EVALUATION OF SOYBEAN GERMPLASM AND IDENTIFICATION OF QUANTITATIVE TRAIT LOCI CONFERRING SOYBEAN CYST NEMATODE RESISTANCE

by

DUNG THUY TRAN

(Under the Direction of Zenglu Li)

ABSTRACT

Soybean cyst nematode is the most destructive pest causing yield reduction in soybean. However, majority of resistant cultivars in the U.S. have their resistance derived from PI 88788 (*rhg1*) and 'Peking' (*rhg1/Rhg4*). It is critical to identify new resistance genes to combat SCN. Through greenhouse evaluation of 462 accessions for SCN resistance and by using SoySNP50K data, a genome-wide association study identified 13 SNPs on five chromosomes that are significantly associated with SCN resistance. Of those, the regions on chromosomes 18 and 8 were known to be *Rhg1* and *Rhg4* loci. Based on this discovery, an $F_{2:3}$ population derived from 'Lee' x PI 567488B was used for genetic mapping. The result indicated that PI 567488B carries two loci for SCN resistance, which were located on chromosomes 16 and 20. The QTL identified could be used for deployment and stacking of resistance alleles to improve SCN resistance in soybean.

INDEX WORDS: *Glycine max*, *Heterodera glycines* (HG), Soybean cyst nematode (SCN), Resistance to *Heterodera glycines* gene (*Rhg*).

EVALUATION OF SOYBEAN GERMPLASM AND IDENTIFICATION OF QUANTITATIVE TRAIT LOCI CONFERRING SOYBEAN CYST NEMATODE RESISTANCE

by

DUNG THUY TRAN

B.S., Vietnam National University of Agriculture, Vietnam, 2013

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2017

EVALUATION OF SOYBEAN GERMPLASM AND IDENTIFICATION OF QUANTITATIVE TRAIT LOCI CONFERRING SOYBEAN CYST NEMATODE RESISTANCE

by

DUNG THUY TRAN

Major Professor: Zenglu Li

Committee: Jim Noe
Esther van der Knaap

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

DEDICATION

This dissertation is dedicated to my mother who always believes and support me to do my best and pursue my dream.

ACKNOWLEDGEMENTS

I would like to give my deep appreciation to my advisor, Dr. Zenglu Li for his guidance, advice, and partially financial support in completing this project and my master study. I also would like to thank my committee members Dr. Esther van der Knapp for her suggestion and support throughout my master study and research proposal; and Dr. Jim Noe for his support regrading some aspects of nematode knowledge and professional development. I want to extend my sincere gratitude to Steve Finnerty for spending much time working with me. He kindly shares all his experiences related to nematode with me. My interest in nematode originated from working with him. Also, I want to thank to two student workers: Emily Walls and Dillion Brown for helping screening in greenhouse and Sarah Jackson for taking care of my plants. I could not have completed this project without their helps. I would thank the members in Soybean Breeding program who serve as a source of advice as well as their help in the field, greenhouse, laboratory and friendship. They make soybean program a great working place for me. Importantly, I would like to acknowledge Vietnam Education Foundation Fellowship for financial support to help me pursue master degree at the University of Georgia. Specially, I would like to thank my mother for always understanding and supporting me. Finally, I would like to thank all my friends at UGA and in my hometown for all their love and encourage when I study aboard.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Soybean *(Glycine max)* introduction

Soybean [*Glycine max* (L.) Merr.)] is a cultivated plant in the Fabaceae family, which contains 650 genera and 18,000 species (Zhu et al., 2005). The domestication event of soybean has been traced back to ancient China about 5,000 years ago (Hymowitz, 1970). Some soy foods such as tofu, douchi and doujiang were popular in ancient times and are still staples of Chinese cuisine today (Stacey, 2008). Records indicate that soybean was introduced into North America through Savannah, Georgia, in 1765 by Samuel Bowen (Hymowitz and Harlan, 1983) but wasn't planted in mid-west area until 1851 (Cumo, 2015). Soybean has many nutrients and is an especially rich source of protein, accounting for 40% of dry soybean weight (Hassan, 2013). Soybean is also the world's largest oilseed crop, as soybean oil is used in baking and frying fats and many industrial products. Thus, soybean has been grown as a main crop for oil, animal feed, food products, and biodiesel (Stacey, 2008). Nowadays, soybean has become the second most grown crop in the United States behind maize, totaling 33.8 million hectares planted, which grossed $40 billion (USD) in 2016 (NASS, 2017). Beyond domestic consumption, nearly 50% of soybean produced in the U.S. in 2016 was exported, making the U.S. the top soybean exporter in the world (http://ers.usda.gov).

Soybean is an annual short-day plant. The root system includes a tap root and many lateral roots. Additionally, soybean nodulation improves soil fertility through nitrogen fixation with rhizobia bacteria (Siczek and Lipiec, 2011). Soybean is classified into 13 different maturity

groups (MG), which may be affected by the environmental conditions of each growing region in the U.S. There are two growth types: determinate and indeterminate. An indeterminate growth habit indicates that the vegetative part will continue to grow after flowering (Boerma and Specht, 2004). Indeterminate soybeans (typically MG 00 to IV) are commonly planted in the northern and central U.S. (McWilliams, 2015). A determinate growth habit indicates that vegetative growth will stop at flowering (Boerma and Specht, 2004). Determinate soybeans (typically MG V to IX) are commonly planted in the southern U.S. (McWilliams, 2015). There are two growth developmental stages for soybean: vegetative and reproductive. The vegetative stages begin with emergence of two cotyledons, followed by the growth of primary unifoliate leaves which are of ovate shape and located oppositely and finally trifoliate leaves, which are compound with three leaflets. The reproductive stage begins with flowering and proceeds until pod maturity (Boerma and Specht, 2004). Soybean is a self-pollinated plant with a perfect papilionaceous flower and four complete parts: five sepals, one petal, nine stamens, and one pistil. One node may produce as many as 2 to 20 pods with the number of seeds per pod ranging from 1 to 5 (Boerma and Specht, 2004).

Soybean belongs to the *Glycine* genus consisting of two subgenera, *Glycine* Willd. and *Soja* (Moench) F. J. Hermann, with a total of 28 species (Chang et al., 2014). The subgenus *Soja* contains the domesticated soybean species, *G. max* which has a chromosome number of 2n = 40 (Ratnaparkhe et al., 2011). The *G. max* genome sequence was released with estimated genome size of 1,115 megabases (Mb) with 46,430 predicted proteins (Schmutz et al., 2010). The reference genome was assembled using the variety, ‘William 82’.

Soybean cyst nematode (*Heterodera glycines*)

Morphology

The soybean cyst nematode (SCN; *Heterodera glycines*) is an obligate endoparasite that requires a living root to complete its life cycle (Niblack et al., 2006). The morphology of SCN changes during its life cycle. At the second stage juvenile, the cyst nematode is worm shaped with offset head and tapering tail (EPPO, 2008). Body length is from 375 to 520 µm with a body diameter 18 µm (Davis and Tylka, 2000). The head of SCN could be recognized by the presence of a strong stylet, approximately 22-26 µm long (EPPO, 2008). The stylet has three parts: conical conus-end, cylindrical shaft and three basal knobs (Sharma, 1998). SCN adults are dissimilar in appearance. The female is swollen and sedentary while the male is vermiform in shape and motile (Niblack et al., 2006). The male will emerge from the root while the female will become larger within the root although her vulva is displayed on the root exterior (Niblack et al., 2006). After fertilization, the fertilized egg remains inside of the female body. When the female dies, the female cuticle will become dark and tough in order to protect the eggs (Sharma, 1998). The female body also forms lemon shaped cysts with variation in color (white, yellow or brown) based on maturity level (Niblack et al., 2006). Therefore, cyst shape and color are used to identify SCN on soybean roots.

Life cycle

The SCN is an obligate parasite with six life stages: egg, four juvenile (J), and adult (Figure 1.1) (Opperman and Bird, 1998). The life cycle ranges from 25 to 40 days depending on environmental conditions and suitable hosts, and the optimum temperatures (Opperman and Bird, 1998). Inside the egg, the J1 hatches into J2, which is the only infective stage. Depending on suitable environment conditions, J2 will develop a stylet to pierce the egg shell and emerge

from the cyst (Niblack et al., 2006). The J2 will travel through the soil in search of plant roots through chemoreception (Perry, 1996). Using its robust stylet to penetrate plant cells, the J2 also uses a pump mechanism near the esophagus zone to inject secreted proteins that affects the plant nucleus and other organelles (Niblack et al., 2006). The SCN effectors will alter the morphology of cortical and pericycle cells of host plant cells (Davis and Mitchum, 2005). Cell walls of neighboring cells will dissolve, then fuse into a large cytoplasmic feeding space called the syncytium. The primary goal of the nematode is to absorb nutrients from the plants at the feeding site (Davis et al., 2004). Once the feeding site within the host has been established, the J2 then migrates and becomes the J3 and J4 (Niblack et al., 2006).

After receiving nutrition, the J4 then undergoes sex differentiation. The J4 will elongate during maturity to form adult males with a vermiform shape. The motile adult male does not feed on the root, but rather exits the root to search for females (Niblack et al., 2006). After fertilization, the female will die and her body is hardened to protect the eggs; some eggs enter dormancy while others are left outside in a gelatinous matrix, called an "egg sac" (Niblack et al., 2006). The female can produce as many as 600 eggs, depending on the health of host plant (Moore et al., 1981). Dormant SCN eggs are protected within the dead female body, called a "cyst" that may survive for up to eight years or more in the soil (Moore et al., 1981). Eggs will only hatch after meeting appropriate environmental conditions, with an optimum temperature between 25 to 30°C (Moore et al., 1981).

History and distribution

The SCN was first reported in northeast China in 1899 (Liu, 1997) and in Japan and Korea in 1915 and 1936, respectively (Ichinohe, 1959; Kim et al., 2013). The pest subsequently spread to many other soybean producing region/countries such as Taiwan, Columbia, and Brazil.

SCN was first found in the United States in 1954 in North Carolina. Then SCN spread to Tennessee and Missouri in 1956 and to Arkansas and Kentucky in 1957 (Riggs, 1975). It was first thought that SCN was introduced into the USA with *Rhizobium* in 19th century. SCN damage was not noticed until large scale soybean production began (Noel, 1986). In 2014, SCN was recorded in most soybean producing states except for New York and West Virginia (Tylka and Marett, 2014) (Figure 1.2). In 2017, SCN was reported in New York (Wang et al., 2017).

