feature-based statistical framework. *PLoS Computational Biology*, in revision.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792-1797.

Eleftherianos, I., Atri, J., Accetta, J., and Castillo, J.C. 2013. Endosymbiotic bacteria in insects: guardians of the immune system? *Frontiers in Physiology* 4: 46.

Eves-van den Akker, S., Laetsch, D., Thorpe, P., Lilley, C., Danchin, E., Da Rocha, M., et al. 2016. The genome of the yellow potato cyst nematode, *Globodera rostochiensis*, reveals insights into the basis of parasitism and virulence. *Genome Biology* 17:124.

Finn, R.D., Clements, J., and Eddy, S.R. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research* 39: W29-W37.

Fosu-Nyarko, J., Nicol, P., Naz, F., Gill, R., and Jones, M.G.K. 2016. Analysis of the transcriptome of the infective stage of the beet cyst nematode, *H. schachtii*. *PLOS ONE* 11: e0147511. DOI:10.1371/journal.pone.0147511

Ganji, S., Jenkins, J., Wubben, M. 2014. Molecular characterization of the reniform nematode C-type lectin gene family reveals a likely role in mitigating environmental stresses during plant parasitism. *Gene* 537: 268-78.

Gardner, M., Verma, A., and Mitchum, M.G. 2015. Emerging roles of cyst nematode effectors in exploiting plant cellular processes. In: Escobar, C. and Fenoll, C. (eds.), *Advances in Botanical Research Vol 73, Plant Nematode Interactions: A View on Compatible Interrelationships*. Elsevier Publishing, Oxford p. 259–291.

Gao, B., Allen, R., Maier, T., Davis, E.R., Baum, T.J., Hussey, R.S. 2003. The parasitome of the phytonematode *Heterodera glycines*. *Molecular Plant Microbe Interactions* 2003: 720-726.

Gheysen, G., and Mitchum M.G. 2011. How nematodes manipulate plant development pathways for infection. *Current Opinion in Plant Biology* 14: 415-421.

Grabherr, M.G. et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644-652.

Gúzman, P. 2012. The prolific ATL family of RING-H2 ubiquitin ligases. *Plant Signaling & Behavior* 7: 1014-1021.

- Haas, B.J. Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P., Bowden, et al. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8: 1494-1512.
- Hageman, A., Bauters, L., Kyndt, T., Rahman, M.M., and Gheysen, G. 2013. Identification of candidate effector genes in the transcriptome of the rice root knot nematode *Meloidogyne graminicola*. *Molecular Plant Pathology* 14: 379-390.
- Hall, M.A. 1999. Correlation-based feature selection for machine learning. Doctoral Dissertation, The University of Waikato.
- Jaouannet, M., Magliano, M., Arguel, M.J., Gourgues, M., Evangelisti, E., Abad, P., and Rosso, M.N. 2013. The root-knot nematode calreticulin Mi-CRT is a key effector in plant defense suppression. *Molecular Plant-Microbe Interactions* 26: 97-105.
- Jones, J.D.G., and Dangl, J.L. 2006. The plant immune system. *Nature* 444: 323-329.
- Kikuchi, T. et al. 2011. Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathogens* 7: e1002219.
- Koenning, S.R., and Wrather, J.A. 2010. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. *Plant Health Progress* doi:10.1094/PHP-2010-1122-01-RS.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* 305: 567-580.
- Kumar, M., Gantasala, N.P., Roychowdhury, T., Thakur, P.K., Banakar, P., Shukla, R.N., et al. 2014. *De novo* transcriptome sequencing and analysis of the cereal cyst nematode, *Heterodera avenae*. *PLoS ONE* 9: e96311.
- Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*: msw054.
- Langmead, B., and Salzberg, S. 2012. Fast gapped-read alignment with Bowtie2. *Nature Methods* 9: 357-359.
- Magrane, M. et al. 2011. UniProt knowledgebase: a hub of integrated protein data. *Database(Oxford)*: bar009.
- Minic, Z., Jouanin, L. 2006. Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiology and Biochemistry* 7: 435-449.

- Mitchum, M.G., Hussey, R.S., Davis, E.L., Baum, T.J., Wang, X., Elling, A.A., et al. 2013. Nematode effector proteins: an emerging paradigm of parasitism. *New Phytologist* 199: 879–894.
- Na, R., and Gijzen, M. 2016. Escaping host immunity: new tricks for plant pathogens. *PLoS Pathogens* 12: e1005631.
- Nakamura, Y., Kawai, S., Yukuhiro, F., Ito, S., Gotoh, T., Kisimoto, R., et al. 2009. Prevalence of *Cardinium* bacteria in planthoppers and spider mites and taxonomic revision of “*Candidatus Cardinium hertigii*” based on detection of a new *Cardinium* group from biting midges. *Applied and Environmental Microbiology* 75: 6757-6763.
- Navarro, L., Zipfel, C., Rowland, O., Keller, I., Robatzek, S., Boller, T., et al. 2004. The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiology* 135: 1113-1128.
- Nguyen Ba, A.N., Pogoutse, A., Provar, N., and Moses, A.M. 2009. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* 10:202.
- Noel, G.R., and Atibalentja, N. 2006. ‘*Candidatus Paenicardinium endonii*’, an endosymbiont of the plant-parasitic nematode *Heterodera glycines* (Nemata: Tylenchida), affiliated to the phylum *Bacteroidetes*. *International Journal of Systematic and Evolutionary Microbiology* 56: 1697-1702.
- Noon, J.B., Hewezi, T., Maier, T.R., Simmons, C., Wei, J.Z., Wu, G., et al. 2015. Eighteen new candidate effectors of the phytonematode *Heterodera glycines* produced specifically in the secretory esophageal gland cells during parasitism. *Phytopathology* 105: 1362-72.
- Patel, N., Hamamouch, N., Li, C., Hewezi, T., Hussey, R.S., Baum, T.J., et al. 2010. A nematode effector protein similar to annexins in host plants. *Journal of Experimental Botany* 61:235-248.
- Penz, T., Schmitz-Esser, S., Kelly, S.E., Cass, B.N., Müller, A., Woyke, T., et al. 2012. Comparative genomics suggests an independent origin of cytoplasmic incompatibility in *Cardinium hertigii*. *PLoS Genetics* 8: e1003012. doi: 10.1371/journal.pgen.1003012.
- Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8: 785-786.
- Petitot, A., Dereeper, A., Agbessi, M., Da Silva, C., Guy, J., Ardisson, M., and Fernandez, D. 2016. Dual RNA-seq reveals *Meloidogyne graminicola* transcriptome and candidate effectors during the interaction with rice plants. *Molecular Plant Pathology* 17: 860-874.

- Punta, M. et al. 2012. The Pfam protein families database. *Nucleic Acids Research* 40: D290-301.
- Qin, L., Kudla, U., Roze, E., Govere, A., Popeijus, H., Nieuwland, J., et al. 2004. Plant degradation: A nematode expansin acting on plants. *Nature* 427: 30.
- Quentin, M., Abad, P., Favery, B. 2013. Plant parasitic nematode effectors target host defense and nuclear functions to establish feeding cells. *Frontiers in Plant Science* 4: 53.
- Rehman, S., Postma, W., Tytgat, T., Prins, P., Qin, L., Overmars, H., et al. 2009. A secreted SPRY domain-containing protein (SPRYSEC) from the plant-parasitic nematode *Globodera rostochiensis* interacts with a CC-NB-LRR protein from a susceptible tomato. *Molecular Plant-Microbe Interactions* 22: 330-40.
- Robertson, L., Robertson, W.M., and Jones, J.T. 1999. Direct analysis of the secretions of the potato cyst nematode *Globodera rostochiensis*. *Parasitology* 119: 167-176.
- Robertson, L., Robertson, W.M., Sobczak, M., Helder, J., Tetaud, E., Ariyanyagam, M.R., et al. 2000. Cloning, expression and functional characterization of a preoxiredoxin from the potato cyst nematode *Globodera rostochiensis*. *Molecular and Biochemical Parasitology* 111: 41-49.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139-40.
- Ruark, C.L., Koenning, S.R., Davis, E.L., Opperman, C.H., Lommel, S.A., Mitchum, M.G., and Sit, T.L. 2017. Soybean cyst nematode culture collections and field populations from North Carolina and Missouri reveal high incidences of infection by viruses. *PLOS ONE* 12: e0171514.
- Simão, F., Waterhouse, R., Ioannidis, P., Kriventseva, E., and Zdobnov, E. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31: 3210-3212.
- Sonnhammer, E., von Heijne, G., and Krogh, A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*, 175-182.
- Thorpe, P., Mantelin, S., Cock, P., Blok, V., Coke, M., Eves-van den Akker, S., et al. 2014. Genomic characterization of the effector complement of the potato cyst nematode *Globodera pallida*. *BMC Genomics* 15:923.

Toruño, T.Y., Stergiopoulos, I., and Coaker, G. 2016. Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. *Annual Review of Phytopathology* 54: 419-441.

Weinstein, L., and Albersheim, P. 1979. Purification and partial characterization of a wall-degrading endo-arabanase and an arabinosidase from *Bacillus subtilis*. *Plant Physiology* 63: 425-432.

Wieczorek, K., Golecki, B., Gerdes, L., Heinen, P., Szakasits, D., Durachko, D., et al. 2006. Expansins are involved in the formation of nematode-induced syncytia in roots of *Arabidopsis thaliana*. *Plant Journal* 48: 98-112.

WormBase web site, <http://www.wormbase.org>, release WS246, date 04-Nov-2014.

Zheng, J., Peng, D., Chen, L., Liu, H., Chen, F., Xu, M., et al. 2016. The *Ditylenchus destructor* genome provides new insights into the evolution of plant parasitic nematodes. *Proceedings of the Royal Society - Biological Sciences* 283: 20160942.

CHAPTER 4

VIRULENT TRANSCRIPTOME ASSEMBLY AND COMPARATIVE ANALYSIS OF KNOWN *HETERODERA GLYCINES* EFFECTORS

ABSTRACT

The soybean cyst nematode (SCN) *Heterodera glycines* represents the most devastating economic pathogen of soybean, able to cause billions of dollars in yield loss each year. This pathogen is normally controlled through the use of resistant soybean cultivars, most deriving their resistance from three sources: PI 88788, PI 548402 (Peking), and PI 437654. Over time the repeated use of these resistance sources has resulted in the generation of virulent populations of SCN able to reproduce on these resistances. To identify what differentiates these virulent populations of SCN from closely related avirulent SCN populations, we sequenced nematodes from an avirulent population and three virulent populations to identify genes and molecular pathways corresponding to virulence. In addition, the expression of known SCN effectors was compared between virulent and avirulent populations and a suite of effectors was found to have similar expression patterns in multiple virulent populations. These effectors represent potential markers for virulence and may also be pursued to identify their molecular function and what role they may play in contributing to SCN virulence.

