

# Active and inactive forms of biotin synthase occur in *Heterodera glycines*

Khee Man Kwon,<sup>1,4\*</sup> Sadia Bekal,<sup>2</sup>  
Leslie L. Domier<sup>1,3</sup> and  
Kris N. Lambert<sup>1</sup>

<sup>1</sup>Department of Crop Sciences,  
University of Illinois, Urbana, IL.

<sup>2</sup>Department of Agricultural and  
Biological Engineering, University of  
Illinois, Urbana, IL.

<sup>3</sup>United States Department of  
Agriculture – Agricultural Research  
Service, Urbana, IL.

<sup>4</sup>Department of Plant Pathology  
and Center for Applied Genetic  
Technologies, University of  
Georgia, Athens, GA.

\*E-mail: kkwon@uga.edu

This paper was edited by  
Axel Elling.

Received for publication April 29,  
2019.

## Abstract

*Heterodera glycines*, the soybean cyst nematode (SCN), is a plant-parasitic nematode capable of manipulating host plant biochemistry and development. Many studies have suggested that the nematode has acquired genes from bacteria via horizontal gene transfer events (HGTs) that have the potential to enhance nematode parasitism. A recent allelic imbalance analysis identified two candidate virulence genes, which also appear to have entered the SCN genome through HGTs. One of the candidate genes, *H. glycines* biotin synthase (*HgBioB*), contained sequence polymorphisms between avirulent and virulent inbred SCN strains. To test the function of these *HgBioB* alleles, a complementation experiment using biotin synthase-deficient *Escherichia coli* was conducted. Here, we report that avirulent nematodes produce an active biotin synthase while virulent ones contain an inactive form of the enzyme. Moreover, sequencing analysis of *HgBioB* genes from SCN field populations indicates the presence of diverse mixture of *HgBioB* alleles with the virulent form being the most prevalent. We hypothesize that the mutations in the inactive *HgBioB* allele within the virulent SCN could result in a change in protein function that in some unknown way bolster its parasitic lifestyle.

## Keywords

Biochemistry, Biotin, Biotin synthase, Dethiobiotin, *Heterodera glycines*, Horizontal gene transfer, Soybean cyst nematode, Vitamin B<sub>7</sub>.

The soybean cyst nematode (SCN), *Heterodera glycines*, is an obligate plant parasite that poses a serious threat to soybean production worldwide. It is the most damaging pest of soybean in the USA and is estimated to cause more yield loss than any other disease (Koenning and Wrather, 2010; Wrather *et al.*, 2010; Allen *et al.*, 2017). Yield losses are attributed to the fact that this nematode injects a complex mixture of proteins and metabolites into the plant root cells (Niblack *et al.*, 2006) and transforms them into a metabolically active feeding structure called a syncytium (Hussey, 1989). The syncytium becomes a nutrient sink where the nematode feeds throughout its life cycle (Jones, 1981; Endo, 1998). This disrupts the plant's vascular system and prevents normal functioning of the root system leading to lower soybean yields.

What makes matters worse is the nematode's ability to persist in soils for 10 or more years even in the absence of a soybean host (Riggs, 2004). The most common and practical control strategy for SCN is to plant resistant cultivars and rotate to non-host crops (Niblack *et al.*, 2006). Although the management of SCN has been initially successful thanks to the use of resistant cultivars, the repeated use of the same resistance source has selected for virulent nematode populations that can reproduce on resistant cultivars (Niblack and Riggs, 2004; Mitchum *et al.*, 2007; Niblack *et al.*, 2008; Mitchum, 2016; McCarville *et al.*, 2017). Understanding how SCN evade or suppress host plant resistance at the molecular level may allow the development of more effective sources of SCN-resistant soybean.

To this end, tremendous effort was put in the past decades to understand the complex interactions between the nematode and its host plant. The majority of studies on plant-parasitic nematode-host interactions were mainly focused on discovering genes that code for secreted effector proteins originating from the esophageal glands involved in manipulating host plant biochemistry, development, and defenses (Hussey, 1989; Smant *et al.*, 1998; Lambert *et al.*, 1999; Davis *et al.*, 2000). Not all nematode secretory proteins, however, originate from the esophageal glands and some of the genes that encode secretory proteins can act as either virulence or avirulence genes. An esophageal gland-secreted protein/enzyme chorismate mutase (CM) is thought to suppress host defense by lowering the levels of chorismate-derived compounds required for plant defense (Lambert *et al.*, 1999; Doyle and Lambert, 2003). Some SCN CM genes contain polymorphisms correlating with virulence on resistant soybean cultivars (Bekal *et al.*, 2003; Lambert *et al.*, 2005), also implicating these genes in SCN virulence. Similarly, a cuticle-secreted lipid-binding protein Gp-FAR-1 from *Globodera pallida* may be able to inhibit the host's jasmonic acid signaling pathway (Prior *et al.*, 2001). In contrast, the SPRYSEC effector protein Gp-RBP-1 secreted from an esophageal gland in *G. pallida* (Sacco *et al.*, 2009), the venom allergen-like effector protein Gr-VAP1 secreted from an esophageal gland in *G. rostochiensis* (Lozano-Torres *et al.*, 2012), MAP-1 secreted from amphids in *Meloidogyne incognita* (Semblat *et al.*, 2001), and a SNARE-like protein (*HgSLP-1*) secreted from an esophageal gland in *H. glycines* (Bekal *et al.*, 2015) appear to act as avirulence genes. Although not a secreted protein, Cg-1 from *M. javanica* is also thought to function as an avirulence gene (Gleason *et al.*, 2008).

Some of the secreted proteins are closely related to bacterial proteins, suggesting that they have been acquired from bacteria through horizontal gene transfer events (HGTs) (Smant *et al.*, 1998; Lambert *et al.*, 1999; Davis *et al.*, 2000; Bird *et al.*, 2003; Bekal *et al.*, 2015). While these studies focusing on secreted proteins and HGT candidates have been productive for understanding the molecular basis of nematode-host interactions, non-secreted proteins, enzymes, and metabolites have been overlooked; hence, horizontally transferred genes coding for non-secreted proteins/enzymes with bacterial homology could also play a critical role in SCN biology, parasitism, and (a)virulence.