Symptoms and damage

SCN damage may not be detected when the SCN population is low. Above ground symptoms of SCN infection including stunting and chlorosis, which are found at high SCN population densities. However, these symptoms may be confused with other nutrient deficiency problems (Moore et al., 1981), because SCN infected plants have poorly developed root systems that cannot absorb enough nutrients and water (Lambert and Bekal, 2002). The digestive enzymes secreted by SCN interfere with normal plant growth (Davis and Tylka, 2000). Additionally, with minor SCN induced injury, above ground symptoms aren't usually observed until pod emergence. Therefore, above ground symptoms may not be an obvious and reliable method to scout for SCN. To properly identify responses to infection, plants should be removed to see below ground symptoms on the roots (Davis and Tylka, 2000). Females are lemon-shaped and less than 1 mm in diameter on roots (Niblack et al., 2006). Breaking down and washing the root gently or analyzing soil samples are common methods to confirm the presence of SCN. In terms of crop yield losses, Bradley and Allen (1996) estimated a 25% yield loss (3.4 million metric tons) in U.S soybean production in 2014 (Figure 1.3). SCN damage may cause a 30% yield reduction before above ground symptoms are observed (Young, 1996).

Race system and HG type

Race system and HG (*Heterodera glycines*) type were developed to classify SCN virulence. The race system was first developed with four soybean lines: 'Peking', 'Pickett', PI 88788 and PI 90763 (Table 1.1) (Riggs and Schmitt, 1988). Races are determined by measuring and comparing the ability of cyst production on these four genotypes with the susceptible cultivar 'Lee' (Riggs and Schmitt, 1988), which is defined as female index (FI). If an indicator line has FI greater than 10%, then this line is considered to be susceptible (Riggs and Schmitt, 1988). However, this classification system did not suit the variability of SCN because of its complexity (Niblack et al., 2002). Therefore, Niblack et al. (2002) developed a new classification system called HG type to differentiate *H. glycines* virulence based on the nematode fecundity on seven indicator lines instead of four lines (Table 1.2). With HG type system, the susceptible check 'Lee' was replaced with 'Lee 74' because the number of cysts counted on 'Lee' was often inconsistent (Niblack et al., 2002). Classification of HG type is based on a list of indicator lines with designated numbers. For instance, an HG Type 2.5.7 score indicates that PI 88788 (No. 2), PI 209332 (No. 5), and PI 548316 (No. 7) had a FI greater than 10%; by contrast, HG Type 0 indicates that no indicator line had a FI greater than 10% (Niblack et al., 2002). The new system is considered to be more efficient because of the ability to identify new races of SCN and to characterize new sources of SCN resistant genotypes.

Resistance to SCN

Planting resistant cultivars has been the most effective method for managing SCN in soybean. Screening for SCN resistance in the U.S. has occurred since the 1950s (Ross, 1957). At least 158 soybean accessions have been confirmed to be resistant to SCN (Rincker et al., 2017). The first resistant cultivar, 'Pickett' selected from 'Peking' resistant to SCN race 1 and 3, yielded

(2,900 kg/ha) more than susceptible cultivars, 'Tracy' (2,080 kg/ha) in SCN race 3 infested fields (Young, 1992). Using resistant cultivars has prevented soybean yield loss when grown in the SCN infested areas. Chen et al. (2001) evaluated soybean yields of resistant and susceptible cultivars in Minnesota. In six surveyed fields, 56 resistant cultivars mostly derived from PI 88788 and 'Peking' had 28.4% of the increased yield (676 kg/ha) than susceptible cultivars. Also, Tylka and Mullaney (2016) evaluated yield of 45 resistant varieties and 4 susceptible varieties at two locations in Iowa. Planting resistant varieties increased up to 1.02 t/ha. However, most of commercial SCN resistant cultivars have been derived from only two plant introductions (PIs), 'Peking' and PI 88788 (Concibido et al., 2004). It is estimated that greater than 95% of commercial SCN resistant cultivars in the U.S. are derived from PI 88788. Specifically, 95% of resistant cultivars in Illinois and most of the resistant cultivars planted in Iowa are derived from PI 88788 (Melito et al., 2010; Tylka and Mullaney, 2016). As a result, the ability of SCN populations to overcome PI 88788 derived resistance is thought to be considerable with some SCN populations already shifting (Table 1.4). In 2005, 70% of SCN samples overcame PI 88788 derived resistance in Illinois (Niblack et al., 2008). Similarly, Mitchum et al. (2007) observed that 78% of collected SCN populations evaluated overcame PI 88788 based resistance, and that 70% of sampled SCN populations overcame the resistance from PI 209332 and PI 548316 which belonged to PI 88788 based resistance, and that 30% of collected SCN populations overcame 'Peking'. Also, based in one screening survey conducted in Kentucky, 12 out of the 20 SCN populations sampled overcame PI 88788 based resistance (Hershman et al., 2008). Therefore, it is obvious that the effectiveness of PI 88788 derived resistance will continue to decrease if new SCN resistant sources are not utilized for cultivar development.

Rhg1 and *Rhg4*, major genetic loci for SCN resistance

Inheritance of soybean resistance to SCN is complex. Early inheritance studies showed recessive and dominant genes were involved in SCN resistance: *rhg1, rhg2* and *rhg3* (Caldwell et al., 1960), *Rhg4* (Matson and Williams, 1965), and *Rhg5* (Rao-Arelli et al., 1992). More than thirty quantitative trait loci (QTL) have been reported on 20 chromosomes since 1994 (Figure 1.4) (Concibido et al., 2004). A major QTL on chromosome 18 (LG-G) is known as the *Rhg1 (Resistance to H. glycine)* locus. Also, this locus has been designated by the Soybean Genetics Committee as *cqSCN-001* (Glover et al., 2004). Because the resistance alleles are present in most resistant germplasm sources including the seven indicator lines 'Peking', PI 90763, PI 88788, PI 437654, PI 209332, PI 89772 and PI 404198A used in the HG type test (Table 1.3), it becomes the most prevalent SCN resistance locus used in soybean cultivar development (Concibido et al., 1996; Concibido et al., 1997 ; Guo et al., 2006). Concibido et al. (1997) estimated that this locus alone accounts for 50% of the phenotypic variation for SCN race 1, 3, and 6 resistances. Although *rhg1* was identified as a recessive allele, its gene action is considered to be incompletely dominant because cysts could be formed in the lines that are heterozygous at this locus (Concibido et al., 2004; Melito et al., 2010).

Additionally, it was observed that the responses to SCN varied among the resistant lines possessing the *rhg1* resistance allele. The 'Peking'-type *rhg1* and PI 88788-type *rhg1* had different cellular responses during SCN infection (Matsye et al., 2011). They compared SCN resistance mechanisms between 'Peking' and PI 88788 and reported that the 'Peking'-type *rhg1* allele caused nematodes lethality at the second juvenile stage from cell wall appositions that lead to a thicker cell wall, whereas PI 88788-type *rhg1* resistance did cause cell wall thickening, but only delayed SCN death until the third or fourth juvenile stage (Matsye et al., 2011).

Additionally, when studying the effect of an *α-SNAP* gene located within *Rhg1* locus, they found that polymorphism was present among 'Peking', PI 437654, PI 88788, and 'Williams 82'. This suggested that allelic differences caused cell wall apposition in 'Peking', serving as a barrier to SCN penetration. As a result, the resistance allele of *rhg1* from PI 88788 was given a new name, *rhg1-b* by the Soybean Genetics Committee (Kim et al., 2010).

The *rhg1-b* allele in PI 88788 was fine mapped within a 67-kb interval between two SSR markers on chromosome 18 (LG G): BARCSOYSSR_18_0090 and BARCSOYSSR_18_0094. A 31.2 kb genome segment of *rhg1-b* with multiple copies causing phenotypic differences in SCN resistant lines was identified (Cook et al., 2012). There are three distinct genes contributing to SCN resistance found within each repeat: *Glyma18g02580*, *Glyma18g02590*, and *Glyma18g02610*. The *Glyma18g02580* gene encodes a predicted amino acid transporter (*GmAAT*). The *Glyma18g0290* gene encodes an *α-SNAP* protein (*GmSNAP18*), whereas the *Glyma18g02610* gene encodes a wound-inducible protein 12 (*GmWI12*). Silencing any of these three genes reduces SCN resistance (Cook et al., 2012). However, overexpression of each gene individually did not enhance SCN resistance, while overexpression of all three genes simultaneously improved resistance (Cook et al., 2012). According to the number of repeats present, cultivars were classified into three categories: single copy, low copy number (2-4 copies), and high copy number (more than 6 copies) (Cook et al., 2014) . PI 88788; PI 548316 ('Cloud') and PI 209332 belonged to the high copy number group with 9 and 10 copies, respectively, whereas 'Peking', PI 90763, PI 89772 and PI 437654 were classified into the low copy group with just three copies (Cook et al., 2014). Susceptible genotypes, such as 'Williams 82', only have a single copy at *Rhg1* locus (Cook et al., 2012). The study also characterized the relationship between copy number and resistance, finding that both PI 209332 (10 copies) and PI

88788 (9 copies) were more resistant than ‘Cloud’, which possessed seven copies. Additionally, it was found that PI 209332 (10 copies) was resistant to races 3, 5, 14, whereas PI 88788 (9 copies) was resistant to races 3 and 14 only (Lee et al., 2015). However, the distinction of reactions to SCN between low and high copy number groups was not clear because the *Rhg4* gene also conveys SCN resistance in the lines with low copy numbers (Cook et al., 2014). Although PI 438489B carries only two copies at *Rhg1* locus (Lee et al., 2015), it was characterized having resistance to five SCN populations (1, 2, 3, 5, 14) (Diers et al., 1997), suggesting that expression of *Rhg1* locus maybe mediated by other resistance loci (Lee et al., 2015).