INTRODUCTION

Soybean cyst nematode (SCN) virulence is an issue that continues to expand throughout the production of soybean in the United States. By constantly growing soybean cultivars that derive their resistance from a very limited pool of resistance sources, over time the virulence of SCN field isolates has grown dramatically (Niblack et al., 2008). Currently, the overwhelming majority of soybean cultivars derive their resistance from a single resistance source, PI 88788. Because of this dramatic increase in virulence it has become more important than ever to determine the genes underlying virulence and better understand how virulence on different resistance sources is related so that more informed management decisions can be made.

Resistance to SCN typically manifests at the nematode feeding site, resulting in the degeneration of the syncytium and subsequent demise of the nematode. However, many characteristics of a resistance response differ depending on the source of SCN resistance. Resistance derived from the cultivar Peking is typically characterized as a more rapid response with dilation of the rough endoplasmic reticulum occurring as early as two days post infection (Kim et al., 2012). This resistance response is typically characterized by nuclear degeneration and necrosis initiated inside the syncytium (Riggs et al., 1973). This may result in termination of nematode development as early as five days post infection (Kim et al., 1987). In contrast, resistance derived from PI 88788 involves a slower response with necrosis initiating outside the syncytial cell wall and cytoplasmic degeneration occurring by 8-10 days post infection (reviewed in Mitchum, 2016). Finally, resistance derived from PI 437654 appears similar in mechanism to

Peking-derived resistance, however an additional mechanism must be present to distinguish between the two, as Peking-virulent nematodes are not guaranteed to be virulent on PI 437654 (Gardner et al., 2017; Niblack et al., 2002). In order for the nematode to successfully reproduce it must find a way to circumvent these responses and maintain a feeding site for the duration of its life cycle.

Previous studies have attempted to identify the genes responsible for virulence in SCN. One of the first studies of SCN virulence genes identified the *ror* (for reproduction on a resistant host) genes (Dong and Opperman, 1997). These are genes responsible for virulence on PI 88788 and PI 90763, and were shown to be inherited in a dominant and a recessive manner respectively. However, while these genes have been identified in a genetic study, the actual function and role they play in nematode parasitism was not further characterized. In more recent years, studies have utilized more modern tools to identify virulence factors in SCN. A recent genomic SNP array study identified two SCN genes, one encoding a biotin synthase (*HgBioB*) and the other encoding a bacterial-like protein containing a SNARE domain (*HgSLP-1*), correlated with virulence on PI 88788 (Bekal et al., 2015). *HgBioB* contained polymorphisms between the avirulent and virulent populations while *HgSLP-1* was absent in virulent populations, which the authors hypothesize allows for evasion of host resistance responses. Follow up studies will need to be performed to characterize the function of these two virulence candidates and ensure their role in virulence goes beyond correlation. In addition, the correlation of these candidates with virulence remains to be confirmed in field populations. This study showed that not only are multiple genes necessary for virulence, but the mechanisms to grant that virulence may differ as well. To better understand SCN virulence this type of

analysis needs to be expanded to the other two types of SCN virulence, PI 437654 and Peking, so that we can understand what differentiates these mechanisms of virulence and prevent the emergence of a nematode virulent on all three sources of resistance. Recent genetic studies suggest that virulence on PI 437654 is likely a combination of virulence genes, whereas virulence on Peking is likely due to a recessive gene (Gardner et al., 2017). The identity and molecular function of these genes, however, remains unknown.

The objective of this study was to perform a comparative analysis across multiple populations virulent on the three primary sources of resistance to SCN (Peking, PI 88788, and PI 437654) to identify molecular pathways that either overlap or serve to differentiate each mechanism of SCN virulence. In order to perform this analysis a *de novo* virulent SCN transcriptome was generated and annotated. Differential expression analysis using this transcriptome led to the identification of enriched gene categories that link virulence in a population virulent on PI 88788 (MM8) and one virulent on PI 437654 (MM10) together. In addition, expression of known SCN stylet-secreted effectors was analyzed and a suite of differentially expressed effectors was observed. This information can be utilized to pursue further molecular and functional characterization of genes involved in virulence. Differentially expressed effectors represent potential markers for virulence that can be used to improve identification of virulent SCN populations.

RESULTS

Transcriptome sequencing and assembly

Four populations were chosen representing virulence on the three primary sources of SCN resistance and one avirulent population (Table 4.1). The populations MM1 and MM2 are the result of mass selection of the HG type 0 population PA3 on the soybean lines ExF63 (susceptible) and ExF67 (resistant). ExF63 and ExF67 are recombinant inbred soybean lines that only differ at the *Rhg4* locus. The SCN populations MM1 and MM2 are highly related and any differences between the two are likely to be related to the resistance found at that locus. The populations MM8 and MM10 are the result of a cross between PA3 (HG type 0) and TN20 (HG type 1-7) that have been out-selected on the resistant soybean lines PI 88788 and PI 437654, respectively (Gardner et al., 2017). In both these populations virulence is derived from the TN20 population, however the MM8 population only received genes allowing virulence on PI 88788-type resistance whereas MM10 is virulent on all sources of SCN resistance. mRNA sequencing of the four populations (three virulent, one avirulent) in three total replicates for each yielded a total of approximately 649 million paired end reads used for transcriptome assembly. The pJ2 MM1 sample was sequenced using 100bp reads while pJ2 MM2, pJ2 MM8, and pJ2 MM10 samples were all sequenced using 50bp reads. The pJ2 MM1, pJ2 MM2, pJ2 MM8, and pJ2 MM10 samples were all used as input for Trinity *de novo* transcriptome assembly. The resulting assembly consisted of 185,137 transcripts with a total assembly length of 59.2 Mb and an N50 of 1,415 bp (Table 4.2). Based on the number of input reads and the size of the assembly this would result in an estimated 21-fold coverage of the transcriptome. The assembly also contained 367 (309 complete and 61 fragmented) of the total 429 BUSCOs (Benchmarking Universal Single-Copy Orthologs) in the eukaryote dataset (Simão et al., 2015).

Population	Host Cultivar	Resistance Source	Population Development
MM1	ExF63	None	PA3 mass selected on ExF63
MM2	ExF67	Peking	PA3 mass selected on ExF67
MM8	PI 88788	PI 88788	PA3 x TN20 cross out-selected on PI 88788
MM10	PI 437654	PI 437654	PA3 x TN20 cross out-selected on PI 437654

Table 4.1. Description of *Heterodera glycines* populations used for transcriptome assembly. Population histories for the four input SCN populations used in transcriptome assembly.

	SCN Virulent pJ2 <i>de novo</i> assembly
Number of transcripts	185,137
Total Assembly length (Mb)	59.2
Number of Trinity ‘genes’	100,233
N50 (bp)	1,415
Avg. contig length (bp)	852.73
Predicted proteins	97,740
BUSCO* score	C: 71%, F: 14.5%, M: 14.5%

Table 4.2. *de novo* transcriptome assembly statistics for the virulent SCN transcriptome assembly. A *de novo* transcriptome assembly was generated using Trinity from the parasitic second-stage juveniles of four different SCN populations virulent on each of the three main sources of SCN resistance. The transcriptome was assessed for completeness using the tool BUSCO (benchmarking universal single-copy orthologs) to identify complete (C), fragmented (F), and missing (M) sequences representing conserved orthologs found in all eukaryotes.

Annotation of transcripts

Transcripts for the virulent SCN transcriptome were annotated following the Trinotate pipeline (Haas et al., 2013). Transcripts were compared to the Swiss-Prot database (Boutet et al., 2016) using BLASTX and identifying available homologues for 60,576 (31%) of the 185,137 transcripts at an e-value cutoff of $1e-5$ (Figure 4.1). The homologues identified represent 926 different species ranging from model organisms such as *Arabidopsis thaliana* and *Caenorhabditis elegans* to more diverse organisms such as *Aphelenchus avenae* and *Globodera rostochiensis*. The Swiss-Prot database was chosen as it is highly curated and provides detailed annotation information for each sequence represented, which facilitated down-stream gene ontology analysis.

Differential expression analysis

Differentially expressed transcripts were predicted between each of the four populations using DESeq2 at a minimum 2-fold expression change and p-value cutoff of 0.05 (Figure 4.2.A). The MM1 and MM2 populations followed a very similar expression pattern, resulting in only 32 differentially expressed transcripts between these populations. In contrast, the MM10 population had the most drastic change in expression from the MM1 and MM2 populations, resulting in 8353 and 7953 predicted differentially expressed transcripts respectively.