Studies on non-secreted plant-parasitic nematode proteins have been few in number until Craig *et al.* (2008, 2009) showed that a certain group of non-secreted HGT proteins can include enzymes involved

in vitamin B<sub>1</sub>, B<sub>5</sub>, B<sub>6</sub>, and B<sub>7</sub> biosynthesis and their corresponding salvage pathways. More recently, a whole genome allelic imbalance analysis, conducted for the purpose of identifying single nucleotide polymorphisms (SNPs) associated with SCN genes implicated in virulence, found two candidate virulence genes, *HgSLP-1* and biotin synthase (*HgBioB*), both of which have been acquired through HGTs (Bekal *et al.*, 2015). *HgBioB* is a non-secreted protein/enzyme responsible for catalyzing the conversion of dethiobiotin (dethiobiotin or DTB) to biotin (vitamin B<sub>7</sub>). Furthermore, *HgBioB* contained SNPs that generated amino acid polymorphisms between avirulent and virulent inbred SCN strains, suggesting the allelic forms of *HgBioB* may have different activity (Bekal *et al.*, 2015).

We postulated that these sequence polymorphisms could alter *HgBioB* activity and play an important role in SCN parasitism or virulence. Because it seemed reasonable that a nematode's ability to produce more biotin might counteract a plant resistance mechanism that restricts the nematode from accessing free biotin or its immediate precursor, we hypothesized that virulent nematodes, which can reproduce on resistant cultivars, might produce a more active form of *HgBioB* than their avirulent counterparts, which cannot parasitize the resistant hosts. To support or refute this hypothesis, we tested the function of the two *HgBioB* alleles to ascertain if any functional differences could be detected between the virulent and avirulent forms of the enzyme.

## Materials and methods

### Construction and cloning of recombinant plasmids containing *HgBioB*

Two different DNA fragments *HgBioB*-avr (avirulent; accession number MH251885) and *HgBioB*-vir (virulent; accession number MH251886) that code for *HgBioB* in inbred SCN strains TN10 (avirulent; Hg-type 0; grown on Roma tomato) and TN20 (virulent; Hg-type 1, 2, 3, 4, 5, 6, 7; grown on plant introduction 437654), respectively, were synthesized with codons optimized for *E. coli* expression (Life Technologies, Carlsbad, CA). Also, two different mutagenized forms of *HgBioB*-vir, *HgBioB*-vir-M1 (virulent, mutagenized allele 1; accession number MH251887) and *HgBioB*-vir-M2 (virulent, mutagenized allele 2; accession number MH251888), were prepared by generating (i) an alanine to proline mutation (M1) at amino acid position 24 (A24P); and (ii) a glutamine to arginine mutation (M2) at amino acid position 44 (Q44R), respectively. M1 was produced via site-directed

mutagenesis following the manufacturer's instructions for In-Fusion cloning (Clontech, Mountain View, CA), using *HgBioB-vir* as the template. A pair of inverse polymerase chain reaction (PCR) primers (Table 1), BioB-mut1-F and BioB-mut1-R, was designed to share a 15-bp homology with each other and include a single-base substitution to generate an A24P mutation. Unlike M1, M2 was prepared by custom-ordering the synthetic DNA fragment containing the desired Q44R mutation (Life Technologies).

To perform In-Fusion cloning (Clontech), two pairs of PCR primers (Table 1) were designed for the insert and the vector so that they would share a 15-bp homology with each other. All four DNA fragments were amplified using the PCR primers BioB-F and BioB-R; and pBAD.LIC.8A cloning vector (Addgene plasmid No. 37501) was amplified with PCR primers BioB-Vec-F and BioB-Vec-R. All PCRs were done using an iProof HF Mastermix (Bio-Rad, Hercules, CA) following the manufacturer's recommended PCR conditions. The resulting PCR products were gel-purified using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA), joined together using an In-Fusion HD Cloning Kit (Clontech), and transformed

into Stellar Competent Cells (Clontech) following the manufacturer's instructions. Cells containing successful transformants were screened on Luria-Bertani (LB) agar plates with ampicillin (100mg/ml) and a single colony was picked to inoculate 5ml of LB broth with ampicillin (100mg/ml) which was incubated overnight at 37°C. Recombinant plasmid DNA was extracted from the overnight culture using a QIAprep Spin Miniprep Kit (Qiagen) and the resulting plasmids were sequenced for verification at the University of Illinois Roy J. Carver Biotechnology Center.

### Complementation of $\Delta$ bioB *Escherichia coli* with avirulent and virulent alleles of *HgBioB*

A kanamycin-resistant *E. coli* vitamin B<sub>7</sub> auxotroph (K-12 Keio Collection: JW0758-1) was obtained from the Coli Genetic Stock Center at Yale University and used to prepare electrocompetent cells as described by Dower *et al.* (1988) and Hanahan (1983). Four different recombinant plasmids – pBAD.LIC.8A:: *HgBioB-avr*, pBAD.LIC.8A:: *HgBioB-vir*, pBAD.LIC.8A:: *HgBioB-vir-M1*, and pBAD.LIC.8A:: *HgBioB-vir-M2* – were generated and independently transformed into the kanamycin-resistant  $\Delta$ bioB mutant electrocompetent cells. For control purposes, an empty vector (pBAD.LIC.8A) lacking any inserts was also transformed into the electrocompetent cells. The resulting transformants were screened on LB agar plates with ampicillin (100mg/ml) and kanamycin (50mg/ml), and a single colony was picked to inoculate 10ml of LB broth with ampicillin (100mg/ml) and kanamycin (50mg/ml) for overnight incubation at 37°C. The overnight cultures of these four strains were thoroughly washed in liquid M9 minimal media to remove trace amounts of biotin that may be present, and then streaked onto M9 minimal media plates containing ampicillin (100mg/ml), kanamycin (50mg/ml), and either avidin or biotin. Plates containing avidin (1mg/ml) were prepared to sequester trace amounts of biotin in the M9 media. Positive control plates were M9 media supplemented with biotin to a final concentration of 0.2mM, and a negative control plate consisted of  $\Delta$ bioB *E. coli* transformed with an empty pBAD.LIC.8A vector plated on M9 media lacking biotin. Some plates were also treated with 2% arabinose to induce *HgBioB* expression and monitor its potential effect on *E. coli* growth. All plates were initially incubated at 28°C for 48hr, and then allowed to grow at room temperature for 10d. *E. coli* growth was monitored daily.