The *Rhg4* locus, designated as *cqSCN-002,* and mapped to Chr 8 (LG-A) (Webb et al., 1995), is the second major locus conferring SCN resistance. This locus was identified in some resistant sources including ‘Peking’, PI 89772, PI 90763, and PI 437654 and contributed 28% of the total phenotypic variance observed for SCN resistance (Table 1.3). The *Rhg4* locus is situated at 0.35 cM from the *I* locus that controls black seed color (Weisemam et al., 1992). Their research found two genes functioning at the *Rhg4* locus: the first encodes a *serine hydroxy methyltransferase* (*SHMT*), whereas the second encodes a subtilisin-like protease, but *SHMT* was reported for providing SCN resistance (Liu et al., 2012). They also reported that SCN resistant cultivars became susceptible when silencing this gene by virus-induced gene silencing and RNA interference. Sequencing the *SHMT* genes from 28 soybean lines revealed eight different haplotypes, and PI 90763, PI 437654 and PI 89772 were grouped with ‘Peking’ in carrying *SHMT* resistance alleles and *rhg1*. The two loci, *Rhg1* and *Rhg4* have an interaction or a complementation effect (Brucker et al., 2005). For ‘Forrest’, both resistance loci must be present to provide resistance to SCN, which explained 70% of the phenotypic variation (Meksem et al.,

2001). However, lines with a high-copy number at the *Rhg1* locus, such as PI 209332, PI 88788, and PI 548316, were still resistant to SCN without the *Rhg4* locus.

Other QTLs for SCN resistance

Many other QTL have been identified for SCN resistance, but most of them only showed a minor effect (Guo et al., 2006). However, two genes: *rhg2* and *rhg3* that were reported with *rhg1* in the inheritance study (Caldwell et al., 1960) have not yet been characterized. A third QTL locus, designated as *cqSCN-003* by Soybean Genetics Committee, confers resistance to two SCN populations: PA3 (HG Type 7, race 3) and PA 14 (HG Type 1. 3. 5. 6. 7, race 14). This QTL contains *Rhg5* locus, which is located on Chr 16 (LG-J) (Glover et al., 2004). Using near-isogenic lines of 'Bell' derived from PI 88788, *Rhg5* was mapped between SSR markers Satt547 and Satt431 (Glover et al., 2004). A fourth QTL for HG Type 0 resistance, designated *cqSCN-005*, was mapped on Chr 17 (LG D2) from 'Hartwig', a cultivar derived from PI 437654 (Kazi et al., 2010). Two other QTLs have been reported from a wild soybean accession PI 468916: *cqSCN-006* and *cqSCN-007* (Wang et al., 2001). *cqSCN-006* was then fine-mapped to a 212.1 kb interval on Chr 15 (LG E) with flanking SNPs, ss715621232 and ss715621239 (Yu and Diers, 2017), explaining 23% of total phenotypic variation for HG type 2.5.7 resistance (Kim and Diers, 2013). Locus *cqSCN-007* was fine-mapped on Chr 18 (LG-G), between BARC18_1669 and ss715631888 within a 103.2 kb interval (Yu and Diers, 2017), which explained 27% of phenotypic variation for HG Type 2.5.7 resistance (Kim and Diers, 2013). Although located on Chr 18, *cqSCN-007* was mapped to a different location than *rhg1*.

PI 567516C, originating from China, was reported to have resistance to race 1, 2, and 3, and LY1 (a highly virulent combination of race 2 and 3) (Arelli et al., 2010). Two QTLs were mapped in an $F_{2:3}$ population derived from Magellan × PI 567516C: one on Chr 10 (LG O) and

the other on Chr 18 (LG G), located 80 cM from *Rhg1* locus (Vuong et al., 2010) . In addition, the *rhg1* allele from PI 567516C was reported to provide reniform nematode resistance in a study examining resistance to reinform nematode (Jiao et al., 2015). Because previous SCN mapping did not identify the *Rhg1* locus as being responsible for SCN resistance, whole genome sequencing was conducted indicating that the *Rhg1* locus is present in PI 567516C and PI 567516C belongs to the 'Peking'-type resistance group with a low copy number for *rhg1* (Vuong et al., 2010). The result was confirmed by genotyping with the KASP SNP marker (GSM381 and GSM383) at *Rhg1* locus reported by Shi et al. *(*2015*)* (Tran et al., unpublished data).

Genome-wide association study (GWAS) for SCN resistance

GWAS is known as association mapping to detect the association between genetic variants and phenotypes. Different from bi-parental QTL mapping, GWAS is used to examine genetic variants at a whole genome level using a panel of diverse lines, in which both phenotypes and high-density genomic data are used. The use of unrelated genotypes increases mapping resolution and reduces time in comparison to QTL mapping (Korte and Farlow, 2013). Linkage disequilibrium (LD) and population structure are known to impact GWAS results, since LD and recombination rate have reversed proportions. Population structure should be considered in GWAS because Type I errors may increase when related individuals are used (Zhu et al., 2008). To account for this, a compressed Mixed Linear Model is applied in many GWAS applications (Korte and Farlow, 2013). When using linear regression, there are three basic assumptions: (1) that the phenotype is normally distributed; (2) the variance is the same, and (3) that sampling is independent. The significant threshold *p*-value is calculated as 0.05/ number of markers (Bonferroni correction) or 0.001 (false discovery rate *p*-value), which is used to reduce Type 1 errors (Sham and Purcell, 2014). The readily available SoySNP50K Infinium Chip dataset has

increased the use of GWAS because of the larger number of markers available on every chromosome. In soybean, GWAS has been applied to identify significant regions controlling seed composition (Vaughn et al., 2014), iron deficiency (Mamidi et al., 2014), flowering and maturity (Zhang et al., 2015) and biotic resistance (Rincker et al., 2016).

GWAS has also been used for identification of loci for SCN resistance (Table 1.5; Figure 1.5). A GWAS for HG Type 0 resistance was performed using 282 *G. max* accessions and 1,247 Universal Soy Linkage Panel 1.0 (USLP) SNPs by Bao et al. (2014). A total of six significant SNPs residing on Chr 18 was found. Of these six SNPs, four SNPs were located at the known *Rhg1* locus; one SNP was located within *FGAM1* gene, 1.1 Mb away from *Rhg1*; and the last one was located at the end of Chr 18 (*Glyma18g46201*)., Vuong et al. (2015) performed a GWAS for HG Type 0 resistance with 553 *G. max* accessions in maturity group from III to V using 45,000 SNP markers from SoySNP50K iSelect Beadchip (soybase.org). They reported a total of 60 significant SNPs from 14 genomic regions, and 13 of them were previous mapped QTL including *Rhg1*, *Rhg4*, and *qSCN10* on Chr 18, 8 and 10 respectively. One novel QTL was located at on Chr 2 (at ~ 13.66 Mb). Using a total of 440 *G. max* landraces and elite cultivars that were genotyped by Specific Locus Amplified Fragment Sequencing (SLAF-seq) and phenotyped with two HG Types: 0 and 1.2.3.5.7 (race 3 and race 4 respectively), Han et al. (2015) reported a total of 19 SNPs associated with two HG Type (12 SNPs for HG Type 0 and 7 SNPs for HG Type 1.2.3.5.7). Of 19, eight SNPs were located at known *Rhg1* and *Rhg4* loci and eight SNPs were overlapped or linked with reported QTLs. Three remaining SNPs on Chrs 2 (rs21804864: 21,804,864 bp); 14 (rs30581306: 30,581,306) and 20 (rs2085816: 2,085,816) were considered to be located in novel QTL regions for HG Type 0. Besides the HG type 0, GWAS analysis was also reported for HG Type 2.5.7 (race 1) resistance. A panel of 200 *G. max* accessions including

180 accessions from China were genotyped by > 33,000 SNP markers using the Specific Locus Amplified Fragment Sequencing (SLAF-seq) (Zhao et al., 2017). Thirteen significant SNPs on Chrs 7,8, 14, 15 and 18 responsible for HG Type 2.5.7 resistance were found. Also, 120 Chinese *G. max* accessions were screened with HG Type 2.5.7 and genotyped with 7,189 SNPs (Zhang et al., 2017). Three of SNPs on Chr 8, 13, and 17 were considered as novel QTL for HG Type 2.5.7 resistance. These two GWAS studies mapped HG Type 2.5.7 resistance to the known *Rhg1* and *Rhg4* loci.

Marker-assisted selection for SCN resistance

Molecular markers are often applied as a tool to select the traits of interest to save time and resources in plant breeding programs. With advances in genetic marker technology, many types of markers such as restriction fragment length polymorphism (RFLP), random amplified polymorphic (RAPD), simple sequence repeats (SSR) and single nucleotide polymorphism (SNP) were used in QTL mapping studies, which have identified the *Rhg1* locus on Chr 18 and the *Rhg4* locus on Chr 8 in many SCN resistant cultivars (Concibido et al., 2004; Shi et al., 2015). However, these markers were found to be incapable of distinguishing between the resistant lines, such as PI 88788 and PI 209332, and susceptible lines (Cregan et al., 1999; Concibido et al., 2004). With discovery of candidate genes controlling to SCN resistance at the *Rhg1* and *Rhg4* loci (Cook et al.,2012; Liu et al., 2012), three SNP markers were developed to be used for selection of PI 88788 and 'Peking'-type SCN resistance (Shi et al, 2015). Two of the SNPs markers (GSM381 and GSM383) reside at the *Glyma18g02590* gene at the *Rhg1* locus, whereas the other SNP marker (GSM191) is located at the *Glyma08g11490* gene at the *Rhg4* locus. At the *Rhg1* locus, marker GSM383 is able to distinguish between 'Peking' and PI 88788-

type resistance in the screened genotypes (Shi et al., 2015). Therefore, these markers have been useful in deploying marker-assisted selection for SCN resistance in breeding programs.

Summary

Soybean cyst nematode, a cosmopolitan obligate endoparasite, has been the most devastating pest in the U.S., causing over a 25% reduction in yield loss. SCN has been observed in most soybean producing states. To date, 16 different races of SCN populations were determined in screening studies using four SCN resistant soybean cultivars. Due to the variation observed among the races, the new system of classification was developed to refer SCN populations as HG Types. Seven genotypes were used as indicators to determine HG Types by comparing cyst counts on each genotype compared to those on the susceptible cultivar 'Lee 74'. Results revealed that Race 3 (HG Type 0) is the most predominant SCN population in the U.S. Although numerous SCN resistant accessions have been identified, most commercial varieties are derived from PI 88788. The seven resistant sources have been genetically divided into two types of resistance groups: 'Peking' type, with 2-4 copies at *rhg1* and presence of the resistance allele at *Rhg4* locus, and PI 88788 type with more than six copies at *rhg1* but no *Rhg4* resistance allele. Due to the narrow genetic diversity used by breeding programs for SCN resistance, it has been observed that some SCN populations have overcome those resistance mechanisms. Therefore, identifying new sources of resistance as well as new resistance genes is of paramount importance to soybean breeders.

The main goals of this study are to: 1) screen soybean germplasm to search for new sources of SCN resistance that are different from 'Peking' and PI 88788 types; and 2) to identify QTLs derived from these new sources of SCN resistance using GWAS and bi-parental mapping approaches.