Comparisons between the three virulent populations and the avirulent MM1 population were conducted to identify potential overlap in virulence mechanisms and any highly conserved virulence genes that may be good targets for broad resistance (Figure 4.2.B). This analysis identified five (two up-regulated and three down-regulated)

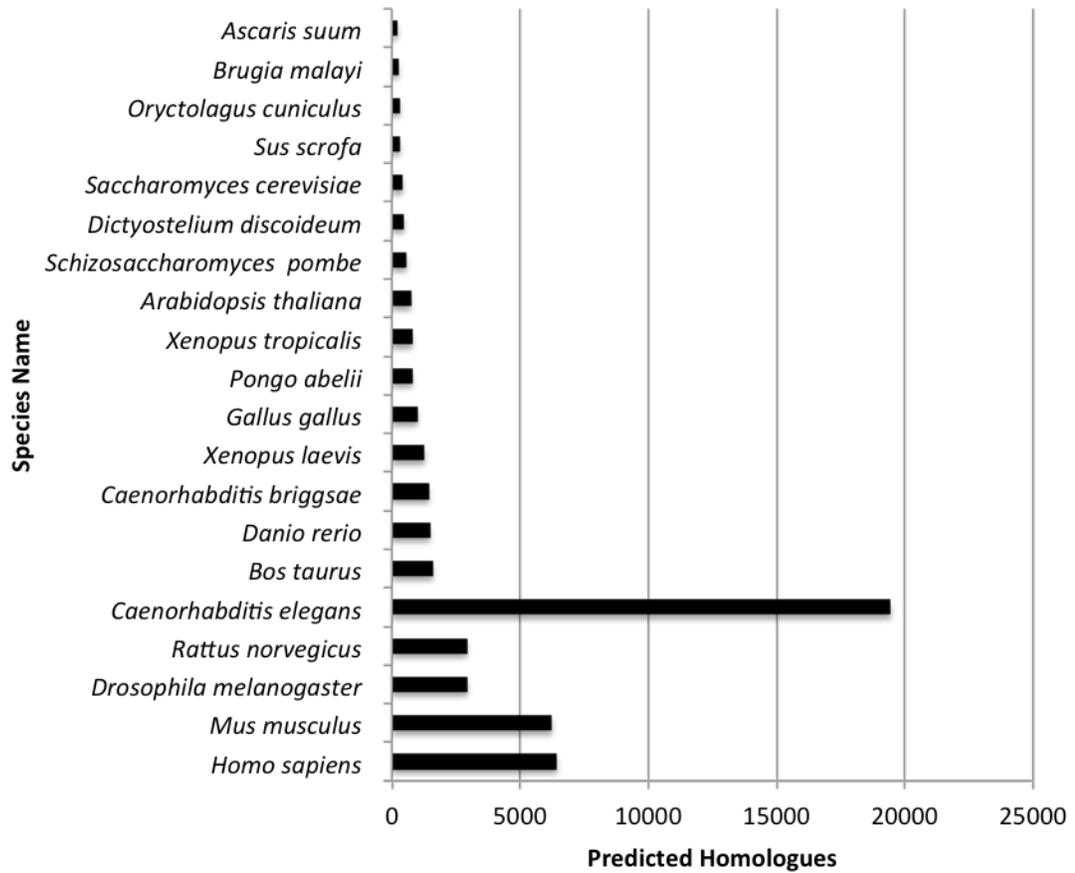


Figure 4.1. Species distribution of predicted homologues to the virulent pJ2 transcriptome of *H. glycines*. Homologues were predicted using a BLAST search to protein databases at an e-value cutoff of $1e-5$. The top twenty species in terms of quantity of predicted homologues are shown here.

A

	MM1	MM2	MM8	MM10
MM1	0	32	4902	8353
MM2		0	4584	7953
MM8			0	5302
MM10				0

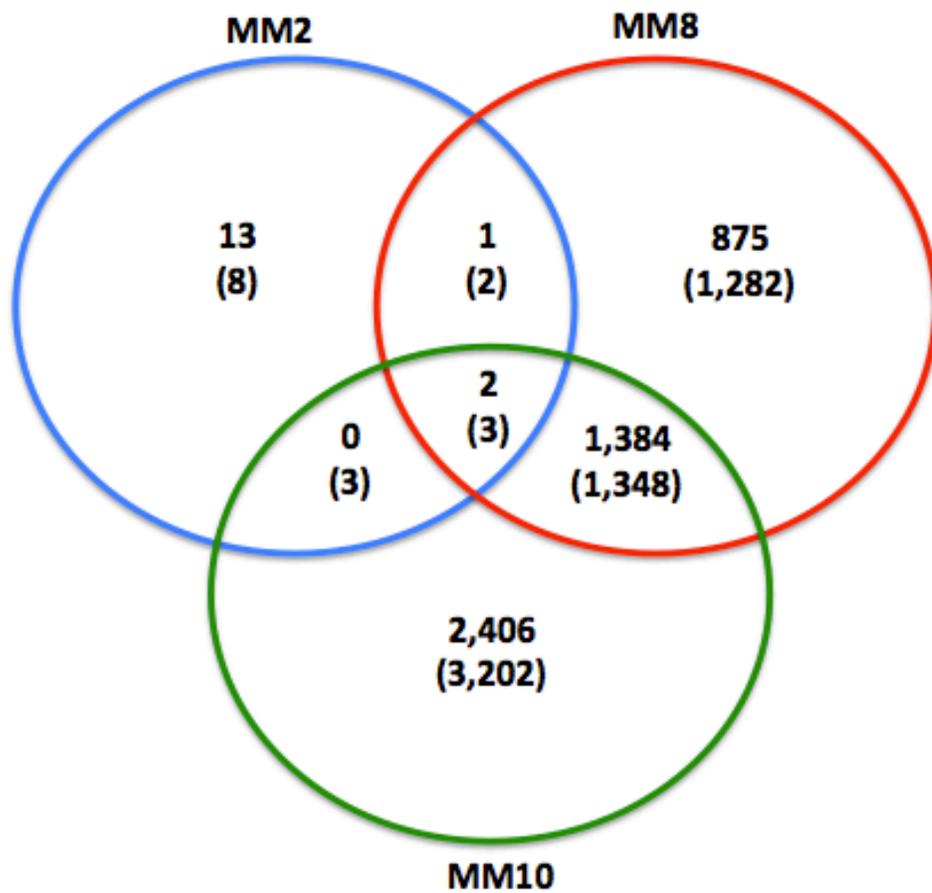
B

Figure 4.2. Differential expression analysis of the virulent SCN transcriptome. A) pairwise comparison of differentially expressed transcripts between each of the four populations. B) Venn diagram of differentially expressed transcripts in each of the virulent populations compared to the avirulent population. Numbers in parentheses are down-regulated in the virulent sample relative to the avirulent sample.

transcripts with expression patterns conserved between all three virulent populations (Figure 4.3.A). The up-regulated transcripts are annotated as an extracellular serine/threonine protein kinase and a novel protein (Figure 4.3.B). The three down-regulated transcripts encode a viral nucleoprotein, an Argonaute (AGO) protein, and a large subunit GTPase 1 homolog. These three transcripts exhibit a dramatic reduction in expression in all three virulent populations, particularly the MM8 and MM10 populations where they are not detected at all (Figure 4.3.C). Interestingly, there appeared to be significantly more overlap between differential expression in the MM8 and MM10 populations than in the MM2 population, with 2,732 transcripts overlapping in their expression pattern, of which 1,384 are up-regulated from MM1 and 1,348 are down-regulated from MM1. This may be an indication of a closer relatedness between the mechanism of these two types of virulence, the PI 88788 virulence and PI 437654 virulence, than that of Peking virulence but may also reflect the similarities between these two populations due to their virulent parent TN20.

Differential expression of known SCN effectors was also examined between the virulent populations and MM1. Transcripts with the highest sequence similarity to each of the previously published SCN effectors were identified (Bekal et al., 2015; Gao et al., 2003; Noon et al., 2015) through a BLASTN search, and those effectors mapping to the same transcript were grouped into an effector family for this analysis. Interestingly, the comparison of MM1-MM2 yielded no differentially expressed SCN effectors above a 2-fold minimum expression change and p-value cutoff of 0.05. In contrast, both MM8 and MM10 have suites of effectors either up-regulated or down-regulated relative to the avirulent population (Figure 4.4). Both MM8 and MM10 show up-regulation of the

A

Transcript ID	Annotation	MM1	MM2	MM8	MM10
c46098_g6_i6	Novel	0	8.3	119.1	96.8
c48103_g2_i8	Extracellular serine/threonine protein kinase	4.2	8.0	8.7	8.7
c51328_g1_i10	Argonaute protein wago-1	27.4	11.9	0	0
c51368_g4_i3	Viral Nucleoprotein	1.9	0	0	0
c53746_g1_i5	Large subunit GTPase 1 homolog	2.21	0	0	0

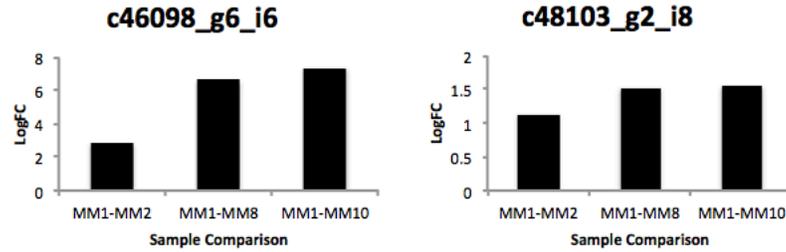
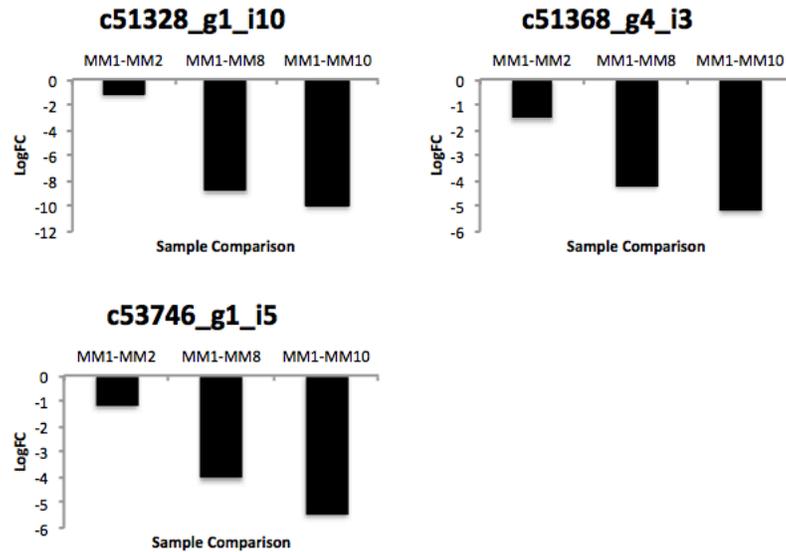
B**C**

Figure 4.3. Transcripts with conserved expression patterns across virulent populations of *H. glycines*. Transcripts were identified with conserved differential expression between the avirulent SCN population MM1 and all three virulent populations MM2, MM8, and MM10. The corresponding annotations and TMM-normalized fragments per kilobase of transcript per million mapped reads (FPKM) were retrieved (A) and DESeq2 calculated differential expression values for both up-regulated (B) and down-regulated (C) transcripts.