**Table 1. Oligonucleotides for plasmid construction and sequencing analysis.**

Primer name	Sequence
BioB-F	5'-GAA GGA GAT ATA GAT ATG CCT CCG CCT ATT GGT AGC-3'
BioB-R	5'-TTA TGG AGT TGG GAT TTA CAG ATT CAG GGT CAC TTT TTC ATC GT-3'
BioB-mut1-F	5'-CCA TTT CCG GAA CTG ATT TTT CGT GCA CAG AAT GTT C-3'
BioB-mut1-R	5'-CAG TTC CGG AAA TGG CAG GCT AAA AAC GCT CAG TGC-3'
BioB-Vec-F	5'-ATC CCA ACT CCA TAA GGA TCC CTT G-3'
BioB-Vec-R	5'-ATC TAT ATC TCC TTC TTA AAG TTA AAC AAA ATT ATT TCT AGA TGT AGA TGT TAG CC-3'
BioB-flank-F1a	5'-GGA GAG GAA TGA TAT GAT GAA-3'
BioB-flank-R1a	5'-CAT CTT CTG CTT CTG TTC TG-3'

## Nematode culture and DNA extraction of SCN field populations

A collection of soil samples containing diverse SCN field populations representative of the *H. glycines* diversity was randomly chosen and prepared for nematode culture. The majority of these soil samples was from Iowa, a kind gift from Dr Gregory Tylka at Iowa State University (Ames, IA), and some of them were from Illinois. At least two to three different field populations originating from the same geographical region (e.g., at least two to three field populations from the east central region) were randomly selected for cyst extraction. Cysts from each field population were collected as described by Niblack *et al.* (1993) and crushed over a 0.25-mm (60 mesh) sieve to release eggs. Eggs were suspended in 2% carboxymethyl cellulose solution for an easy delivery of an approximately equal number of eggs during inoculation. This egg suspension was used to inoculate soybean cv. Essex plants, previously pre-germinated and planted into 50-ml Falcon tubes with the absorbent wick described in Bekal *et al.* (2015). Inoculated plants were grown for 8wk to allow eggs to hatch and infect the plants. The resulting cysts from each field population, approximately 2 to 200 cysts from each plant, were harvested in a 2-ml cryotube vial, and stored at  $-80^{\circ}\text{C}$  until use. The cyst-containing vials were placed in a pre-cooled rack in liquid nitrogen and two stainless steel beads (3.2 mm diameter) were added into each vial. These frozen cysts were then freeze-fractured with a bead beater homogenizer and total nucleic acid was extracted using an RNeasy Mini Kit (Qiagen).

## Sequencing and statistical analysis of different SCN populations

In addition to 30 DNA samples from field populations, 4 DNA samples from inbred SCN strains – 1 avirulent (OP25) and 3 virulent inbred strains (OP20, OP50, and TN20) – were also included in the analysis. The OP strains were previously maintained as described in Dong and Opperman (1997) by Dr Terry Niblack (Ohio State University, Columbus, OH), and TN20 was maintained on PI 437654. From each extracted 50- $\mu\text{l}$  DNA sample from a single population, 2.5  $\mu\text{l}$  was used as a template for amplification using a GenomiPhi Kit (GE Healthcare, Piscataway, NJ) following the manufacturer's protocols. The amplified DNA was treated with ExoSAP-IT (Affymetrix, Santa Clara, CA) for the degradation of primers and nucleotides that may interfere with downstream applications. A set of PCR primers (Table 1), BioB-flank-F1a and BioB-flank-R1a,

flanking the polymorphic region of the *HgBioB* gene was designed for PCR amplification and subsequent sequencing analysis. All PCRs were done using a CloneAmp Hi-Fi PCR Premix (Clontech) following the manufacturer's recommended conditions. The resulting PCR product was gel-purified using a QIAquick Gel Extraction Kit (Qiagen) and then sent for sequencing at the University of Illinois Roy J. Carver Biotechnology Center. The sequences were aligned to the original *HgBioB* sequence using Sequencher (Gene Codes Corporation, Ann Arbor, MI), and the relative chromatogram peak heights corresponding to amino acid positions 24 and 44 were measured with a 15-cm ruler on the computer screen to estimate the percentage of each polymorphic nucleotide(s).

Peak heights were statistically analyzed using both descriptive and inferential methods. First, a box plot was generated to compare the median values and understand the distributional characteristics of these alleles at both polymorphic regions, and then a Wilcoxon signed rank test was performed using the SAS Proc Univariate procedure (SAS Institute, Cary, NC) to test if the differences between avirulent and virulent alleles were statistically significant. For the signed rank test, a new variable 'diff' was set as the difference between the virulent and avirulent alleles ( $\text{diff} = \text{vir} - \text{avir}$ ) at each amino acid for each SCN sample prior to running the Proc Univariate procedure on that variable. Finally, a Spearman rank-order correlation was conducted using the SAS Proc Corr procedure (SAS Institute) to calculate the coefficients for the relationship between the two variables, amino acid positions 24 and 44, for each set of virulent and avirulent pairs.

## Results

### Complementation of $\Delta\text{bioB}$ *E. coli* with avirulent and virulent alleles of *HgBioB*

We had observed differences in the predicted amino acid sequences between avirulent and virulent *HgBioB* alleles (Bekal *et al.*, 2015), and thus hypothesized that such differences could lead to an altered enzyme function. To test for *HgBioB* function, we opted to use genetic complementation of a BioB-deficient *E. coli* mutant. Specifically, we wanted to learn if amino acid(s) changes, P24A or R44Q, the only two amino acids that are different between the *HgBioB* alleles, would alter BioB enzyme function.

To conduct this experiment, four plasmid constructs were generated to express the different forms of *HgBioB*; the virulent allele (*HgBioB*-vir, A24/Q44), the avirulent allele (*HgBioB*-avr, P24/R44), and

the two single mutant alleles (*HgBioB*-vir-M1, P24/Q44; and *HgBioB*-vir-M2, A24/R44) (Fig. 1A). These two single mutants prepared from the virulent allele were included in the complementation studies to identify which amino acid location(s) were responsible for a functional enzyme activity. These plasmids, along with an empty vector control, were transformed into mutant *E. coli* lacking endogenous BioB function and were then allowed to grow on minimal agar with and without supplemented biotin.