References

Arelli, P.R., V.C. Concibido and L.D. Young. 2010. QTLs associated with resistance in soybean PI 567516C to synthetic nematode population infecting cv. *Hartwig*. J. Crop Sci Biotech. 13: 163-167. doi:10.1007/s12892-010-0060-z.

Bao, Y., T. Vuong, C. Meinhardt, P. Tiffin, R. Denny, S. Chen, et al. 2014. Potential of Association Mapping and Genomic Selection to Explore PI 88788 Derived Soybean Cyst Nematode Resistance. The Plant Genome. 7. doi:10.3835/plantgenome2013.11.0039.

Boerma, H.R. and J.E. Specht, editors. 2004. Soybeans: Improvement, Production, and Uses. 3rd ed. ASA, CSSA, and SSSA, Madison, WI.

Bradley, C., T. Allen and A. Wrather. 1996. Estimates of soybean yield reductions caused by diseases in the United States. UIUC, Illinois, USA.

Brucker, E., S. Carlson, E. Wright, T. Niblack and B. Diers. 2005. *Rhg1* alleles from soybean PI 437654 and PI 88788 respond differentially to isolates of *Heterodera glycines* in the greenhouse. Theor Appl Genet 111: 44-49. doi:10.1007/s00122-005-1970-3.

Caldwell, B.E., C. A. Brim, and J. P. Ross. 1960. Inheritance of resistance of soybeans to the cyst nematode, *Heterodera glycines*. J. Agron. 52: 635-636.

Chang, S., C.S. Thurber, P.J. Brown, G.L. Hartman, K.N. Lambert and L.L. Domier. 2014. Comparative mapping of the wild perennial *Glycine latifolia* and soybean (*G. max*) reveals extensive chromosome rearrangements in the genus *Glycine.* PLoS One 9: e99427. doi:10.1371/journal.pone.0099427.

Chen, S. Y., Porter, P. M., Orf, J. H., Reese, C. D., Stienstra, W. C., Young, N. D., Walgenbach,D. D., Schaus, P. J., Arlt, T. J., and Breitenbach, F. R. 2001. Soybean cyst nematode population development and associated soybean yields of resistant and susceptible cultivars in Minnesota. Plant Dis. 85:760-766.

Concibido, V.C., B.W.Diers and P.R.Arelli. 2004. A decade of QTL mapping for cyst nematode resistance in soybean. Crop Sci.44: 1121-1131.

Concibido, V.C., D.A. Lange, R.L.Denny, J. H. Orf, N. D. Young. 1997. Genome mapping of SCN resistance genes in Peking, PI 90763 and PI 88788 using DNA markers. Crop Sci. 37: 258-264.

Concibido, V.C., R. L. Denny, D. A. Lange, J. H. Orf, and N. D. Young. 1996. RFLP mapping and marker-assisted selection of soybean cyst nematode resistance in PI 209332. Crop Sci.36: 1643-1650.

Cook, D.E., T. G. Lee et al. 2012. Copy number variation of multiple genes at *rhg1* mediates nematode resistance in soybean. Science 338: 1206-1209.

Cook, D.E., A.M. Bayless, K. Wang, X. Guo, Q. Song, J. Jiang, et al. 2014. Distinct Copy Number, Coding Sequence, and Locus Methylation Patterns Underlie *Rhg1*-Mediated Soybean Resistance to Soybean Cyst Nematode. Plant Physiol. 165: 630-647. doi:10.1104/pp.114.235952.

Cregan, P.B., J. Mudge, E. W. Fickus, D. Danesh, R. Denny, and N. D. Young. 1999. Two simple sequence repear markers to select for soybean cyst nematode resistance conditioned by the *rhg1* locus. Theor Appl Genet: 811-818.

Cumo, C. 2015. Plant Biotechnology. Plants and People. CRC Press: 130-133.

Davis, E.L., and G. L. Tylka. 2000. Soybean cyst nematode disease. The Plant Health Ins. doi:10.1094/PHI-I-2000-.

Davis, E.L., R.S. Hussey and T.J. Baum. 2004. Getting to the roots of parasitism by nematodes. Trends Parasitol. 20: 134-141. doi:10.1016/j.pt.2004.01.005.

Davis, E.L. and M.G. Mitchum. 2005. Nematodes. Sophisticated parasites of legumes. Plant Physiol. 137: 1182-1188. doi:10.1104/pp.104.054973.

Diers, B.W., H. T. Skorupska, P. R. Arelli and S. R. Cianzo. 1997. Genetic relationships among soybean plant introduction with resistant to soybean cyst nematode. Crop Sci. 37: 1966-1972.

EPPO. 2008. *Heterodera glycines*. EPPO Bulletin 38: 379-389. doi:10.1111/j.1365-2338.2008.01249.x.

Glover, K.D., D. Wang, P. R. Arelli, S. R. Carlson, S. R. Cianzio, B. W. Diers. 2004. Near isogenic lines confirm a soybean cyst nematode resistance gene from PI 88788 on linkage group J. Crop Sc. 44: 936-941.

Guo, B., D.A. Sleper, H.T. Nguyen, P.R. Arelli and J.G. Shannon. 2006. Quantitative Trait Loci underlying Resistance to Three Soybean Cyst Nematode Populations in Soybean PI 404198A. Crop Sci. 46: 224. doi:10.2135/cropsci2004.0757.

Han, Y., X. Zhao, G. Cao, Y. Wang, Y. Li, D. Liu, et al. 2015. Genetic characteristics of soybean resistance to HG type 0 and HG type 1.2.3.5.7 of the cyst nematode analyzed by genome-wide association mapping. BMC Genom. 16: 598. doi:10.1186/s12864-015-1800-1.

Hassan, S.M. 2013. Soybean, Nutrition and Health. Licensee InTech. doi:10.5772/54545.

Hershman D. E., R. D. Heinz, B.S. Kennedy. 2008. Soybean cyst nematode, *Heterodera glycine*s, populations adapting to resistant soybean cultivars in Kentucky. Plant Dis. 92: 1475.

Hymowitz, T. 1970. On the domestication of the Soybean. Econ. Bot 24: 408-421.

Hymowitz, T., and J.R. Harlan. 1983. Introduction of soybean to North America by Samuel Bowen. Econ. Bot 37: 371-379.

Ichinohe, M. 1959. Studies on the soybean cyst nematode *Heterodera glycines* and its injury to soybean plants in Japan. Plant Dis. Rep. USDA, USA.

Jiao, Y., T.D. Vuong, Y. Liu, Z. Li, J. Noe, R.T. Robbins, et al. 2015. Identification of quantitative trait loci underlying resistance to southern root-knot and reniform nematodes in soybean accession PI 567516C. Mol Breed. 35: 131. doi:10.1007/s11032-015-0330-5.

Kazi, S., J. Shultz, J. Afzal, R. Hashmi, M. Jasim, J. Bond, P. R. Arelli and D. A. Lighfoot. 2010. Iso-lines and inbred-lines confirmed loci that underlie resistance from cultivar 'Harwig' to three soybean cyst nematode populations. Theor Appl Gene. 120: 633-644. doi: 10.1007/s00122-009-1181-4.

Kim, M., D.L. Hyten, A.F. Bent and B.W. Diers. 2010. Fine Mapping of the SCN Resistance Locus *rhg1-b* from PI 88788. J. Plant Genome. 3: 81. doi:10.3835/plantgenome2010.02.0001.

Kim, M. and B.W. Diers. 2013. Fine Mapping of the SCN Resistance QTL and from PI 468916. Crop Sci. 53: 775. doi:10.2135/cropsci2012.07.0425.

Kim, M.S., M.K. Sung, M.W. Kim, H.J. Seo, D.G. Kim and J.I. Chung. 2013. Investigation of Soybean Cyst Nematode *Heterodera glycines* Type and Evaluation of Resistance on Soybean Varieties and Germplasms in Korea. Korean J. Crop Sci. 58: 161-168. doi:10.7740/kjcs.2013.58.2.161.

Korte, A., and A. Farlow. 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods. 9: 29.

Lambert, K.N., and S. Bekal. 2002. Introduction to Plant-Parasitic Nematodes. The Plant Health Instructor. doi:10.1094/PHI-I-.

Lee, T.G., I. Kumar, B.W. Diers and M.E. Hudson. 2015. Evolution and selection of *Rhg1*, a copy-number variant nematode-resistance locus. Mol Ecol. 24: 1774-1791. doi:10.1111/mec.13138.

Liu, S., P.K. Kandoth, S.D. Warren, G. Yeckel, R. Heinz, J. Alden, et al. 2012. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. Nature 492: 256-260. doi:10.1038/nature11651.

Liu, X. 1997. History and status of soybean cyst nematode in China. Int. J. Nematol. 7: 18-25.

MacGuidwin, A. E. 2012. Distribution of SCN HG Types in Wisconsin. 2011. Wiscosin Crop Production Association Distinguished Service Awards: 107.

Mamidi, S., R. K. Lee, J. R. Goos, P. E., McClean. 2014. Genome-wide association studies identifies seven major regions responsible for iron deficiency chlorosis in soybean

(*Glycine max*). PLoS One. 9 (9): e107469. https://doi.org/10.1371/journal. pone. 0107469.

Matson, A. L., and Williams, L. F. 1965. Evidence of a fourth gene for resistance to the soybean cyst nematode. Crop Sci. 5: 477.

Matsye, P.D., R. Kumar, P. Hosseini, C.M. Jones, A. Tremblay, N.W. Alkharouf, et al. 2011. Mapping cell fate decisions that occurs during soybean defense response. Plant Mol Biol. 77: 513-528. doi:10.1007/s11103-011-9828-3.

McWilliams, D.A., D. R. Berglund, and G. Endres. 2015. Soybean growth and management quick guide. A-1174. North Dakota State University.

Meksem, K., P. Pantazopoulos, V.N. Njiti, L.D. Hyten, P.R. Arelli and D.A. Lightfoot. 2001. 'Forrest' resistance to the soybean cyst nematode is bigenic: saturation mapping of the *Rhg1* and *Rhg4* loci. TAG Theoretical and Applied Genetics. 103: 710-717. doi:10.1007/s001220100597.

Melito, S., A.L. Heuberger, D. Cook, B.W. Diers, A.E. MacGuidwin and A.F. Bent. 2010. A nematode demographics assay in transgenic roots reveals no significant impacts of the *Rhg1* locus LRR-Kinase on soybean cyst nematode resistance. BMC Plant Biol 10: 104. doi:10.1186/1471-2229-10-104.