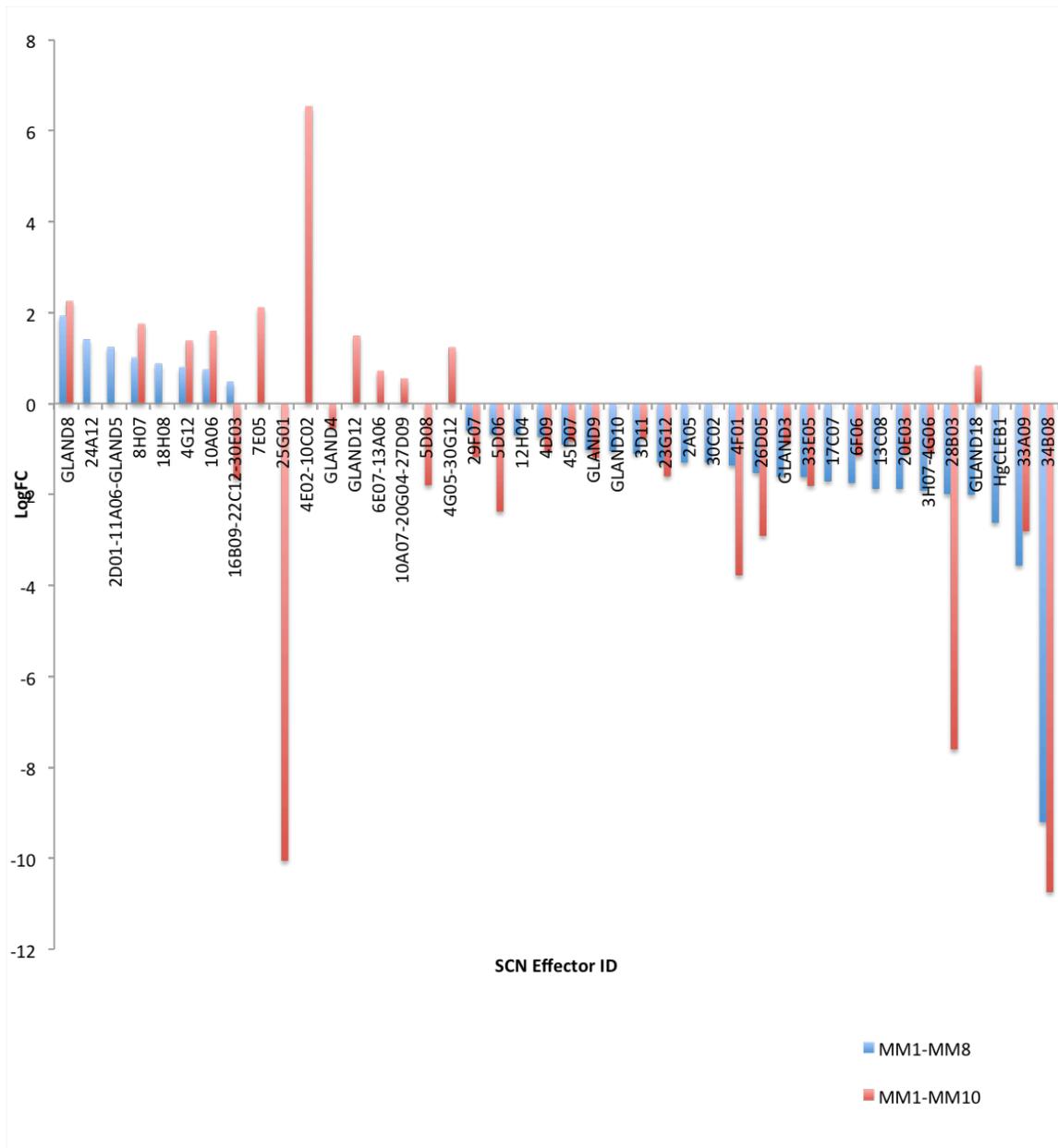


Figure 4.4. Differential expression of known SCN effectors across the virulent populations MM8 and MM10 relative to the avirulent population MM1. Transcripts corresponding to the 72 known SCN effectors were identified via a BLASTN search to the virulent SCN transcriptome. Expression for the resulting transcripts was then compared between each virulent population and the avirulent population MM1 using DESeq2 to identify and differential expression over a 2-fold minimum expression change and above a p-value cutoff of 0.05.

effectors 8H07, 4G12, GLAND8, and 10A06 compared to MM1. The down-regulated effector pool shared between both of these populations is much larger, containing 29F07, 5D06, 4D09, 45D07, GLAND9, 3D11, 23G12, 4F01, 26D05, GLAND3, 33E05, 6F06, 20E03, 3H07-4G06, 28B03, 33A09, and 34B08. In addition, there are several effectors with unique expression patterns in one of the two virulent populations. The effectors/effector families 24A12, 2D01-11A06-GLAND5, 18H08, 2A05, 30C02, and 17C07 are all predicted to be differentially expressed only between MM1 and MM8 while the effectors/effector families 16B09-22C12-30E03, 7E05, 25G01, 4E02-10C02, GLAND4, GLAND12, 6E07-13A06, 10A07-20G04-27D09, 5D08, and 4G05-30G12 are only predicted to be differentially expressed between MM1 and MM10.

Gene ontology analysis

The list of differentially expressed genes between virulent and avirulent populations was then examined for enrichment of functional gene categories using the bioinformatics tool GSeq (Young et al., 2010). GSeq is an application designed to perform gene ontology analysis on RNASeq and was chosen here due to its proven high accuracy and ability to correct for selection bias due to sequence length. Gene ontology information was parsed from the annotations available for each transcript and gene categories over-represented in each of the comparisons of virulent to avirulent population were considered enriched (Appendix 4). The top fifteen most enriched functional gene categories as determined by p-value were examined for potential roles they may play in soybean cyst nematode parasitism and virulence. The virulent population MM2 had no enriched gene categories in either the up- or down-regulated genes relative to MM1. This

is potentially a reflection of how closely related these two populations are. Alternatively, virulence on Peking resistance may be due to variation in a small number of genes and so not detected through this analysis or may not be due to differential gene expression at all.

The virulent MM8 population had many enriched functional gene categories in both the up- and down-regulated genes relative to the avirulent population MM1 (Figure 4.5). The up-regulated genes in MM8 are enriched in many pathways related to cell wall organization and negative regulation of conjugation, as well as external encapsulating structure and extracellular region. Of the down-regulated genes in MM8 many of the top functional categories are related to cellulose metabolism, cation binding, and metal ion binding. The gene categories related to cellulose catabolism may be related to initial feeding site establishment and migration. The fact that these genes are down-regulated in MM8 is a potential indication that those nematodes in the virulent population have progressed past the earliest stages of host invasion and initial feeding site establishment. In addition, functional gene categories related to DNA integration and DNA polymerase activity are also down-regulated.

The MM10 population has some very intriguing enriched functional categories represented in the up-regulated genes (Figure 4.6). These include plasma membrane organization, oocyte construction, germarium-derived female germ-line cyst formation, and several categories related to retinal cell programmed cell death. The down-regulated genes are enriched in several categories similar to MM8, including those related to cellulose processing, cation and metal ion binding, and several other metabolic processing categories. Once again, this may be potentially linked to a more rapid developmental progression than the avirulent population MM1.

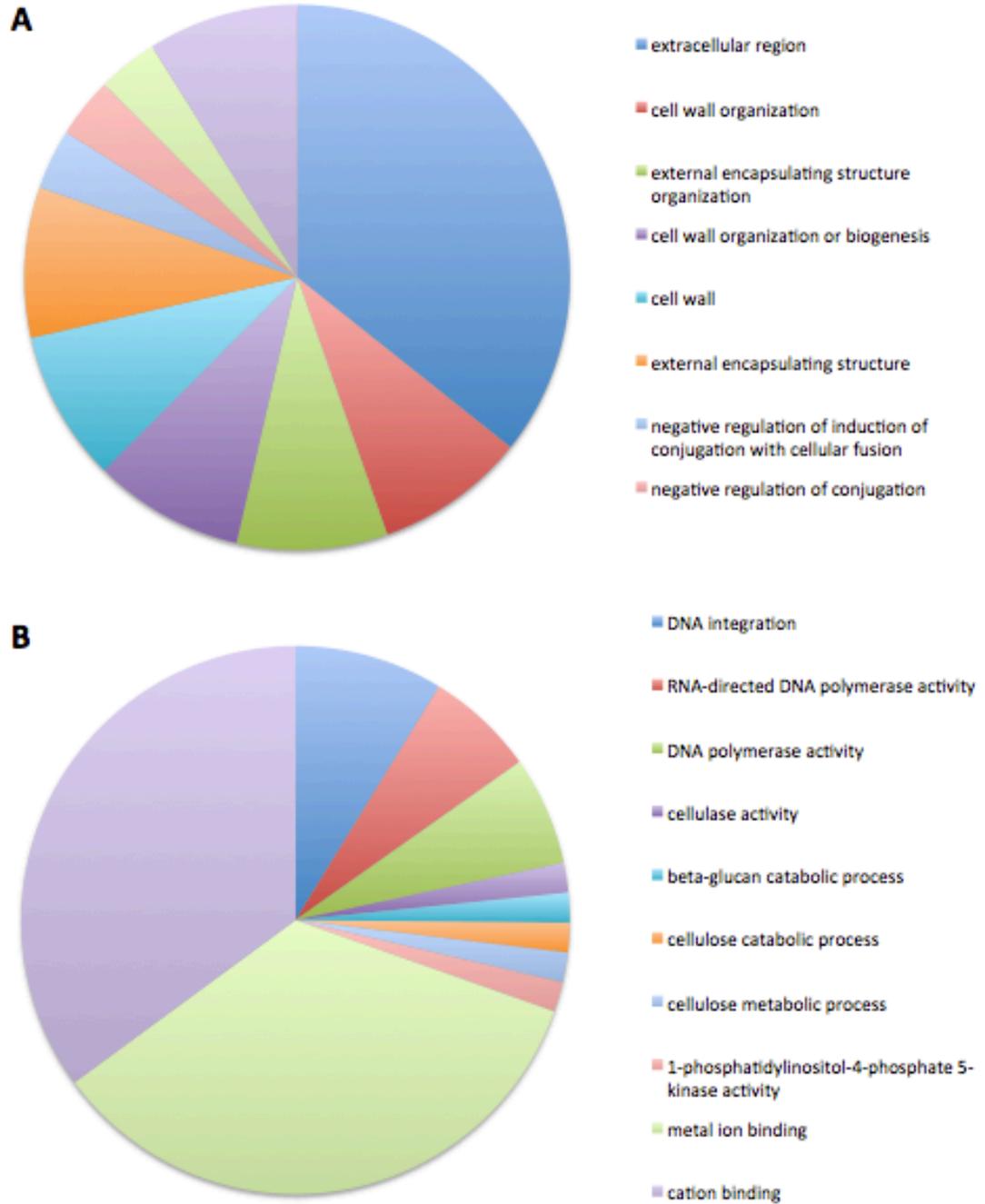


Figure 4.5. Gene enrichment analysis of virulent MM8 population compared to the avirulent population MM1. Available gene ontology terms were pulled following the Trinotate pipeline and gene enrichment analysis was performed using GOSeq to identify the top ten most highly enriched genes in the up-regulated (A) and down-regulated (B) gene sets.