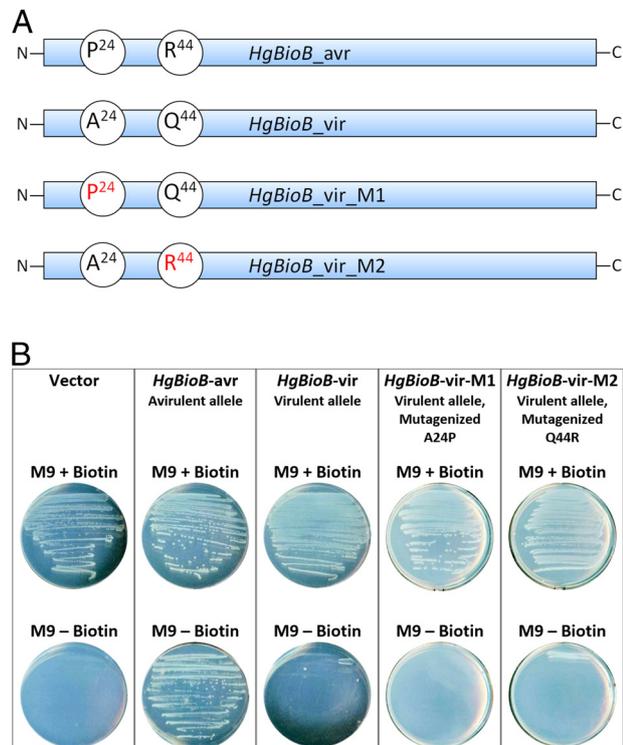


Figure 1: Complementation by expression of *HgBioB* in biotin synthase-deficient *E. coli* strain. (A) *HgBioB* construct diagrams showing polymorphic amino acids, from N-terminus to C-terminus, at positions 24 and 44. *HgBioB*-vir was mutagenized to produce *HgBioB*-vir-M1, by generating an alanine to proline mutation at position 24 (A24P); and *HgBioB*-vir-M2, by generating a glutamine to arginine mutation at position 44 (Q44R). (B) The expression of the vector control, *HgBioB*-avr, *HgBioB*-vir, and two mutagenized forms of the virulent allele *HgBioB*-vir-M1 and *HgBioB*-vir-M2 on minimal media with and without supplemental biotin.

Within 48hr after streaking the *E. coli* containing the BioB constructs onto minimal media plates, all five biotin-supplemented plates (minimal media +biotin; positive controls) showed vigorous growth of *E. coli* as expected. Among the five plates lacking biotin (minimal media alone), however, growth only occurred for the bacteria containing the avirulent allele containing the P24/R44 amino acid combination (*HgBioB*-avr) (Fig. 1B). The other four constructs, *E. coli* carrying the empty vector (negative control), the virulent allele (*HgBioB*-vir), and the two mutagenized forms of the virulent allele (*HgBioB*-vir-M1 and *HgBioB*-vir-M2) failed to complement the growth of the mutant *E. coli* (Fig. 1B).

Our complementation test showed that avirulent SCN produces an active biotin synthase while virulent SCN contains an inactive form of the enzyme. Additionally, site-directed mutagenesis experiments confirmed that changing either of these two amino acids caused the *HgBioB* protein to become inactive, losing its ability to complement the mutant; therefore, a proline at position 24 (P24) and an arginine at position 44 (R44) are required for a functional BioB activity (Fig. 1A,B). This experiment was repeated and produced the same results.

### Sequencing, statistical analysis, and confirmation of SNPs from SCN field populations

Since our previous studies were based on just two inbred SCN strains, we were interested to learn if sequence variation in the *HgBioB* gene was common in different field populations of SCN. To assess how common *HgBioB*-vir alleles are in different populations of SCN, the *HgBioB* gene was PCR-amplified and sequenced from a series of SCN populations. Sequencing analysis of four inbred SCN strains (OP25, OP20, OP50, and TN20) showed that these strains were more likely to be uniform (i.e., have 100% allelic frequencies) carrying single homozygous forms (Table 2). In contrast, sequencing results from 30 SCN field populations showed that in most SCN populations, the *HgBioB* gene was not present in a single homozygous form, but rather in a diverse mixture consisting of both SNPs resulting in both active and inactive forms of *HgBioB* (Table 2). Statistical analysis of the field sequencing data was conducted by generating a box plot and comparing the median values. This analysis showed that the *HgBioB*-vir alleles were present in a higher percentage than the *HgBioB*-avr alleles at both polymorphic regions (i.e., more A24 than P24; more Q44 than R44) (Fig. 2A,B, Table 3). This trend was

Table 2. Sequencing analysis of *HgBioB* nematode populations showing the percentage of amino acids at each polymorphic region.

SCN populations	24th Amino acid			44th Amino acid		
	Threonine (ACT) %A	Alanine (GCT) %G	Proline (CCT) %C	Glutamine (CAA) %A	Arginine (CGA) %G	Proline (CCA) %C
Hg257		<b>100.00</b>		<b>100.00</b>		
Hg0-A		60.00	40.00	53.33	40.00	
Hg0-B	33.33	66.67		<b>100.00</b>		
OP 25 <sup>a</sup>			<b>100.00</b>		<b>100.00</b>	
OP 20 <sup>a</sup>			<b>100.00</b>			<b>100.00</b>
OP 50 <sup>a</sup>	33.33		66.67			<b>100.00</b>
TN 20 <sup>a</sup>		<b>100.00</b>		64.29	21.42	
Carbondale 6 <sup>b</sup>		75.00	25.00	75.00	25.00	
Carbondale 30 <sup>b</sup>			<b>100.00</b>		<b>100.00</b>	
C 104 <sup>c</sup>	10.00	40.00	50.00	87.50	12.50	
C 432 <sup>c</sup>		83.33	16.67	<b>100.00</b>		
EC 138 <sup>c</sup>		66.67	33.33	83.33	16.67	
EC 146 <sup>c</sup>		66.67	33.33	75.00	25.00	
EC 406 <sup>c</sup>		60.00	40.00	75.00	25.00	
NC 228 <sup>c</sup>		75.00	25.00	87.50		12.50
NC 238 <sup>c</sup>		66.67	33.33	60.00	20.00	20.00
NC 305 <sup>c</sup>		<b>100.00</b>		<b>100.00</b>		
NC 408 <sup>c</sup>		50.00	50.00	60.00	40.00	
NE 336 <sup>c</sup>		83.33	16.67	87.50	12.50	
NE 433 <sup>c</sup>		<b>100.00</b>		<b>100.00</b>		
NW 138 <sup>c</sup>		75.00	25.00	85.71	14.28	
NW 209 <sup>c</sup>		57.14	42.85	66.67	33.33	
NW 308 <sup>c</sup>		75.00	25.00	83.33	16.67	
NW 421 <sup>c</sup>		66.67	33.33	83.33	16.67	
SC 349 <sup>c</sup>		75.00	25.00	83.33	16.67	
SE 132 <sup>c</sup>		60.00	40.00	66.67	33.33	
SE 148 <sup>c</sup>			<b>100.00</b>		<b>100.00</b>	
SE 217 <sup>c</sup>			<b>100.00</b>		<b>100.00</b>	
SW 207 <sup>c</sup>		40.00	60.00	50.00	50.00	
SW 242 <sup>c</sup>		45.45	54.54	60.00	40.00	
SW 251 <sup>c</sup>		85.71	14.28	<b>100.00</b>		
WC 216 <sup>c</sup>			<b>100.00</b>		<b>100.00</b>	
WC 229 <sup>c</sup>		83.33	16.67	90.91	9.09	
WC 350 <sup>c</sup>		90.90	9.09	<b>100.00</b>		

<sup>a</sup>Inbred SCN strains provided by Charlie Opperman (OP) and Terry Niblack (TN); <sup>b</sup>SCN field populations from Carbondale, Illinois; <sup>c</sup>SCN field populations provided by Gregory Tylka. The sampling of these populations was conducted in 2015 from different regions of Iowa. The geographical origins are indicated as follows: C=central; EC=east central; NC=north central; NE=northeast; NW=northwest; SC=south central; SE=southeast; SW=southwest; WC=west central; \* Amino acids with 100 percent allelic frequencies are shown in boldface.