Moore, W.F et al. 1981. Soybean cyst nematode. [United States] : Prepared and distributed under the auspices of the Soybean Industry Resource Committee, [1981].

Mitchum, M.G., J.A. Wrather, R.D. Heinz, J.G. Shannon and D. Gene. 2007. Variability in distribution and virulence phenotypes of *Heterodera glycines* in Missouri during 2005. Plant Dis. 91: 1473-1476.

Mudge, J.P., B. Cregan, J.P. Kenworthy, W. J. Kenworthy, J. H. Orf, and N. D. Young. 1997. Two mucrosatellite markers that flank the major soybean cyst nematode resistance locus. Crop Sci. 37: 1611-1615.

NASS. 2017. Crop Prod 2015 Summary. USDA.

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

Niblack, T.L., K. N. Lambert, and G. L. Tylka. 2006. A Model Plant Pathogen from the Kingdom Animalia: *Heterodera glycines*, the Soybean Cyst Nematode. Annu Rev Phytopathol. 44: 283-303. doi:10.1146/.

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

Noel, G.R. 1986. The Soybean Cyst Nematode. In: F. Lamberti and C. E. Taylor, editors, Cyst Nematodes. Springer US, Boston, MA. p. 257-268.

Opperman, H.C., and D. M. Bird. 1998. The soybean cyst nematode, *Heterodera glycines*: a genetic model system for the study of plant-parasitic nematodes. Curr Opin Plant Biol. 1.

Perry, N.R. 1996. Chemoreception in plant parasitic nematodes Annu. Rev. Phytopathol. 34.

Rao-Arelli, A. P., Anand, S. C., and Wrather, J. A. 1992. Soybean resistance to soybean cyst nematode race is conditioned by an additional dominant gene. Crop Sci. 32: 862-864.

Ratnaparkhe, M.B. et al. 2011. *Glycine*. Wild crop relatives: genomic and breeding resources legume crops and forages. editor C. Kole. Berlin. Heidelberg. Springer. 83-116.

Riggs, R.D. 1975. Worldwide distribution of soybean-cyst nematode and its economic importance. Ann Meet Nematol & Am Phytopathol. Soc. Houston, Texas, U.S.A.

Riggs, R.D. and D.P.Schmitt. 1988. Complete characterization of the race scheme for *Heterodera glycines.* J. Nematol. 20: 392-395.

Rincker, K., et al. 2016. Genome-wide association study of brown stem rot resistance in soybean across multiple populations. Plant Genome. 9 (2): 68-78.

Rincker, K., T. Cary and B.W. Diers. 2017. Impact of Soybean Cyst Nematode Resistance on Soybean Yield. Crop Sci. doi:10.2135/cropsci2016.07.0628.

Ross, J.P. 1957. Resistance of soybeans to the soybean cyst nematode as determined by a double-row method. Plant Dis Rep. 41: 923-924.

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, et al. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463: 178-183. doi:10.1038/nature08670.

Sham P.C., and S. M. Purcell. 2014. Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet. 15 (5): 335-346.

Sharma, S.B. 1998. Hatch and emergence. In: S. B. Sharma, editor The Cyst Nematodes. Springer Netherlands, Dordrecht. p. 191-216.

Shi, Z., S. Liu, J. Noe, P. Arelli, K. Meksem and Z. Li. 2015. SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. BMC Genomics. 16: 314. doi:10.1186/s12864-015-1531-3.

Siczek, A. and J. Lipiec. 2011. Soybean nodulation and nitrogen fixation in response to soil compaction and surface straw mulching. Soil Tillage Res. 114: 50-56. doi:10.1016/j.still.2011.04.001.

Stacey, G. 2008. Genetics and genomics of soybean. Plant Genetics and Genomics: Crops and Models: 3-15.

Tylka, G.L. and C.C. Marett. 2014. Distribution of the Soybean Cyst Nematode, H*eterodera glycines*, in the United States and Canada: 1954 to 2014. Plant Health Prog. doi:10.1094/php-br-14-0006.

Tylka, G. L., and M. P. Mullaney. 2016. Soybean cyst nematode-resistant soybeans for Iowa. Ext. Punl. PM 1649. Iowa State Univ., Ames.

Vaughn, N. J., R. L. Nelson, Q. Song, P.B. Cregan, and Z. Li. 2014. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. G3: Gene, Genomes, Genetics. 4: 2283-2294. doi: 10. 1534/g3.114.013433.

Vuong, T.D., D.A. Sleper, J.G. Shannon and H.T. Nguyen. 2010. Novel quantitative trait loci for broad-based resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe) in soybean PI 567516C. Theor Appl Genet. 121: 1253-1266. doi:10.1007/s00122-010-1385-7.

Vuong, T.D., H. Sonah, C.G. Meinhardt, R. Deshmukh, S. Kadam, R.L. Nelson, et al. 2015. Genetic architecture of cyst nematode resistance revealed by genome-wide association study in soybean. BMC Genomics 16: 593. doi:10.1186/s12864-015-1811-y.

Wang, D., P. R. Arelli, R. C. Shoemaker, and B. W. Diers. 2001. Loci underlying resistance to race 3 of SCN in glyine soja plant introduction 468916. Theor Appl Genet. 103.

Wang, W., X., G. Bergstrom, S. Chen, D. M. Thurston, J. Cummings, Z. Handoo, et al. 2017. First Report of the Soybean Cyst Nematode, *Heterodera glycines*, in New York. The Am. Phytopathol. Soc. 101 (11): 1957.

Webb, D.M., B.M.Baltazar, A.P. Rao-Arelli, J. Schupp, K. Clayton, P. Keim, and W.D. Beavis. 1995 .Genetic mapping of SCN race 3 resistance loci in the soybean PI 437654. Theor Appl Genet. 91: 574-581.

Weisemam, J.M., B. F. Matthews, and T. E. Devine. 1992. Molecular markers located proximal to the soybean cyst nematode resistance gene, *Rhg4*. Theor Appl Genet. 85.

Winter, S.M., B.J. Shelp, T.R. Anderson, T.W. Welacky and I. Rajcan. 2007. QTL associated with horizontal resistance to soybean cyst nematode in Glycine soja PI 464925B. Theor Appl Genet. 114: 461-472. doi:10.1007/s00122-006-0446-4.

Wrather, A., G. Shannon, R. Balardin, L. Carregal, R. Escobar, G.K. Gupta, et al. 2010. Effect of Diseases on Soybean Yield in the Top Eight Producing Countries in 2006. Plant Health Prog. doi:10.1094/php-2010-0125-01-rs.

Young, L. D. 1992. Management of the soybean cyst nematode, *Heterodera glycines* in soybean. In: Pest management in soybean. L. G. Copping, M. B. Green and R. T. Rees. Doordrecht. Springer: 137-146.

Young, L.D. 1996. Yield loss in soybean caused by *Heterodera glycines*. Suppl. J. Nematol. 28: 604-607.

Yu, N. and B.W. Diers. 2017. Fine mapping of the SCN resistance QTL *cqSCN-006* and *cqSCN-007* from *Glycine soja* PI 468916. Euphytica 213. doi:10.1007/s10681-016-1791-2.

Yue, P., D. A. Sleper, and P. R. Arelli. 2001. Mapping Resistance to Multiple Races of *Heterodera glycines* in Soybean PI 89772 Journal series number: 13,029 of Missouri Agric. Exp. Stn . Crop Sci. 41:1589-1595. doi:10.2135/cropsci2001.4151589x

Zhang, H., Q. Song, J.D. Griffin and B.H. Song. 2017. Genetic architecture of wild soybean (*Glycine soja*) response to soybean cyst nematode (*Heterodera glycines*). Mol Genet Genomics. doi:10.1007/s00438-017-1345-x.

Zhang, J., et al. 2015. Genome-wide association study for flowering times, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. BMC Genomics 16 (1): 217.

Zheng, J., Y. Li., and S. Chen. 2006. Characterization of the virulence phenotypes of *Heterodera glycines* in Minnesota. J. Nematol. 38 (3): 383-390.

Zheng, J.W., and S.Y.Chen. 2011. Estimation of virulence type and level of soybean cyst nematode field populations in response to resistant cultivars. J. Nematol 3: 37-43.

Zhu, C.S., et al. 2008. Status and prospects of association mapping in plants. Plant Genome 1 (1): 5-20.

Zhu, H., H.K. Choi, D.R. Cook and R.C. Shoemaker. 2005. Bridging model and crop legumes through comparative genomics. Plant Physiol 137: 1189-1196. doi:10.1104/pp.104.058891.

Table 1.1 List of four indicator lines for SCN race determination.

No	Indicator lines	Race 1[a]	Race 2	Race 3
1	'Pickett'	-	+	-
2	'Peking'	-	+	-
3	PI 88788	+	+	-
4	PI 90763	-	-	-

[a] Race determination based on "+" (susceptible reaction) and "-" (resistant reaction). Standard susceptible check is 'Lee'.

Table 1.2 List of 7 indicator lines for HG Type determination and 2 examples of HG Type.

No	Indicator lines	HG Type 0	HG Type 2.5.7
1	'Peking'	-	-
2	PI 88788	-	+[a]
3	PI 90763	-	-
4	PI 437654	-	-
5	PI 209332	-	+
6	PI 89772	-	-
7	PI 548316	-	+

[a] "+" is susceptible reaction (FI >10%) and "-" is resistant reaction (FI <10 %). HG Type is determined by lines that have susceptible reactions. 'Lee 74' is susceptible check.

Table 1.3 List of seven sources of SCN resistance.

No	SCN resistant sources	Resistance allele	References
1	'Peking'	*rhg1, Rhg4*	Caldwell et al., 1960
2	PI 88788	*rhg1*	Concibido et al., 1997
3	PI 90763	*rhg1, Rhg4*	Concibido et al., 1997
4	PI 437654	*rhg1, Rhg4*	Webb et al., 1995
5	PI 209332	*rhg1*	Concibido et al., 1996
6	PI 89772	*rhg1, Rhg4*	Yue et al., 2001
7	PI 548316	*rhg1*	Kim et al., 2010

Table 1.4 Shift of SCN populations reported in the U.S.

State	Surveyed year	'Peking' (%)	PI 88788 (%)	References
Minnesota	2002	1[a]	15	Zheng et al., 2006
Missouri	2005	29	78	Mitchum et al., 2007
Illinois	2005	8	70	Niblack et al., 2008
Kentucky	2006-2007	25	60	Hershman et al., 2008
Wisconsin	2011	26	78	MacGuidwin, 2012

[a] The percentage of surveyed nematode populations reproduced on soybean varieties with 'Peking' and PI 88788 resistance.