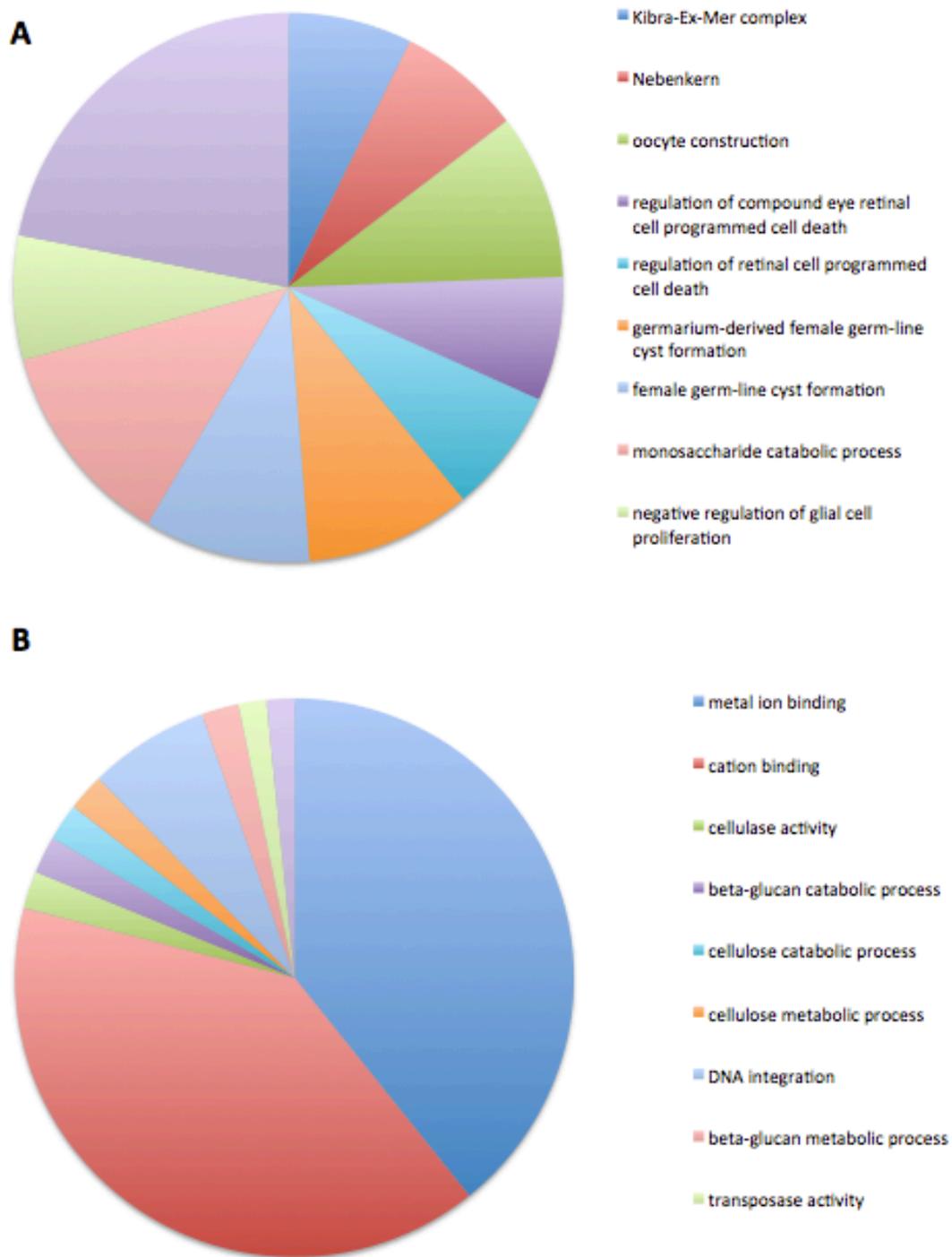


Figure 4.6. Gene enrichment analysis of the virulent MM10 population compared to the avirulent population MM1. Available gene ontology terms were pulled following the Trinotate pipeline and gene enrichment analysis was performed using GOSeq to identify the top ten most highly enriched genes in the up-regulated (A) and down-regulated (B) gene sets.

DISCUSSION

In this study we describe construction and analysis of a virulent SCN transcriptome, allowing for comparisons between different populations of this parasitic nematode. This transcriptome contains transcripts from populations virulent on each of the main sources of resistance to SCN: Peking, PI 88788, and PI 437654. By assembling these transcripts together we were able to perform comparative analysis between these populations and a population of SCN avirulent on all three sources of resistance to identify differences in expressed sequences that may correlate to virulence. The transcriptome assembly focuses on an early parasitic stage of SCN, specifically the parasitic second-stage juvenile. At this point in its life cycle, the nematode has successfully invaded the host plant and begun to establish its feeding site. This time point was specifically chosen because resistance to SCN occurs fairly early during infection, often killing avirulent nematodes within 5-8 days, depending on the resistance (Mitchum, 2016). Any adaptations that virulent nematodes have must be enacted at this early stage in order to ensure a successful infection. *De novo* transcriptome assembly resulted in a transcriptome of nearly 200,000 transcripts, which reflects the large amount of diversity between and within the populations sequenced, even at one specific life-stage and time point. The fact that this transcript diversity exists even after years of inbreeding suggests that the nematode has mechanisms to ensure genetic variability, potentially as a survival mechanism that would aid in the development of virulence.

To compare the four populations and identify features responsible for contrasting outcomes of parasitism, transcripts were quantified and a differential expression analysis

was performed. This resulted in the identification of subsets of transcripts with conserved expression patterns between all three virulent populations as well as unique transcripts in each of them, including five transcripts conserved across all three virulent populations. The five transcripts with conserved expression across all three virulent SCN populations have a broad array of predicted functions based on their annotations, ranging from a viral nucleoprotein to an extracellular protein kinase. None of the five transcripts are predicted to contain a signal peptide but they may still be secreted from the nematode through non-classical mechanisms, much like the detoxifying enzymes produced by *Globodera rostochiensis* (Robertson et al., 2000). This is likely the case for the extracellular serine/threonine kinase. Extracellular protein kinases have been characterized in several mammalian pathogens, including both bacteria and the nematode *Trichinella spiralis* (Arden et al., 1997; Canova and Molle, 2014). In *T. spiralis* a nucleoside diphosphate kinase was confirmed to be secreted by the infective stage and predicted to be involved in regulation of host cell proliferation and suppression of innate immune responses in the host cell nucleus (Afferson et al., 2012; Gounaris et al., 2002). It is not unreasonable to predict a similar function in SCN, however the transcript identified here has no predicted nuclear localization signal, meaning it would not be imported to the host cell nucleus and likely functions elsewhere in the interior of the cell.

The three down-regulated transcripts shared between the virulent populations shared homology with Argonaute protein wago-1, large subunit GTPase 1, and a viral nucleoprotein. The worm-specific argonautes (WAGOs) are part of the RNAi pathway and have been shown to be very diverse in nematodes, potentially playing roles in environmental sensing and transgenerational memory (Buck et al., 2013). In addition,

these genes are enriched in the genomes of plant-parasitic nematodes and certain members of the WAGO family are up-regulated at different infective stages of the potato cyst nematode *Globodera rostochiensis* (Eves-van den Akker et al., 2016). The virulent populations show a marked reduction in expression of the WAGO-1 homolog, which would reduce the silencing effect it has on genes regulated by this homolog, causing an increase in expression. The genes regulated by the WAGO-1 homolog need to be elucidated in order to determine the effect this down-regulation has on SCN parasitism. Further characterization is necessary to confirm whether or not this gene is epigenetically regulated and if that regulation is due to the host plant. The second down-regulated transcript shared among all three virulent populations shows homology to a viral nucleoprotein from the genus *Bornavirus*. Nematodes have been known to harbor plant viruses and more recently were shown to be infected by viruses themselves (Brown et al., 1995; Felix et al., 2011). In *H. glycines* specifically viruses from multiple families have been identified, including viruses related to nyaviruses, bornaviruses, flaviviruses, and bunyaviruses (Bekal et al., 2011; Bekal et al., 2014). The down-regulation of the viral nucleoprotein in the virulent populations compared to the avirulent population MM1 is a potential indication of reduced viral titer in the virulent populations. Little is known about the effects of viruses on *H. glycines* but it is possible reduced virus levels result in increased health of the nematode, allowing it to respond to the plant defense response and successfully infect. These five transcripts need to be further investigated and their functions confirmed in order to determine whether they function in virulence or are merely correlative.