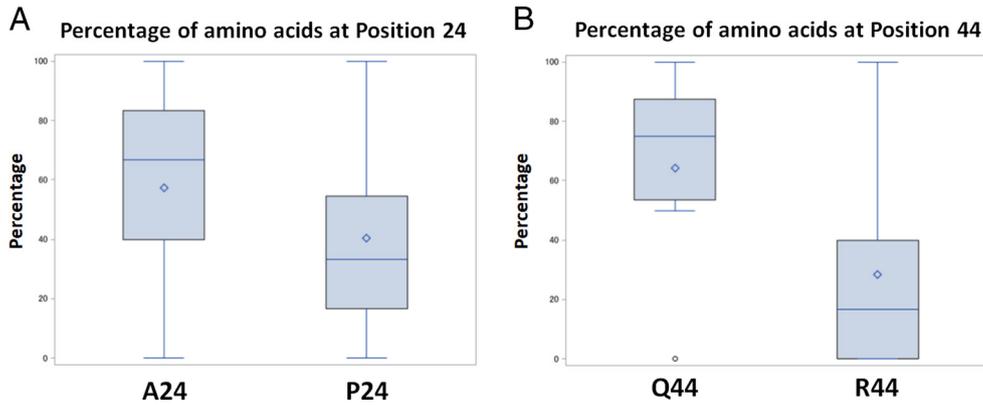


Figure 2: Box plots of *HgBioB* amino acid polymorphism trends among SCN field populations, showing the percentage of amino acid differences for A24P (A) and Q44R (B).

also supported by conducting a Wilcoxon signed rank test, the nonparametric alternative to the paired *t*-test, which was chosen because the polymorphic nucleotide percentage data for avirulent and virulent alleles were not normally distributed. The data were considered paired since we compared the two amino acid forms at each polymorphic region (i.e., A24 vs P24 and Q44 vs R44). The Wilcoxon signed rank test results showed that *HgBioB*-vir allele ranks were significantly higher than *HgBioB*-avr allele ranks for both amino acid positions 24 ( $Z=105$ ,  $p<0.05$ ) and 44 ( $Z=121.5$ ,  $p<0.01$ ). Therefore, from both descriptive and inferential statistical results, we concluded that the virulent alleles are more prevalent than the avirulent alleles in the field sites tested.

In addition, it is interesting to note that these SNP changes occur simultaneously in pairs, at about the same frequency, between amino acid positions 24 and 44. For example, if a SNP in *HgBioB* results in a P24 then it would be paired with an R44; if it were

an A24 then it would be paired with a Q44. This linkage of SNP frequency was supported by testing the correlations between the virulent allele pair (A24 and Q44) as well as the avirulent allele pair (P24 and R44) using a Spearman rank-order correlation, the nonparametric alternative to the Pearson product-moment correlation, chosen because the assumptions of normality could not be met. The Spearman correlation results showed that there was a positive correlation for both virulent ( $r=0.8683$ ,  $n=30$ ,  $p<0.001$ ) and avirulent ( $r=0.9184$ ,  $n=30$ ,  $p<0.001$ ) pairs. Scatterplots for both avirulent and virulent pairs also summarize that there is a positive linear relationship between amino acid positions 24 and 44 (Fig. 3A,B). The strong, positive correlation between these positions suggests that these amino acid changes must both occur in the same protein to exert their desired, but unknown function. Although a rarity, alternative SNPs that result in a threonine at position 24 (T24) and a proline at position 44 (P44)

Table 3. Statistical analysis of amino acid percentages from nematode field populations.

	Amino acid position 24		Amino acid position 44	
	Alanine (A24)	Proline (P24)	Glutamine (Q44)	Arginine (R44)
Mean	57.28	40.46	64.07	28.47
SD	33.51	32.88	36.02	33.21
Median	66.67	33.33	75.00	16.67
Min	40.00	9.09	50.00	9.09
Max	100.00	100.00	100.00	100.00

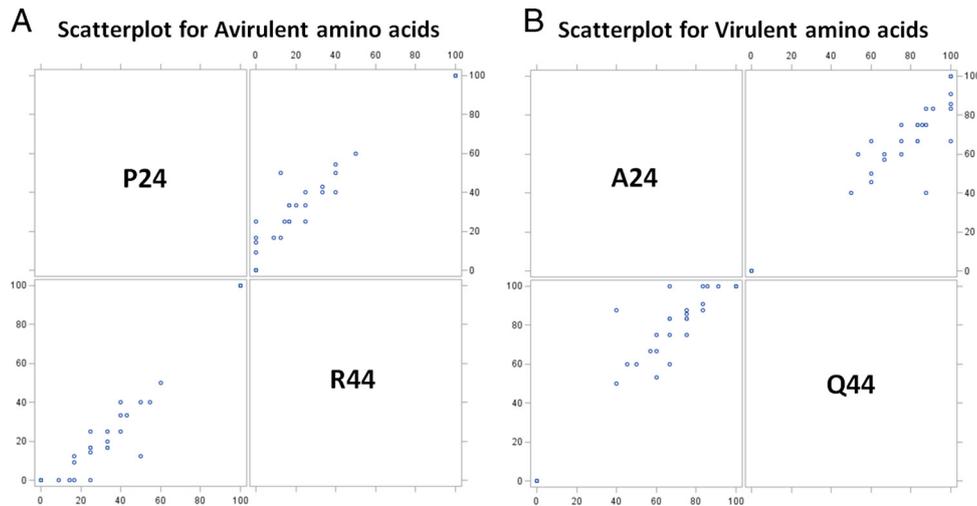


Figure 3: Scatterplots showing the *HgBioB* avirulent and virulent polymorphic amino acid pairs. (A) The avirulent pair shows a positive linear relationship between amino acid positions P24 and R44. (B) Similarly, the virulent pair also shows a positive linear relationship between amino acid positions A24 and Q44.

were detected as well, suggesting an important but unknown function of these two amino acid positions in the *HgBioB* enzyme.