Table 1.5 Summary of QTL for SCN resistance reported in genome-wide association studies.

HG Type	Marker	No. of markers used	Population size	No of mapped loci	References
HG Type 0	SNP (USLP array)	1247	282 (*G.max*)	3	Bao et al., 2014
HG Type 0	SNP (SoySNP50K)	45,000	553 (*G.max*)	14	Vuong et al., 2015
HG Type 1.2.3.5.7 and 0	SNP (SoySNP50K)	36,976	440 (*G.max*)	19	Han et al., 2015
HG Type 2.5.7	SNP (iSelect Bead Chip)	41,087	235 (*G. soja*)	4	Zhang et al., 2017
HG Type 2.5.7	SNP (8K iSelect Bead Chip)	7189	120 (*G.max*)	5	Zhang et al., 2017
HG Type 2.5.7	SNP (SLAF-seq)	33,194	200 (*G.max*)	5	Zhao et al., 2017

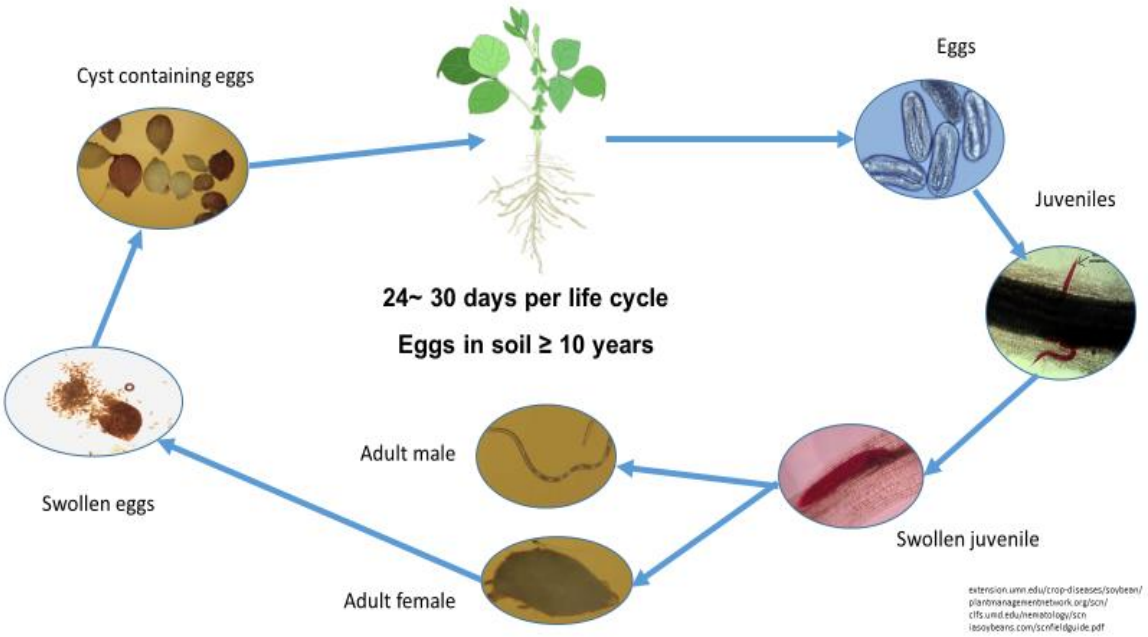

Figure 1.1 Life cycle of the soybean cyst nematode.

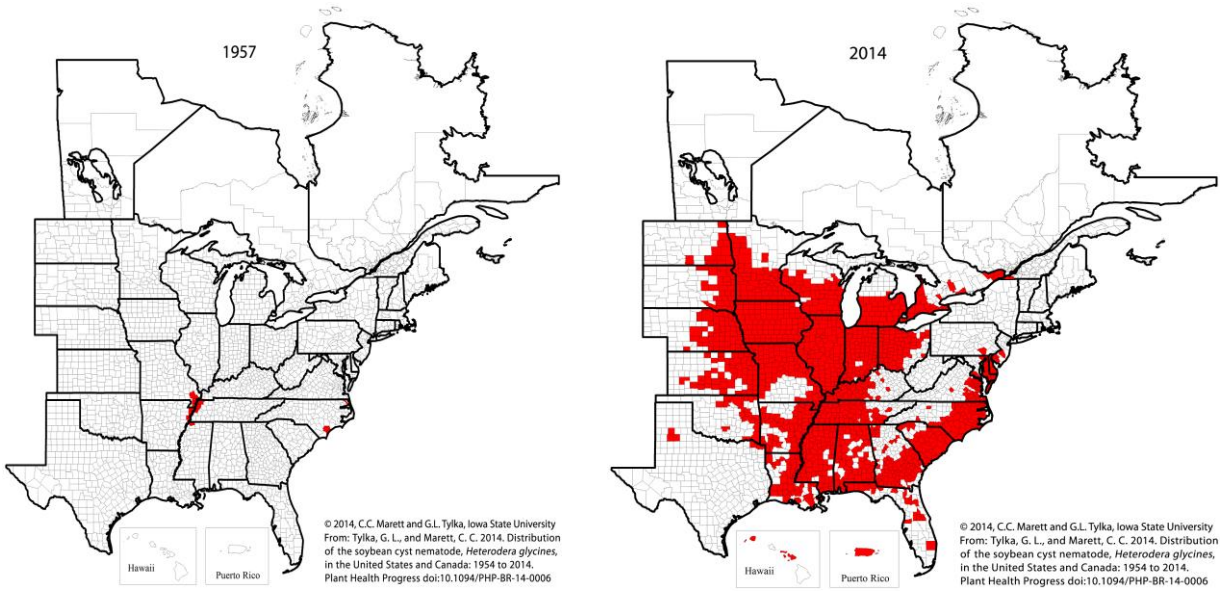

Figure 1.2 Distribution maps of SCN in the U.S. in 1957 (left) and 2014 (right) showed the fast spread of SCN to soybean production states. (Source: Tylka, 2014)

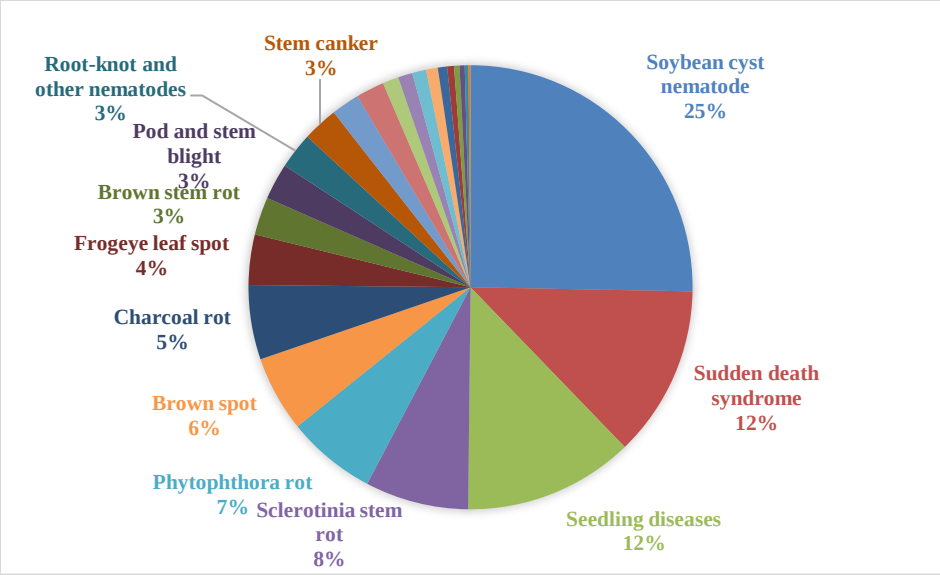

Figure 1.3 Estimated yield loss caused by disease in soybean in the U.S. in 2014.
(Data source: http://extension.cropsciences.illinois.edu/fieldcrops/diseases/yield_reductions.php)

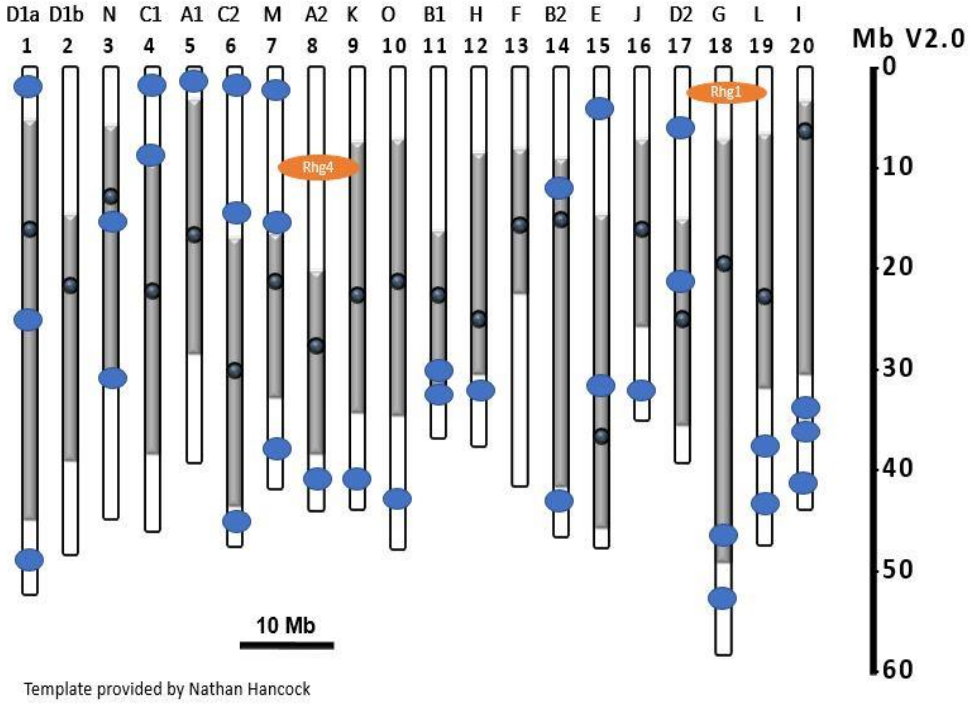

Figure 1.4. Summary of QTLs for SCN resistance reported using bi-parental mapping.
Illustration depicting the two major QTLs (*Rhg1* and *Rhg4* loci, orange dots) and minor QTLs (blue dots) conveying soybean cyst nematode (SCN) resistance in soybean from former mapping studies. The numbers and letters at the top of the figure correspond to the assigned letter and number for each soybean chromosome, and the scales located to the right and at the bottom of the figure representing the physical length of the chromosomes in base pairs.