Effectors are another aspect of nematode biology that may explain the differences between virulent SCN populations. The soybean cyst nematode has an established repertoire of 72 stylet-secreted effectors (Bekal et al., 2015; Gao et al., 2003; Noon et al., 2015). These effectors have been shown to play multiple roles in the infection process including suppression of host defenses (Gheysen and Mitchum, 2011). To analyze the expression of SCN effectors between the four populations we identified the corresponding transcripts for each of the known effectors using BLASTN and then isolated those effectors differentially expressed between the avirulent population MM1 and the three virulent populations. None of the known SCN effectors were predicted to be differentially expressed between the MM1 and MM2 populations, which could be the result of these two populations being so similar or an indication that virulence on Peking-type resistance is not mediated by any of these known stylet-secreted effectors. The MM8 and MM10 populations both show a greater overlap in differentially expressed genes compared to the MM1 population. There are multiple common genes up-regulated in both of these virulent populations, including 4G12, GLAND8, 8H07, and 10A06. While GLAND8 and 8H07 both have relatively little functional characterization, a homolog of 10A06 has been shown to be involved in defense suppression in the closely related beet cyst nematode *Heterodera schachtii* (Hewezi et al., 2010). The *H. schachtii* homolog of 10A06 (Hs10A06) was expressed in Arabidopsis, leading to an increase in susceptibility of the plant to *H. schachtii* as well as the bacterial pathogen *Pseudomonas syringae* pv tomato (Pst DC3000) and the viral pathogen *Cucumber mosaic virus* (CMVY). Hs10A06 was also found to interact with Arabidopsis spermidine synthase 2 (SPDS2), an enzyme involved in polyamine synthesis and this interaction is thought to

result in disruption of salicylic acid signaling and increased antioxidant production in the syncytium. Presumably this effector plays a similar role in *H. glycines* and when the virulent nematodes from MM8 and MM10 encounter host resistance they produce more of this effector to protect the nematode from host defense responses. Alternatively, these nematodes may also be developing faster than their avirulent counterparts and up-regulating effectors accordingly. The *Heterodera glycines* effector 4G12, which is up-regulated in both MM8 and MM10 relative to the MM1 population, has been identified as a CLAVATA3/endosperm surrounding region-related (CLE) ligand mimic and is known to modulate host developmental programs for feeding cell formation (Olsen and Shriver, 2003; Wang et al., 2001; Wang et al., 2005). Once secreted into the plant these CLEs are thought to function as ligand mimics of endogenous CLE peptides and so producing more of these mimics may be required in order to overcome any defense responses mounted by the plant. In addition, both MM8 and MM10 populations have up-regulated effectors unique to the interaction with their respective resistance (PI88788 and PI437654). The MM8 population showed an upregulation of effectors 24A12, 18H08, and the effector family of 2D01-11A06-GLAND5. None of these effectors have been functionally characterized in depth and so it remains to be seen how they contribute to virulence on PI88788-derived resistance. In contrast, the MM10 population shows unique up-regulation of the effectors 7E05, 4E02-10C02, GLAND12, 6E07-13A06, 10A07-20G04-27D09, and 4G05-30G12. These effectors again lack much in the way of functional characterization, however, it should be noted that the effector family 4E02-10C02 shows a dramatic up-regulation with over 6-fold higher expression in the virulent MM10 population compared to the MM1 population. This high expression is a likely indication

of an important role in virulence that can be better understood once more is known about this effector. The virulent populations MM8 and MM10 also share a pool of down-regulated effectors in comparison to MM1. Effectors of note in this list include 4F01 and 45D07. Both of these effectors have been implicated in host defense suppression, with the effector 4F01 encoding an annexin-like protein while 45D07 encodes a chorismate mutase (Bekal et al., 2003; Gao et al., 2003). The virulent nematodes may be overcoming the host defense response through other effectors up-stream of the targets of these effectors and so need to produce lower amounts of these effectors to produce the same effect. Each of the virulent populations also had unique down-regulated genes as well, with the most highly down-regulated effector being 25G01 in the MM1-MM10 comparison. 25G01 encodes an endoglucanase, one of multiple cell-wall degrading enzymes secreted by the nematode as it invades the host root (Gao et al., 2003; Wang et al., 1999). These enzymes are vital in the migratory process of *H. glycines* and their down-regulation in MM10 suggests perhaps a faster migration into the host or reduced expression to avoid host recognition. Only two effectors had alternating expression patterns between the comparisons of MM1-MM8 and MM1-MM10, those effectors being GLAND18 and the 16B09-22C12-30E03 effector family. These effectors are of particular interest and may serve to differentiate between nematodes virulent on PI88788-type resistance and those virulent on PI 437654-type resistance. The overall similarity of the MM8 and MM10 populations in their effector expression patterns indicates a potential link in the mechanism of virulence. In contrast, the relatively similar expression of effectors between MM1 and MM2 is indicative that the determining factors in

virulence are likely not related to expression of the known effectors and more likely due to polymorphisms or other genes within the nematode.

To provide a broader spectrum view of the comparison between these nematodes gene ontology terms were assigned to differentially expressed genes and a gene enrichment analysis was performed. This identified functionally enriched gene categories in the virulent populations MM8 and MM10 but not MM2 relative to the avirulent population MM1. In each comparison the functionally enriched gene categories can be divided into two categories, either up-regulated or down-regulated from the avirulent to the virulent population. The most highly enriched gene categories in the up-regulated gene set between the MM1 and MM8 population are mainly associated with negative regulation of cellular fusion (GO:0030312, GO:0010515, GO:0031135, GO:0031138) or categories affecting the external region of the nematode, including the cell wall (GO:0071555, GO:0071554, GO:005618) and extracellular region (GO:0005576, GO:0045229). Differentially expressed genes in these categories include many genes related to chitin biosynthesis, a very important process in nematodes. Chitin plays many different roles in nematodes and has been implicated in egg production as well as in the feeding apparatus of the nematode *Oesophagostomum dentatum* and the cuticle of the animal parasite *Ascaris suum* (Neuhaus et al., 1997; Veronico et al., 2001). Based on the life-stage utilized for this transcriptome it seems most likely that the chitin biosynthesis is related to the cuticle and molting in *Heterodera glycines*. This would indicate that the MM8 population is possibly developing faster than MM1, despite the resistance present in its host. The most enriched gene categories in those genes down-regulated between MM1 and MM8 include binding of different ions (GO:0046872,

GO:0043169) and cellulose catabolic and metabolic processes (GO:008810, GO:0030245, GO:0030243). The ion binding categories have been shown to be down-regulated during the molting process of the free-living nematode *Caenorhabditis elegans* and likely play a similar role here in *H. glycines* (Turek and Bringmann, 2014). In addition, the down-regulation of cellulases has been observed before in *H. glycines* typically between 2-5 days post infection (Ithal et al., 2007). Together these two pieces of evidence point towards a more rapid development of the virulent population MM8 than MM1, which may allow it to overcome any resistance put forward by the host plant. Down-regulated genes between MM1 and MM10 fall into similar categories of ion binding and cellulose-related processes, likely indicating advanced development in this population as well. The enriched categories of up-regulated genes, however, are very different in the MM1-MM10 comparison. These include multiple categories related to female nematode development, including oocyte construction (GO:007308) and female germ-line cyst formation (GO:0048135). In addition, functional gene categories related to regulation of programmed cell death are also present (GO:0046669, GO:0046668). Contained within these gene sets are multiple transcripts with homologs associated with suppression of cell death and apoptosis including death-associated inhibitor of apoptosis 1 (Diap1) and SPRY domain-containing SOCS box protein 3 (SPSB3). Diap1 has been shown to disrupt the activities of enzymes promoting apoptosis in *Drosophila melanogaster* and may play a protective role in *H. glycines* (Wang et al., 1999). Meanwhile, members of the SPSB gene family are involved in protein modification, specifically ubiquitination (Kuang et al., 2010). While this gene family has not been well characterized in nematodes, it has been studied thoroughly in mammalian systems and

other members of the family implicated in targeting genes for proteasomal degradation. In addition, a group of nematode effectors containing a SPRY domain has been previously identified and characterized in *Globodera rostochiensis* and *Globodera pallida* and shown to play an important role in both suppressing and triggering resistance (Postma et al., 2012; Sacco et al., 2009). If this transcript were to be secreted it may play a role in virulence for the MM10 nematode population. Both of these genes are interesting candidates for further potential characterization in order to understand the molecular mechanisms underlying virulence.

This study compared populations of the parasitic nematode *H. glycines* virulent on the primary resistance sources used to control this pathogen (Peking, PI 88788, PI 437654). To do this a transcriptome was generated of the three virulent populations and one avirulent population at five days post infection and then comparative analysis was performed. It appears that the populations virulent on PI 88788 and PI 437654 are developing more quickly than the avirulent population, as confirmed by the gene ontology analysis. In addition, effector expression patterns were found to be variable depending on the population, with some very highly up-regulated to perhaps compensate for host resistance. Meanwhile, the population virulent on Peking was too similar to the avirulent population in expression patterns to perform global analyses. Virulence on this type of resistance is likely linked to polymorphisms or gene variants that allow virulent nematodes to bypass resistance and successfully infect the plant. The insight and tools generated by this study will hopefully lead to further functional characterization of the molecular pathways associated with *H. glycines* virulence and allow for better management of this harmful pathogen.

MATERIALS AND METHODS

Nematode cultivation and isolation

The SCN inbred populations MM1, MM2, MM8, and MM10 were propagated under greenhouse conditions on soybean cultivars ExF63, ExF67, PI 88788, and PI 437654 respectively. The two soybean lines ExF63 and ExF67 are recombinant inbred lines that differ at the *Rhg4* locus, with ExF67 being resistant to SCN while ExF63 is not resistant (Liu et al., 2012). Freshly hatched pre-parasitic second stage juveniles (ppJ2) from each population were inoculated onto 10-day old seedlings of their respective hosts and the inoculated plants were placed in greenhouse conditions. Five days post infection the pJ2 nematodes were isolated from the roots by blending the roots for 30 seconds in a kitchen blender, following which the root homogenate was run over a series of three sieves. These sieves had pore sizes of 850um, 250um, and 25um and once the samples were run through them the nematodes were purified from the remaining root material using sucrose centrifugation (de Boer et al., 1999). The extracted parasitic second-stage juveniles were pelleted and frozen in liquid nitrogen before being stored at -80°C. Three biological replicates were obtained for each sample before proceeding to RNA isolation.

RNA Isolation and Sequencing

RNA was isolated from frozen nematode pellets using the PerfectPure Fibrous Tissue Kit (5Prime) and a modified version of the manufacturer's extraction protocol.

Tissue was homogenized in 30-second intervals in the provided lysis solution containing 0.5 uM TCEP. The homogenization was performed using a bead beater and 1.0 mm zirconia beads and was followed by a 30 second incubation on ice. The process of homogenization and ice incubation was repeated three times, following which the sample was centrifuged briefly at room temperature before transferring the supernatant to a fresh tube. Following lysis and homogenization, 10uL of the provided Proteinase K solution was added and the sample allowed to incubate on ice for 10 minutes, after which the manufacturer's protocol for RNA purification was followed to completion. RNA quality was determined using a Fragment Analyzer (Advanced Analytical) and quantified using a Qubit Fluorometer prior to library preparation. RNA-seq libraries were constructed using the TruSeq mRNA Stranded Library Prep Kit (Illumina) and sequenced on the Illumina HiSeq 2500 platform. pJ2 MM1 samples were sequenced using 100bp reads while pJ2 MM2, pJ2 MM8, and pJ2 MM10 samples were all sequenced with 50bp reads. Library preparation and high throughput sequencing services were performed at the University of Missouri DNA Core Facility.