## Discussion

In this study, we compared the enzymatic activities of biotin synthase genes from virulent and avirulent inbred SCN strains and showed that the *HgBioB*-vir from virulent nematodes lacked activity while the *HgBioB*-avr from avirulent nematodes was fully active, suggesting *HgBioB* is important in SCN virulence. Biotin (vitamin B<sub>7</sub> or vitamin H) is a member of the water-soluble vitamin B family and an essential enzyme cofactor, necessary for all three domains of life (Depeint *et al.*, 2006; Zempleni *et al.*, 2008; Waldrop *et al.*, 2012). This vitamin acts as a carboxyl carrier for biotin-dependent carboxylases, decarboxylases, and transcarboxylases, which are critical for fatty acid, amino acid, and carbohydrate metabolic processes (Knowles, 1989; Alban *et al.*, 2000; Attwood and Wallace, 2002; Cronan, 2014). Biotin is also involved in cell signaling, chromatin structure, and epigenetic gene regulation (Zempleni, 2005). More recent discovery shows that biotin plays a role in bacterial virulence (Feng *et al.*, 2014), supporting its potential importance in nematode virulence and parasitism.

Despite its crucial functions, the ability to synthesize biotin *de novo*, either completely or partially, is only found in bacteria, plants, and some fungi (Hall and

Dietrich, 2007; Alban, 2011). Most of these organisms synthesize biotin from pimeloyl-CoA (ACP) in a highly conserved four-step pathway that consists of 7-keto-8-aminopelargonic acid, 7,8-diaminopelargonic acid, DTB, and biotin (Streit and Entcheva, 2003; Alban, 2011; Lin and Cronan, 2011). However, animals do not have the ability to synthesize biotin (Hall and Dietrich, 2007) and must acquire vitamins from their diet (Craig *et al.*, 2008; Craig *et al.*, 2009). Nevertheless, it is strange that SCN, an animal and a parasite, lost the ability to synthesize biotin and then re-acquired the biotin synthase gene via HGT, because the nematode should be able to acquire biotin from its host. It should be noted that SCN does not have the complete biotin biosynthesis pathway, but a partial pathway that encodes the last key enzyme, biotin synthase (Craig *et al.*, 2009), which converts DTB to biotin (Cleary and Campbell, 1972). The presence of this incomplete *de novo* biotin biosynthesis pathway implies that biotin has an important function in SCN and suggests the nematode is scavenging DTB to make biotin from its host (Craig *et al.*, 2009). Although free biotin can be present at high levels in plant cells, unlike animal cells which have more protein-bound biotin (Alban, 2011), it remains unknown if sufficient free biotin is available to SCN as they feed from syncytial cells of high metabolic activity. It is possible that SCN, a filter feeder that does not consume plants via chewing, cannot obtain enough biotin if it is bound to proteins. That is, their presumed prevalence for withdrawing nutrients via feeding tubes, which act as a molecular

sieve (Böckenhoff and Grundler, 1994), may limit their vitamin ingestion; thus, having the biotin synthase gene, or other vitamin B biosynthetic genes, could be a selective advantage.

We previously predicted that virulent nematodes might gain a competitive advantage by producing a more enzymatically active biotin synthase (Bekal *et al.*, 2015) and utilizing the precursors to expand its salvage pathways (Craig *et al.*, 2008, 2009). In our complementation study, however, only the avirulent allele of *HgBioB* (*HgBioB-avr*) successfully complemented the growth of an *E. coli* strain containing a deletion of *bioB*. This showed that avirulent inbred SCN produces an active biotin synthase while virulent inbred nematodes contain an inactive form of the enzyme, thereby refuting our hypothesis.

However, it was possible that the inbred nature of the nematodes selected for a rare *HgBioB* allele that was inactive. To assess how common *HgBioB*-vir alleles are in natural populations of SCN, the *HgBioB* gene was PCR-amplified and sequenced from a series of nematode populations. The sequencing data showed that both SNPs were present in the field SCN populations, resulting in a mixture of both active and inactive forms of *HgBioB*, which is expected due to the heterozygous nature of most nematode populations. In most SCN populations, the virulent forms were prevalent, mirroring the increased prevalence of virulent SCN described by many nematologists (Niblack and Riggs, 2004; Mitchum *et al.*, 2007; Niblack *et al.*, 2008; Tylka and Mullaney, 2015). This implies that an opposing selection for *HgBioB* activity is occurring in the wild. It would be interesting to find a correlation, if any, between the prevalence of active/inactive forms of *HgBioB* with certain Hg-types in the field populations; however, we were not able to harvest enough cysts to carry out Hg-type bioassays and, therefore, all cysts that were harvested were used for the sequencing analysis. There also seems to be an important, but unknown reason why the SNP changes occur concurrently in pairs with similar frequencies between amino acid positions 24 and 44. Since the avirulent and virulent forms of the *HgBioB* alleles only differ at these two positions, we initially speculated that these amino acid sequence polymorphisms might affect the overall biotin synthase activity, especially if these SNPs were located close to the conserved regions responsible for substrate-binding or catalytic activity of the enzyme. However, these SNPs were not in the conserved C<sup>53</sup>xxxC<sup>57</sup>xxC<sup>60</sup> motif (C, cysteine; x, any amino acid), a key characteristic of biotin synthase and other enzymes in the radical S-adenosyl-methionine

superfamily (Frey and Booker, 2001; Jarrett, 2003; Berkovitch *et al.*, 2004). Both the presence of sequence polymorphisms at amino acid positions 24 and 44, and the evidence from the site-directed mutagenesis indicating the absolute requirement for both proline and arginine at those specific locations (P24 and R44) further stress the importance of these amino acids in spite of their unknown functional role in the *HgBioB* enzyme. Since our sequencing data showed that these SNPs did not occur randomly but changed in a coordinated fashion, these SNPs might lead to a complete change in protein function other than biotin synthase. Because the nematode encodes the last key enzyme involved in the biotin biosynthesis pathway, the nematode could have repurposed the biotin synthase enzyme for a completely different task. It has been reported that the housekeeping glutathione synthetase (GS) gene from *G. pallida* was repurposed to diversify its biochemical function, forming a novel class of GS-like enzymes (Lilley *et al.*, 2018). In a similar manner, it is possible that the biotin synthase gene may also modify its biochemical activity to produce a new family of biotin synthase-like enzymes for the evolution of its parasitic lifestyle.

It is also possible that these SNPs might convert the *HgBioB* enzyme into a biotin-related toxin-binding protein that may protect the virulent nematode from host defense mechanisms. Such biotin-related anti-metabolites do occur in bacteria (Hanka *et al.*, 1966, 1969, 1972), but it is unclear if such toxins play a role in SCN resistance. Regardless of the role the inactive form of *HgBioB* plays in the SCN life cycle, it would have to provide a survival advantage to the nematode to counteract any negative effects such as slower growth or lower reproduction due to the absence of biotin scavenging ability of an active *HgBioB*. It has been shown that a reproductive fitness cost is associated with virulence in *M. incognita* (Castagnone-Sereno *et al.*, 2007; Djian-Caporalino *et al.*, 2011), and therefore this could also be applicable to virulent SCN where low biotin availability could cause biotin deficiency, which may result in low reproductive fitness.