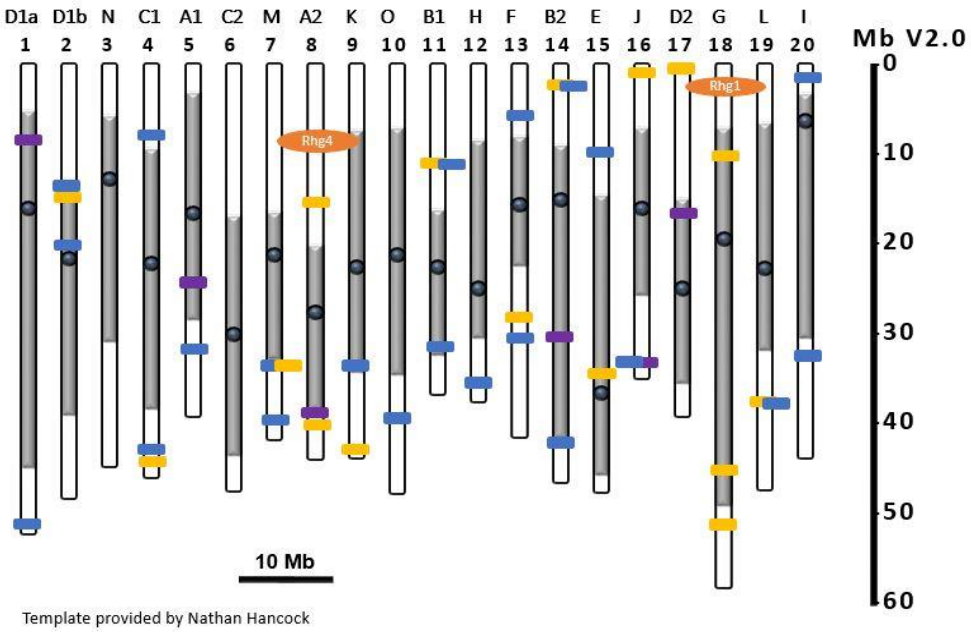

Figure 1.5 Overview of QTLs for SCN resistance using GWAS analysis in soybean.

Rhg1 on Chr 18 and *Rhg4* loci on Chr 8 (orange box) are major QTLs that were identified in most of SCN resistant sources.

The colored bars indicated the QTL for different HG Types: Blue=HG Type 0 (race 3); Yellow=HG Type 2.5.7 (race 1); and Purple=HG Type 1.2.3.5.7 (race 4). The letter at the top of the figure correspond to the assigned linkage groups (LG) in soybean whereas the numbers beneath the letters correspond to the assigned chromosome. The bar to the right corresponds to the Mb length of each chromosome, based on the 'Williams 82' soybean reference genome V.2.0.

CHAPTER 2

SCREENING SOYBEAN GERMPLASM AND IDENTIFYING QUANTITATIVE TRAIT LOCI CONFERRING SOYBEAN CYST NEMATODE RESISTANCE USING A GENOME-WIDE ASSOCIATION STUDY[1]

[1] Tran, D. T., Boehm, J., Arelli, P. R., Noe, J., Li. Z. Identifying quantitative trait loci conferring soybean cyst nematode resistance using a genome-wide association study. To be submitted to Plant Science.

Abstract

Soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is the most destructive pest affecting soybeans (*Glycine max*) in the U.S. Two major resistance alleles, *rhg1* and *Rhg4* residing on chromosomes (Chr) 18 and 8, respectively, provide genetic control of SCN resistance, which were identified in PI 88788 (*rhg1*) and 'Peking' (*rhg1*/*Rhg4*). To date, PI 88788 and 'Peking' have been widely used to develop SCN resistant cultivars in the U.S. for soybean production. However, this has become a major concern for soybean breeders because evolving SCN populations have overcome the PI 88788 and 'Peking' derived resistance. Therefore, it is essential to identify new sources of SCN resistance and deploy them in soybean cultivars. To that end, 462 soybean accessions from various origins were screened using a greenhouse SCN bio-assay and genotyped with three SNPs developed at the *Rhg1* and *Rhg4* loci for SCN resistance as well as a SNP marker for southern root-knot nematode (RKN*:* *Meloidogyne incognita*) resistance. Of 462 accessions, 50 accessions were classified as the 'Peking'-type resistance (*rhg1*/*Rhg4*), while 30 accessions were classified as PI 88788-type resistance (*rhg1*). Additionally, there were 58 accessions that were rated as SCN resistance through greenhouse phenotyping that carried neither the 'Peking' nor the PI 88788 resistance alleles at *Rhg1* and *Rhg4* loci. Based on haplotype analysis at the *Rhg1* and *Rhg4* loci using SoySNP50k Infinium Chip data, these lines were grouped separately from 'Peking' and PI 88788. The genome-wide association study (GWAS) was performed on this panel of 461 accessions using 35,817 SNPs from the Soy50KSNP Infinium Chip data. The GWAS identified 13 SNPs at five genomic regions on Chrs 2, 7, 8, 10, and 18 that were significantly associated with SCN resistance. Of those, three SNPs were located at two known major resistance gene loci, *Rhg1* and *Rhg4*. Thirty-three predicted genes are found near the significant SNPs on Chrs 2, 7,

and 10 that encode various types of protein kinase, receptor-like protein, zinc fingers and RING, suggesting that they might be genes associated with SCN resistance. The identified SNPs as well as candidate genes from this study might be beneficial for the development of DNA markers to be used for marker-assisted breeding and to aid in developing soybean cultivars with novel sources of SCN resistance.

Keywords: Soybean (*Glycine max*), Soybean cyst nematode (*Heterodera glycines*), Resistance to *Heterodera glycines* (*Rhg*) genes, Female index (FI), Genome-wide association (GWAS), Quantitative trait loci (QTL), Linkage disequilibrium (LD), Single nucleotide polymorphism (SNP), SoySNP50K Infinium Chip data.

Introduction

Soybean [*Glycine max* (L.) Merrill] is the most significant economic crop in the legume family for both oil and food products. Soybean has numerous nutrients and is an especially rich source of protein (about 40% by weight) (Hassan, 2013). Currently, soybean has become a second most grown crop in the U.S. behind maize, totaling 33.2 million hectares planted annually, and grossing $34 billion in 2015 in the U.S. (NASS, 2017). Beyond domestic consumption, nearly 50% of soybean production in the U.S. in 2015 went to exports, making the U.S. the top soybean exporter of the world (ers.usda.org). However, soybean production in the U.S. is strongly undermined by SCN. This pest caused yield losses of up to 3.4 million tonnes in 2014 (extension.cropscience.illinois.edu).The SCN is an obligate parasite with a worm-like shape and an unsegmented invertebrate body and requires a living root to complete its life cycle (Niblack et al., 2006). It enters the soybean root at juvenile stages and modifies soybean cells to absorb nutrients thereby preventing the soybean to mature and develop properly. When the female dies, its body called as a cyst to protect the eggs. The cysts may persist in soil for up to 11 years which pose a potentially serious economic threat (Niblack et al., 2006).

SCN was first identified in 1899 in China (Liu, 1997), in Japan in 1915 (Ichinohe, 1959), and in Korea in 1936 (Kim et al., 2013). SCN wasn't reported in the U.S. until 1954 (Riggs, 1975), starting in North Carolina, then spreading westward to Tennessee, Missouri, Arkansas, and Kentucky (Tylka and Marett, 2014). In 2017, SCN was first reported in New York (Wang et al., 2017). Currently, SCN is found in every soybean producing state except West Virginia (Tylka and Marett, 2014, Wang et al., 2017). When SCN is present, the primary above ground symptom is leaf chlorosis. However, this symptom is not the most reliable indicator because the symptom mimics the plant's response to other abiotic and biotic stresses such as nutrient

deficiency (Moore et al., 1981). According to the report by Koenning and Wrather (2010), yield loss can be as high as 30% even without any above ground noticeable symptoms. To properly confirm the presence of SCN, soybean roots or soil should be evaluated for a number of cysts which may be yellow or brown and lemon shaped (Niblack et al., 2006).

Virulence of SCN is variable with 16 possible races determined by four selected differential lines including 'Peking', 'Pickett', PI 88788 and PI 90763 (Riggs and Schmitt, 1988). Of the 16 races, race 3 is considered as the predominant race in the U.S. (Jackson, 2014). Later, to meet diversity population of SCN genotypes, the SCN population classification was modified using seven indicator lines: 'Peking', PI 88788, PI 89772, PI 90763, PI 209332, PI 437654, and PI 548316. The new population classification system termed 'HG Type' is determined by comparing SCN fecundity on each indicator line with the standard susceptible cultivar, 'Lee 74' (Niblack et al., 2002).

Breeding SCN resistant cultivars and rotation with non-host crops is the most effective control method. Screening soybean germplasm for resistance to SCN began as early as 1957 (Ross, 1957). Although many resistant cultivars have been reported, PI 88788 has been primarily used to breed resistant cultivars due to its desirable agronomic traits (Concibido et al., 2004). Tylka et al. (2016) evaluated 51 resistant varieties developed by private seed companies, and found that 97% of them derived their resistance from PI 88788, and only a few from 'Peking' and PI 437654.

Use of a single source of SCN resistance such as PI 88788 or even multiple sources carrying the same SCN resistance gene may lead to a genetic shift in SCN populations by increasing the selection pressure. Several studies have reported that as many as 78% of SCN populations in Missouri and 12% of SCN populations in Minnesota have already overcome PI

88788 and ‘Peking’ resistance (Mitchum et al., 2007; Zheng et al., 2006). The continued planting of soybean varieties with derived PI 88788 and ‘Peking’ SCN resistance will increase SCN populations to overcome these defense mechanisms. Therefore, it is critical to identify new SCN resistant germplasm sources from diverse genetic backgrounds that are different than ‘Peking’ and PI 88788.