De novo transcriptome assembly

Low quality reads were removed prior to assembly using Trimmomatic (Bolger et al., 2013) and then the remaining reads were paired. High-quality paired end reads were used as the input for eventual transcriptome assembly. The transcriptome assembly itself was performed using Trinity, a de Bruijn graph-based tool (Grabherr et al., 2011). In silico normalization was performed during assembly to reduce computational time, following which assembly quality was assessed by mapping reads from each replicate

back to the transcriptome using Bowtie2 at default parameters (Langmead and Salzberg, 2012).

Transcriptome annotation and quantification

The transcriptome was annotated following the established Trinotate pipeline (Grabherr et al., 2011). Homology searches were performed against the highly curated protein sequences contained in the Swiss-Prot (Magrane et al., 2011) database using BLASTX at an e-value cutoff of $1e-5$. Transcripts were translated in six frames and the most likely protein sequence chosen using TransDecoder (Grabherr et al., 2011). HMMER and Pfam databases were used to predict protein domains contained within each transcript (Finn et al., 2011; Punta et al., 2012). Signal peptides were predicted using SignalP 4.0 and TMHMM was utilized to identify predicted transmembrane domains (Petersen et al., 2011; Krogh et al., 2001). All of the ensuing annotation information was then combined and pooled into a SQLite database.

For quantification and differential expression analysis reads from each replicate of each of the four samples used for assembly were mapped and quantified using RSEM (Li and Dewey, 2011). Differential expression analysis was then conducted using DESeq2 (Love et al., 2014) to identify all genes with a minimum 2-fold expression difference and under a p-value cutoff of 0.05 between any of the populations.

Gene ontology enrichment analysis

Gene ontology and enrichment analysis was performed as outlined in the Trinity and Trinotate pipelines. Briefly, available GO assignments were extracted from the

annotation database and then input into the Bioconductor package GSeq (Young et al., 2010). Functional enrichment tests were performed in a pairwise manner between each of the four populations and enriched functional gene categories were determined in both up- and down-regulated genes from each dataset.

ACKNOWLEDGEMENTS

The authors would like to thank Robert Heinz, Amanda Howland, and Clinton Meinhardt for their assistance in nematode cultivation. The authors would also like to thank William Spollen and Christopher Bottoms from the University of Missouri Informatics Research Core Facility for their assistance in providing bioinformatics expertise.

REFERENCES

- Afferson, H., Eleftheriou, E., Selkirk, M., and Gounaris, K. 2012. *Trichinella spiralis* secreted enzymes regulate nucleotide-induced mast cell activation and release of mouse mast cell protease 1. *Infection and Immunity* 80: 3761-3767.
- Arden, S., Smith, A., Booth, M., Tweedie, S., Gounaris, K., and Selkirk, M. 1997. Identification of serine threonine protein kinases secreted by *Trichinella spiralis* infective larvae. *Molecular Biochemistry and Parasitology* 90: 111-119.
- Bekal, S., Niblack, T., and Lambert, K. 2003. A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. *Molecular Plant-Microbe Interactions* 16: 439-446.
- Bekal, S., Domier, L., Niblack, T., and Lambert, K. 2011. Discovery and initial analysis of novel viral genomes in the soybean cyst nematode. *Journal of General Virology* 92: 1870-1879.
- Bekal, S., Domier, L., Gonfa, B., McCoppin, N., Lambert, K., and Bhalerao, K. 2014. A novel flavivirus in the soybean cyst nematode. *Journal of General Virology* 95: 1272-1280.
- Bekal, S., Domier, L.L., Gonfa, B., Lakhssassi, N., Meksem, K., Lambert, K.N. 2015. A SNARE-like protein and biotin are implicated in soybean cyst nematode virulence. *PLoS ONE*: e0145601.
- Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, btu170.
- Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., and Bairoch, A. 2007. UniProtKB/Swiss-Prot. *Methods in Molecular Biology* 406:89-112.
- Brown, D., Robertson, W., and Trudgill, D. 1995. Transmission of viruses by plant nematodes. *Annual Review of Phytopathology* 33: 223-249.
- Buck, A., and Blaxter, M. 2013. Functional diversification of argonautes in nematodes: an expanding universe. *Biochemical Society Transactions* 41: 881-886.
- Canova, M., and Molle, V. 2014. Bacterial serine/threonine protein kinases in host-pathogen interactions. *Journal of Biological Chemistry* 289: 9473-9479.
- De Boer, J.M., Yan, Y., Wang, X., Smant, G., Hussey, R.S., Davis, E.L., and Baum, T.J. 1999. Developmental expression of secretory β -1,4-endoglucanases in the subventral esophageal glands of *Heterodera glycines*. *Molecular Plant-Microbe Interactions* 12: 663-669.

Dong, K., and Opperman, C. 1997. Genetic analysis of parasitism in the soybean cyst nematode *Heterodera glycines*. *Genetics* 146: 1311-1318.

Eves-van den Akker, S., Laetsch, D., Thorpe, P., Lilley, C., Danchin, E., Da Rocha, M., et al. 2016. The genome of the yellow potato cyst nematode, *Globodera rostochiensis*, reveals insights into the basis of parasitism and virulence. *Genome Biology* 17:124.

Félix, M., Ashe, A., Piffaretti, J., Wu, G., Nuez, I., Bélicard, T., et al. 2011. Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to Nodaviruses. *PLoS Biology* 9: e1000586. Doi:10.1371/journal.pbio.1000586.

Finn, R.D., Clements, J., and Eddy, S.R. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research* 39: W29-W37.

Gao, B., Allen, R., Maier, T., Davis, E.R., Baum, T.J., Hussey, R.S. 2003. The parasitome of the phytonematode *Heterodera glycines*. *Molecular Plant-Microbe Interactions* 2003: 720-726.

Gardner, M., Heinz, R., Wang, J., and Mitchum, M.G. 2017. Genetics and adaptation of soybean cyst nematode to broad spectrum soybean resistance. *G3: Genes | Genomes | Genetics* 7:835-841.

Gheysen, G., and Mitchum M.G. 2011. How nematodes manipulate plant development pathways for infection. *Current Opinion in Plant Biology* 14: 415-421.

Grabherr, M.G. et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644-652.

Gounaris, K. 2002. Nucelotidase cascades are catalyzed by secreted proteins of the parasitic nematode *Trichinella spiralis*. *Infection and Immunity* 70: 4917-4924.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P., Bowden, J., et al. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8: 1494-1512.

Hewezi, T., Howe, P., Maier, T., Hussey, R., Mitchum, M., Davis, E., and Baum, T. 2010. Arabidopsis spermidine synthase is targeted by an effector protein of the cyst nematode *Heterodera schachtii*. *Plant Physiology* 152: 968-984.

Ithal, N., Recknor, J., Nettleton, D., Hearne, L., Maier, T., Baum, T., et al. 2007. Parallel genome-wide expression profiling of host and pathogen during soybean cyst nematode infection of soybean. *Molecular Plant-Microbe Interactions* 20: 293-305.

Kim, Y., Riggs, R., and Kim, K. 1987. Structural changes associated with resistance of soybean to *Heterodera glycines*. *Journal of Nematology* 19: 177-187.

Kim, Y., Kim, K., and Riggs, R. 2012. Initial subcellular responses of susceptible and resistant soybeans infected with the soybean cyst nematode. *Plant Pathology Journal* 28: 401-408.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* 305: 567-580.

Kuang, Z., Lewis, R., Curtis, J., Zhan, Y., Saunders, B., Babon, J., et al. 2010. The SPRY domain-containing SOCS box protein SPSB2 targets iNOS for proteasomal degradation. *Journal of Cell Biology* 190: 129.

Langmead, B., and Salzberg, S. 2012. Fast gapped-read alignment with Bowtie2. *Nature Methods* 9: 357-359.

Li, B., and Dewey, C. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.

Liu, S., Kandath, P.K., Warren, S.D., Yeckel, G., Heinz, R., Alden, J., et al. 2012. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492: 256-260.

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

Magrane, M. et al. 2011. UniProt knowledgebase: a hub of integrated protein data. *Database(Oxford)*: bar009.

Mitchum, M. 2016. Soybean resistance to the soybean cyst nematode *Heterodera glycines*: an update. *Phytopathology* 106: 1444-1450.

Niblack, T.L., Arelli, P.R., Noel, G.R., Opperman, C.H., Orf, J.H., Schmitt, D.P., et al. 2002. A revised classification scheme for genetically diverse populations of *Heterodera glycines*. *Journal of Nematology* 34: 279-288.

Niblack, T., Colgrove, A., Colgrove, K., Bond, J. 2008. Shift in virulence of soybean cyst nematode is associated with use of resistance from PI88788. *Plant Health Progress* doi:10.1094/PHP-2008-0118-01-RS.

Neuhaus, B., Bresciani, J., and Peters, W. 1997. Ultrastructure of the pharyngeal cuticle and lectin labeling with wheat germ agglutinin-gold conjugate indicating chitin in the pharyngeal cuticle of *Oesophagostomum dentatum*. *Acta Zoologica* 78: 205-213.

Noon, J.B., Hewezi, T., Maier, T.R., Simmons, C., Wei, J.Z., Wu, G., et al. 2015. Eighteen new candidate effectors of the phytonematode *Heterodera glycines* produced

specifically in the secretory esophageal gland cells during parasitism. *Phytopathology* 105: 1362-72.

Olsen, A., and Shriver, K. 2003. Ligand mimicry? Plant-parasitic nematode polypeptide with similarity to CLAVATA3. *Trends in Plant Science* 8: 55-57.

Postma, W., Slootweg, E., Rehman, S., Finkers-Tomczak, A., Tytgat, T., van Gelderen, K., et al. 2012. The effector SPRYSEC-19 of *Globodera rostochiensis* suppresses CC-NB-LRR-mediated disease resistance in plants. *Plant Physiology* 160: 944-954.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8: 785-786.

Punta, M. et al. 2012. The Pfam protein families database. *Nucleic Acids Research* 40: D290-301.

Riggs, R., Kim, K., and Gipson, I. 1973. Ultrastructural changes in Peking soybeans infected with *Heterodera glycines*. *Phytopathology* 63: 76-84.

Robertson, L., Robertson, W.M., Sobczak, M., Helder, J., Tetaud, E., Ariyanyagam, M.R., et al. 2000. Cloning, expression and functional characterization of a preoxidoreductase from the potato cyst nematode *Globodera rostochiensis*. *Molecular and Biochemical Parasitology* 111: 41-49.