In the future, it would be interesting to see if virulent SCN also have a reproductive fitness cost by correlating a nematode's biotin synthase activity or its biotin content with its reproductive index and/or cyst size. Further study of the biotin synthase genes from other SCN strains and even from other plant-parasitic nematodes may help elucidate how the loss of *HgBioB* activity might benefit SCN parasitism and/or virulence, which may lead to the development of novel methods to minimize soybean yield losses affected by these devastating pathogens.

## Acknowledgments

The authors thank Dr Gregory Tylka (Iowa State University, Ames, IA) for providing the SCN field samples from Iowa. The authors also thank Dr Alex Lipka (University of Illinois, Urbana, IL) for his professional advice on statistical analysis. This study was supported by the United Soybean Board (Proposal Number: 1620-532-5620). The authors have no conflict of interest to declare.

## References

- Alban, C. 2011. Biotin (vitamin B8) synthesis in plants, in Rébeillé, F., Douce, R., Kader, J.-C. and Delseny, M. (Eds), *Advances in botanical research*, vol. 59, Academic Press, London, pp. 39–66.
- Alban, C., Job, D. and Douce, R. 2000. Biotin metabolism in plants. *Annual Review of Plant Biology* 51:17–47.
- Allen, T. W., Bradley, C. A., Sisson, A. J., Byamukama, E., Chilvers, M. I., Coker, C. M., Collins, A. A., Damicone, J. P., Dorrance, A. E. and Dufault, N. S. 2017. Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. *Plant Health Progress* 18:19–27.
- Attwood, P. V. and Wallace, J. C. 2002. Chemical and catalytic mechanisms of carboxyl transfer reactions in biotin-dependent enzymes. *Accounts of Chemical Research* 35:113–20.
- Bekal, S., Niblack, T. L. and Lambert, K. N. 2003. A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. *Molecular Plant-Microbe Interactions* 16:439–46.
- 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* 10:e0145601.
- Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. and Drennan, C. L. 2004. Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. *Science* 303:76–9.
- Bird, D. M., Opperman, C. H. and Davies, K. G. 2003. Interactions between bacteria and plant-parasitic nematodes: now and then. *International Journal for Parasitology* 33:1269–76.
- Böckenhoff, A. and Grundler, F. 1994. Studies on the nutrient uptake by the beet cyst nematode *Heterodera schachtii* by in situ microinjection of fluorescent probes into the feeding structures in *Arabidopsis thaliana*. *Parasitology* 109:249–55.
- Castagnone-Sereno, P., Bongiovanni, M. and Wajnberg, E. 2007. Selection and parasite evolution: a reproductive fitness cost associated with virulence in the parthenogenetic nematode *Meloidogyne incognita*. *Evolutionary Ecology* 21:259–70.
- Cleary, P. P. and Campbell, A. 1972. Deletion and complementation analysis of the biotin gene cluster of *Escherichia coli*. *Journal of Bacteriology* 112:830–39.
- Craig, J. P., Bekal, S., Niblack, T., Domier, L. and Lambert, K. N. 2009. Evidence for horizontally transferred genes involved in the biosynthesis of vitamin b(1), b(5), and b(7) in *Heterodera glycines*. *Journal of Nematology* 41:281–90.
- Craig, J. P., Bekal, S., Hudson, M., Domier, L., Niblack, T. and Lambert, K. N. 2008. Analysis of a horizontally transferred pathway involved in vitamin b6 biosynthesis from the soybean cyst nematode *Heterodera glycines*. *Molecular Biology and Evolution* 25:2085–98.
- Cronan, J. E. 2014. Biotin and lipoic acid: synthesis, attachment and regulation. *EcoSal Plus* 6(1):39.
- Davis, E. L., Hussey, R. S., Baum, T. J., Bakker, J., Schots, A., Rosso, M.-N. and Abad, P. 2000. Nematode parasitism genes. *Annual Review of Phytopathology* 38:365–96.
- Depeint, F., Bruce, W. R., Shangari, N., Mehta, R. and O'Brien, P. J. 2006. Mitochondrial function and toxicity: role of b vitamins on the one-carbon transfer pathways. *Chemico-Biological Interactions* 163:113–32.
- Djian-Caporalino, C., Molinari, S., Palloix, A., Ciancio, A., Fazari, A., Marteu, N., Ris, N. and Castagnone-Sereno, P. 2011. The reproductive potential of the root-knot nematode *Meloidogyne incognita* is affected by selection for virulence against major resistance genes from tomato and pepper. *European Journal of Plant Pathology* 131:431–40.
- Dong, K. and Opperman, C. H. 1997. Genetic analysis of parasitism in the soybean cyst nematode *Heterodera glycines*. *Genetics* 146:1311–18.
- Dower, W. J., Miller, J. F. and Ragsdale, C. W. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Research* 16:6127–45.
- Doyle, E. A. and Lambert, K. N. 2003. *Meloidogyne javanica* chorismate mutase 1 alters plant cell development. *Molecular Plant-Microbe Interactions* 16:123–31.
- Endo, B. Y. 1998. Atlas on ultrastructure of infective juveniles of the soybean cyst nematode, *Heterodera glycines*. United States Department of Agriculture, Washington, DC, pp. 215–20.
- Feng, Y., Napier, B. A., Manandhar, M., Henke, S. K., Weiss, D. S. and Cronan, J. E. 2014. A Francisella virulence factor catalyses an essential reaction of biotin synthesis. *Molecular Microbiology* 91:300–14.
- Frey, P. A. and Booker, S. J. 2001. Radical mechanisms of S-adenosylmethionine-dependent enzymes. *Advances in Protein Chemistry* 58:1–45.
- Gleason, C. A., Liu, Q. L. and Williamson, V. M. 2008. Silencing a candidate nematode effector gene corresponding to the tomato resistance gene Mi-1 leads to acquisition of virulence. *Molecular Plant-Microbe Interactions* 21:576–85.