Sacco, M., Koropacka, K., Grenier, E., Jaubert, M., Blanchard, A., Goverse, A., et al. 2009. The cyst nematode SPRYSEC protein RBP-1 elicits Gpa2- and RanGAP2-dependent plant cell death. *PLoS Pathogens* 5: e1000564.

Simão, F., Waterhouse, R., Ioannidis, P., Kriventseva, E., and Zdobnov, E. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31: 3210-3212.

Turek, M., and Bringmann, H. 2014. Gene expression changes of *Caenorhabditis elegans* larvae during molting and sleep-like lethargus. *PLoS ONE* 9: e113269.

Veronico, P., Gray, L., Jones, J., Bazzicalupo, P., Arbucci, S., Cortese, M., et al. 2001. Nematode chitin synthetases: gene structure, expression and function in *Caenorhabditis elegans* and the plant parasitic nematode *Meloidogyne artiellia*. *Molecular Genetics and Genomics* 266: 28-34.

Wang, S., Hawkins, C., Yoo, S., Müller, H., and Hay, B. 1999. The *Drosophila* caspase inhibitor DIAP1 is essential for cell survival and is negatively regulated by HID. *Cell* 98: 453-463.

Wang, X., Meyers, D., Yan, Y., Baum, T., Smant, G., Hussey, R., and Davis, E. 1999. *In planta* localization of a β -1,4-endoglucanase secreted by *Heterodera glycines*. *Molecular Plant Microbe Interactions* 12: 64-67.

Wang, X., Allen, R., Ding, X., Goellner, M., Maier, T., de Boer, J., et al. 2001. Signal peptide-selection of cDNA cloned directly from the esophageal gland cells of the soybean cyst nematode *Heterodera glycines*. *Molecular Plant-Microbe Interactions* 14: 536-544.

Wang, X., Mitchum, M., Gao, B., Li, C., Diab, H., Baum, T., et al. 2005. A parasitism gene from a plant-parasitic nematode with function similar to CLAVATA3/ESR (CLE) of *Arabidopsis thaliana*. *Molecular Plant Pathology* 6: 187-191.

Young, M., Wakefield, M., Smyth, G., and Oshlack, A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* 11: R14.

CHAPTER 5

SUMMARY AND FUTURE DIRECTIONS

Soybean cyst nematode virulence is the predominant problem faced when generating management plans to control this pathogen. This problem has been exacerbated by repeated use of a few sources of resistance without knowledge of how this immense selection pressure affects the existing SCN populations. Prior research identified putative virulence genes on PI 88788-type resistance but the function of these genes and the identity of the genes for virulence on Peking-type and PI 437654-type resistance was unknown (Dong and Opperman, 1997). In this dissertation the genetics underlying SCN virulence were further examined, leading to a better understanding of the heritability of SCN virulence on Peking-, PI 88788-, and PI 437654-type resistances. Through a controlled crossing strategy it was discovered that virulence on PI 437654-type resistance, the most broad spectrum resistance to SCN available, is most likely a result of a combination of genes and not a single recessive or dominant gene. In addition, interactions between virulence on the different SCN resistance types were examined and counter-selection against virulence was found to exist between the soybean cultivars PI 88788 and PI 90763. The knowledge of this counter-selection can be used to improve management strategies, allowing for a better rotation of resistant varieties to prevent the rise of highly virulent nematodes. The observation that virulence may not decrease in the absence of host resistance may also impact the current rotation schemes. These new rotation schemes will need to be tested in field trials prior to their use on a commercial

scale, but would still be a fairly immediate method for improvement of SCN management.

Effectors represent one of the most likely groups of candidates to lead to SCN virulence. They function in many capacities for the nematode, ranging from assisting in movement to establishment and maintenance of the syncytium and suppression of host stress and defense responses (Gardner et al., 2015; Gheysen and Mitchum, 2011). Previously multiple studies identified candidate effectors for SCN, but were focused on secretions originating in the gland cells or were dependent on the presence of an N-terminal signal peptide (Gao et al., 2003; Noon et al., 2013). In this dissertation a *de novo* transcriptome assembly of the early parasitic stage was generated, the most complete genetic reference currently available for this organism. Using this assembly effector prediction pipelines were established and used to predict a whole new pool of effector candidates, some of which may be involved in virulence. The use of the tool N-Prefector allowed for identification of new putative effectors, some of which may not be secreted through traditional pathways. Evidence exists for the secretion of nematode effectors lacking an N-terminal secretion signal, but the mechanism by which they are secreted remains unknown (Robertson et al., 2000). These effector candidates will need to be localized within the nematode and then confirmed to be secreted before they can be considered true effectors. Then they can be functionally characterized and any role they play in virulence identified.

In other nematode species the difference between a nematode being virulent or avirulent can come down to a single gene. For example, in the potato cyst nematode *Globodera pallida* a secreted SPRYSEC effector is recognized by a host receptor and

triggers resistance (Postma et al., 2012; Sacco et al., 2009). Resistance to SCN and subsequently virulence seems a bit more complicated. The only resistance genes identified to SCN are a serine hydroxymethyltransferase (SHMT) and a three gene cassette including an amino acid transporter, an N-ethylmaleimide-sensitive factor attachment protein (α -SNAP), and a wound-inducible protein (WI12) (Cook et al., 2012; Liu et al., 2012). Virulence mechanisms may involve direct interactions with these proteins or instead manipulate other proteins or pathways within the host cell that compromise the resistance. Recent studies have identified virulence gene candidates such as *HgSLP-1* and *HgBioB*, but no functional evidence exists that these genes are more than correlated with virulence (Bekal et al., 2015). In this dissertation additional sequences were identified that may correlate with virulence, but need to be further characterized in order to validate this correlation, including functional characterization to determine if they play a direct role in virulence. Multiple virulent populations were also characterized in the parasitic second-stage juvenile time point, identifying conserved gene expression between virulence on all three types of SCN resistance. This is, however, just the first step. Follow up from this study should also explore variation in effectors between virulent populations. This could be the result of several possible mechanisms, including alternative splicing and allelic variation. By pursuing where these virulence mechanisms overlap and what they target within the plant it is possible that in the future new broad-spectrum resistance may be engineered.

Future work on SCN virulence will need to focus on using the tools generated here to build on and generate greater genetic resources that can be implemented in improving resistance. The transcriptome data provided here is currently being used to

annotate the eventual genome sequence, improving the genome to allow for comparative analysis both within SCN as well as to other species of plant-parasitic nematodes. By identifying parallels in other plant-parasitic nematodes we can potentially identify conserved pathways targeted by multiple nematodes. These pathways can then be targeted for engineering novel, multi-species resistance.

REFERENCES

- Bekal, S., Domier, L.L., Gonfa, B., Lakhssassi, N., Meksem, K., Lambert, K.N. 2015. A SNARE-like protein and biotin are implicated in soybean cyst nematode virulence. PLoS ONE: e0145601.
- Dong, K., and Opperman, C. 1997. Genetic analysis of parasitism in the soybean cyst nematode *Heterodera glycines*. Genetics 146: 1311-1318.
- Gardner, M., Verma, A., and Mitchum, M. 2015. Emerging roles of cyst nematode effectors in exploiting plant cellular processes. In: Escobar, C., and Fenoll, C. (eds) Advances in Botanical Research Vol. 73, Plant Nematode Interactions: A View on Compatible Interrelationships. Elsevier Publishing, Oxford, p. 259-291.
- Gao, B., Allen, R., Maier, T., Davis, E.R., Baum, T.J., Hussey, R.S. 2003. The parasitome of the phytonematode *Heterodera glycines*. Molecular Plant-Microbe Interactions 16: 720-726.
- Gheysen, G., and Mitchum M.G. 2011. How nematodes manipulate plant development pathways for infection. Current Opinion in Plant Biology 14: 415-421.
- Noon, J.B., Hewezi, T., Maier, T.R., Simmons, C., Wei, J.Z., Wu, G., et al. 2015. Eighteen new candidate effectors of the phytonematode *Heterodera glycines* produced specifically in the secretory esophageal gland cells during parasitism. Phytopathology 105: 1362-72.
- Postma, W., Sloopweg, E., Rehman, S., Finkers-Tomczak, A., Tytgat, T., van Gelderen, K., et al. 2012. The effector SPRYSEC-19 of *Globodera rostochiensis* suppresses CC-NB-LRR-mediated disease resistance in plants. Plant Physiology 160: 944-954.
- Robertson, L., Robertson, W., Sobczak, M., Helder, J., Tetaud, E., Ariyanayagam, M., et al. 2000. Cloning, expression and functional characterization of a preoxidoredoxin from the potato cyst nematode *Globodera rostochiensis*. Molecular and Biochemical Parasitology 111: 41-49.
- Sacco, M., Koropacka, K., Grenier, E., Jaubert, M., Blanchard, A., Goverse, A., et al. 2009. The cyst nematode SPRYSEC protein RBP-1 elicits Gpa2- and RanGAP2-dependent plant cell death. PLoS Pathogens 5:e1000564.

APPENDIX

APPENDIX 1 – Summary of transcript annotation for the SCN transcriptome

The files associated with this table are archived at Mitchum Lab Share/Individual Lab Folders/Michael Gardner/Dissertation and Defense Seminar/Appendixes/.

APPENDIX 2 – Orthologs identified in other nematode genomes

The files associated with this table are archived at Mitchum Lab Share/Individual Lab Folders/Michael Gardner/Dissertation and Defense Seminar/Appendixes/.

APPENDIX 3 – Predicted novel effector candidates from the SCN transcriptome

The files associated with this table are archived at Mitchum Lab Share/Individual Lab Folders/Michael Gardner/Dissertation and Defense Seminar/Appendixes/.

APPENDIX 4 – GO term enrichment analysis between virulent and avirulent populations of SCN

The files associated with this table are archived at Mitchum Lab Share/Individual Lab Folders/Michael Gardner/Dissertation and Defense Seminar/Appendixes/.

VITA

Michael N. Gardner was born on November 17, 1988 in Milwaukee, Wisconsin. He completed his secondary education at Homestead High School in 2007. This education continued at the University of Wisconsin – Madison where he would earn Bachelors of Science degrees in Botany and Plant Pathology in 2011. Following graduation, he began graduate school at the University of Missouri on a Life Sciences fellowship, joining the lab of Dr. Melissa G. Mitchum in the spring of 2012. In the Mitchum lab his work focused on nematode virulence and improving genetic resources available for the soybean cyst nematode.