- Hall, C. and Dietrich, F. S. 2007. The reacquisition of biotin prototrophy in *Saccharomyces cerevisiae* involved horizontal gene transfer, gene duplication and gene clustering. *Genetics* 177:2293–307.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* 166:557–80.
- Hanka, L., Bergey, M. and Kelly, R. 1966. Naturally occurring antimetabolite antibiotic related to biotin. *Science* 154:1667–8.
- Hanka, L., Martin, D. and Reineke, L. 1972. Two new antimetabolites of biotin:  $\alpha$ -methyldehydrobiotin and  $\alpha$ -methylbiotin. *Antimicrobial Agents and Chemotherapy* 1:135–8.
- Hanka, L., Reineke, L. and Martin, D. 1969. Biological studies with  $\alpha$ -dehydrobiotin. *Journal of Bacteriology* 100:42–6.
- Hussey, R. S. 1989. Disease-inducing secretions of plant-parasitic nematodes. *Annual Review of Phytopathology* 27:123–41.
- Jarrett, J. T. 2003. The generation of 5'-deoxyadenosyl radicals by adenosylmethionine-dependent radical enzymes. *Current Opinion in Chemical Biology* 7:174–82.
- Jones, M. 1981. Host cell responses to endoparasitic nematode attack: structure and function of giant cells and syncytia. *Annals of Applied Biology* 97:353–72.
- Knowles, J. R. 1989. The mechanism of biotin-dependent enzymes. *Annual Review of Biochemistry* 58:195–221.
- 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* 11:5.
- Lambert, K. N., Allen, K. D. and Sussex, I. M. 1999. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode *Meloidogyne javanica*. *Molecular Plant-Microbe Interactions* 12:328–36.
- Lambert, K. N., Bekal, S., Domier, L. L., Niblack, T. L., Noel, G. R. and Smyth, C. A. 2005. Selection of *Heterodera glycines* chorismate mutase-1 alleles on nematode-resistant soybean. *Molecular Plant-Microbe Interactions* 18:593–601.
- Lilley, C. J., Maqbool, A., Wu, D., Yusup, H. B., Jones, L. M., Birch, P. R., Banfield, M. J., Urwin, P. E. and Eves-Van Den Akker, S. 2018. Effector gene birth in plant parasitic nematodes: neofunctionalization of a housekeeping glutathione synthetase gene. *PLoS Genetics* 14:e1007310.
- Lin, S. and Cronan, J. E. 2011. Closing in on complete pathways of biotin biosynthesis. *Molecular BioSystems* 7:1811–21.
- Lozano-Torres, J. L., Wilbers, R. H., Gawronski, P., Boshoven, J. C., Finkers-Tomczak, A., Cordewener, J. H., America, A. H., Overmars, H. A., Van'T Klooster, J. W. and Baranowski, L. 2012. Dual disease resistance mediated by the immune receptor cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proceedings of the National Academy of Sciences* 109:10119–24.
- McCarville, M. T., Maret, C. C., Mullaney, M. P., Gebhart, G. D. and Tylka, G. L. 2017. Increase in soybean cyst nematode virulence and reproduction on resistant soybean varieties in Iowa from 2001 to 2015 and the effects on soybean yields. *Plant Health Progress* 18:146–55.
- Mitchum, M. G. 2016. Soybean resistance to the soybean cyst nematode *Heterodera glycines*: an update. *Phytopathology* 106:1444–50.
- Mitchum, M. G., Wrather, J. A., Heinz, R. D., Shannon, J. G. and Danekas, G. 2007. Variability in distribution and virulence phenotypes of *Heterodera glycines* in Missouri during 2005. *Plant Disease* 91:1473–6.
- Niblack, T. and Riggs, R. D. 2004. Variation in virulence phenotypes, in Schmitt, D. P., Wrather, J. A. and Riggs, R. D. (Eds), *Biology and management of soybean cyst nematode*, 2nd ed., Schmitt & Associates of Marceline, Marceline, MO, pp. 57–71.
- 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.
- Niblack, T., Colgrove, A., Colgrove, K. and Bond, J. 2008. Shift in virulence of soybean cyst nematode is associated with use of resistance from PI 88788. *Plant Health Progress* 9:29.
- Niblack, T., Heinz, R., Smith, G. and Donald, P. 1993. Distribution, density, and diversity of *Heterodera glycines* in Missouri. *Journal of Nematology* 25:880–6.
- Prior, A., Jones, J. T., Blok, V. C., Beauchamp, J., Mcdermott, L., Cooper, A. and Kennedy, M. W. 2001. A surface-associated retinol-and fatty acid-binding protein (Gp-far-1) from the potato cyst nematode *Globodera pallida*: lipid binding activities, structural analysis and expression pattern. *Biochemical Journal* 356:387–94.
- Riggs, R. D. 2004. History and distribution, in Schmitt, D. P., Wrather, J. A. and Riggs, R. D. (Eds), *Biology and management of soybean cyst nematode*, 2nd ed., Schmitt & Associates of Marceline, Marceline, MO, pp. 9–40.
- Sacco, M. A., Koropacka, K., Grenier, E., Jaubert, M. J., Blanchard, A., Goverse, A., Smant, G. and Moffett, P. 2009. The cyst nematode SPRYSEC protein rbp-1 elicits gpa2-and rangap2-dependent plant cell death. *PLoS Pathogens* 5:e1000564.
- Semblat, J.-P., Rosso, M.-N., Hussey, R. S., Abad, P. and Castagnone-Sereno, P. 2001. Molecular cloning of a cDNA encoding an amphid-secreted putative avirulence protein from the root-knot nematode *Meloidogyne incognita*. *Molecular Plant-Microbe Interactions* 14:72–9.
- Smant, G., Stokkermans, J. P., Yan, Y., De Boer, J. M., Baum, T. J., Wang, X., Hussey, R. S., Gommers, F. J., Henrissat, B. and Davis, E. L. 1998. Endogenous cellulases in animals: isolation of  $\beta$ -1, 4-endoglucanase

genes from two species of plant-parasitic cyst nematodes. *Proceedings of the National Academy of Sciences* 95:4906–11.

Streit, W. and Entcheva, P. 2003. Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. *Applied Microbiology and Biotechnology* 61:21–31.

Tylka, G. and Mullaney, M. 2015. Soybean cyst nematode-resistant soybeans for Iowa. Extension publication PM 1649, Iowa State University Extension, Ames, IA.

Waldrop, G. L., Holden, H. M. and Maurice, M. S. 2012. The enzymes of biotin dependent CO<sub>2</sub>

metabolism: what structures reveal about their reaction mechanisms. *Protein Science* 21:1597–619.

Wrather, A., Shannon, G., Balardin, R., Carregal, L., Escobar, R., Gupta, G., Ma, Z., Morel, W., Ploper, D. and Tenuta, A. 2010. Effect of diseases on soybean yield in the top eight producing countries in 2006. *Plant Health Progress* 11:29.

Zempleni, J. 2005. Uptake, localization, and non-carboxylase roles of biotin. *Annual Review of Nutrition* 25:175–96.

Zempleni, J., Hassan, Y. I. and Wijeratne, S. S. 2008. Biotin and biotinidase deficiency. *Expert Review of Endocrinology & Metabolism* 3:715–24.