The inheritance of SCN resistance is complex. Early inheritance studies reported three recessive alleles designated as *rhg1, rhg2* and *rhg3* in ‘Peking’ (Caldwell et al., 1960); and two dominant alleles designated as *Rhg4* (Matson and Williams, 1965) and *Rhg5* (Rao-Arelli, 1994) controlling SCN resistance. However, until now, no mapping information about *rhg2* and *rhg3* is available. More than 30 QTLs controlling SCN resistance have been reported since 1994 by linkage mapping, with most of them showing only minor effect on SCN resistance (Concibido et al., 2004). A first major QTL on Chr 18, linkage group (LG) G known as *Rhg1* (Resistance to *H. glycine*) locus was reported to be present in most of resistant sources used for breeding commercial varieties including ‘Peking’, PI 88788, and PI 437654 (Concibido et al., 2004). At the *Rhg1* locus, there were allelic differences detected between ‘Peking’ and PI 88788, so the resistance alleles were denoted as *rhg1-a* (‘Peking’) and *rhg1-b* (PI 88788) (Kim et al., 2010). A 31.2 kb genomic segment of *rhg1-b* with multiple copies causing phenotypic differences in SCN resistant lines was later identified (Cook et al., 2012). Within the 31.2 kb segment, three distinct genes contributing to SCN resistance were found within each repeat: *Glyma.18g02580*; *Glyma.18g02590* and *Glyma.18g02610*. The *Glyma.18g02580* gene encodes a predicted amino acid transporter, the *Glyma.18g0290* encodes an *α-SNAP* protein, and the *Glyma.18g02610* gene encodes a protein with a wound-inducible protein 12 (*WI12*) (Cook et al., 2012). The second major QTL providing SCN resistance, which is the *Rhg4* locus on Chr 8 (LG A2), was identified

as being present in some of resistant sources such as ‘Peking’ and PI 437654 (Concibido et al., 2004). One resistance gene that encodes a *serine hydroxymethyltransferase* (*SHMT*) was later found at the *Rhg4* region (Liu et al., 2012). Beyond that, other QTLs for SCN resistance have also been discovered: *cqSCN-003* (*Rhg5*) located on Chr 16 (LG-J) from PI 88788 (Glover, 2004); *cqSCN-005* located on Chr 17 (LG D2) from ‘Hartwig’ (Kazi et al., 2010); *cqSCN-006* on Chr 15 (LG E) and *cqSCN-007* on Chr18 (LG G) from a wild soybean accessions PI 468916 (Kim and Diers, 2013) as well as *qSCN10* located on Chr 10 (LG O) from PI 567516C (Vuong et al., 2010).

Although QTL mapping in bi-parental populations is a powerful approach to identify genomic regions, only genomic regions that have allelic variation between the two parents may be used to detect resistance, and a low amount of recombination may decrease the mapping resolution. Genome-wide association studies (GWAS) utilize the genetic diversity of a panel of unrelated individuals to capture more recombination events by creating shorter linkage disequilibrium (LD) blocks that allow for identification of significant QTL with higher resolution (Zhu et al., 2008). The development of single nucleotide polymorphism (SNP) genotyping technology has supported the utilization of GWAS. In soybean, GWAS has been applied to dissect QTL controlling seed quality, abiotic tolerance traits, disease resistance, and yield components (Jun et al., 2008; Wang et al., 2008; Bastien et al., 2014 and Vaughn et al., 2015).

GWAS has been previously used to locate genomic regions providing SCN resistance. Using a panel of 159 Chinese soybean accessions genotyped with 55 SSR loci, Li et al. (2011) identified three significant SSRs associated with SCN race 3 resistance on Chrs 7, 17, and 18. Three QTLs on Chr 18, which included *Rhg1, FGAM1* and *Glyma.18g46201* were reported by Bao et al. (2014) to condition HG Type 0 (race 3) resistance in a set of 282 soybean accessions

from University of Minnesota soybean breeding program that were genotyped with 1,536 SNPs markers. Using an association panel of 440 accessions genotyped with Specific Locus Amplified Fragment sequencing (SLAF-sequencing), 12 and seven SNPs significantly associated with resistance to HG Type 0 and 1.2.3.5.7, respectively, were identified. Of these 19 SNPs, eight were linked with two major SCN resistance QTLs *Rhg1* and *Rhg4;* and 11 other SNPs were distributed on 11 chromosomes (Han et al., 2015). Also, Vuong et al. (2015) reported 14 genomic regions associated with HG Type 0 resistance detected among 553 soybean accessions in maturity groups III to V using SoySNP50K iSelect BeadChip data. Of these 14 genomic regions, three were associated with *Rhg1, Rhg4* and *qSCN-10* residing on Chrs 18, 8, and 10, respectively (Table1.4).

The objectives of this study were: 1) identify new source of resistance to HG Type 0 (SCN race 3) by evaluating diverse soybean germplasm, and 2) map the genomic regions associated with HG Type 0 resistance.

Materials and Methods

Plant materials

Five hundred and thirty-five accessions from maturity group (MG) 0 to MG VIII were selected from the USDA Soybean Germplasm Collection for greenhouse screening. These lines originated from various origins, with 61% of them originating from China, which is the center of the domestication of soybean (Hymowitz, 1970). The 535 accessions, along with six known susceptible cultivars ‘Hutcheson’, ‘NC Roy’, ‘NC Raleigh’, ‘CNS’, ‘Lee’, ‘Lee 74’ and one resistant line G93-9009 were evaluated in greenhouse. Additionally, seven indicator lines ‘Peking’, PI 88788, PI 90763, PI 437654, PI 209332, PI 89772 and PI 548316 were also included for HG Type determination.

Greenhouse phenotyping

Due to the large number of accessions, the accessions were split into three sets, which were composed of 82, 204, and 249 accessions in each set, respectively (Table 2.1). Greenhouse screening was performed using SCN race 3 (HG Type 0) in the Plant Pathology Greenhouse at the University of Georgia, Athens, GA in 2016. Plants were grown in cones (20.6 cm length and 4 cm diameter) that were filled with a fumigated sandy loam soil. Cones were arranged into a randomized completed block design (RCBD) with four replications. Four seeds per accession were planted in each cone, and then were thinned to a single seedling at 7-9 d after planting. Each seedling represented a single replicate, which was inoculated with 2,000 eggs placed in 3-4 mL of water with a dispenser machine. Approximately 40-60 d after inoculation when the number of cysts on 'Lee 74' (susceptible check) exceeded 50, cysts on the roots were counted. Plant roots were individually washed free of soil, and then dried for 30 min. The female cysts were counted under a 20X lighted magnifying glass. The level of resistance was defined by the female index (FI) that was calculated based on the ratio between the mean numbers of cysts on a given line and on 'Lee 74', reported as a percentage (Niblack et al., 2002). Rating scale of SCN was based on Schmitt and Shannon (1992): FI < 10% (resistant, R); 10% < FI < 30% (moderately resistant, MR); 30% < FI < 60% (moderately susceptible, MS), and FI > 60% (susceptible, S).

Based on the screening results from three sets, 106 accessions were selected that were rated as resistant or moderately resistant and were rescreened in the greenhouse during Winter 2016. These accessions were then reselected based on three criteria: (1) they were rated as highly or moderately resistance to SCN race 3; (2) fewer than three plants per line were evaluated; and (3) accessions did not carry the 'Peking' or PI 88788-type SNP resistance alleles. A panel of 106

accessions were subsequently sown in 10.16 cm (4 inch) wide clay pots filled with a fumigated sandy loam soil in December of 2016. Pots were arranged in a RCBD with four replications. A heat mat was used underneath the pots to maintain temperature at 28 – 30° C. Four seeds were planted in each pot, and then thinned to a single seedling per pot after 7-9 d. In this confirmation test, a new and more aggressive population of SCN race 3 (compared to HG type 0) that was collected from Collins, Georgia in 2016 was used for inoculation (Table 2.5). According to HG Type designation using seven indicator lines (Table 2.1), the SCN race 3 population collected from Collins, GA was designated as HG Type 5. At 9 d after planting, each pot was then inoculated with 2,000 HG Type 5 (SCN race 3) eggs. At 38 d after inoculation when the cyst counts on susceptible check 'Lee 74' reached approximately 100, all plants were subjected to cyst counts.

SNP marker genotyping and 50K SNP array

Twelve young leaves per line were collected, and then freeze-dried for 48 h. DNA was extracted from soybean leaves by a modified CTAB method (Keim et al., 1988) and stored at minus 20°C until use. DNA concentration was quantified using a TECAN Infinite T1000 Pro (Tecan US, Inc, Morrisville, NC, U.S.) and diluted with water to 10-20 ng/µL for Kompetitive Allele Specific polymerase chain reaction (KASP) assays.

All soybean accessions from screening sets were included for genotyping using SNP markers at the *Rhg1* and *Rhg4* loci that were reported by Shi et al. (2015). SNP marker GSM381 was used for detecting the *rhg1* resistance allele at the *Rhg1* locus, then SNP marker GSM383 was used to distinguish between the 'Peking' (*rhg1-a*) or PI 88788 (*rhg1-b*) allele types at *Rhg1* locus. SNP marker GSM191 was used for identifying the resistance allele at *Rhg4* locus (Shi et al., 2015). The accessions were also genotyped with a SNP marker GSM039A on Chr10 (LG O)

for southern root-knot nematode resistance (Pham et al., 2013). Genotyping was performed using the protocol reported by Pham et al. (2013). Briefly, KASP reactions were run in a 4 µL reaction, which included 2 µL of diluted DNA, 2 µl of KASP master mix, and 0.106 µl primer mix. The PCR fluorescent end reading was performed using a Light Cycler 480 Real Time PCR system (Roche, Germany).

More than 40,000 SNPs of the 461 accessions were obtained from the SoySNP50K Infinium Chip data (source: soybase.org) and Soybean Breeding and Genetics Lab database at the UGA. There was no SNP data available for one accession (PI 670017). The genotyping results of KASP markers GSM381, GSM383, and GSM191 were also included for analysis. SNPs were further eliminated if they had no assigned physical position, if they had greater than 20% missing data, and if they and a minor allele frequency (MAF) less than 0.05. In total, 35, 817 SNPs met these criteria and were used to conduct GWAS. No SNP imputation was performed.

Cluster analysis based on haplotypes at *Rhg1* and *Rhg4* loci

Using the SoySNP50K Infinium Chip (soybase.org), two neighbor-joining trees based on genetic distances at *Rhg1* and *Rhg4* regions on Chrs 18 (LG G) and 8 (LG A2), respectively, were constructed using TASSEL software (Bradbury et al., 2007). The software calculates genetic distance between each genotype using a modified Euclidean distance, where homozygote is 100% similar to itself and 50% similar if heterozygote (Bradbury et al., 2007). Then, using neighbor joining algorithm to create phylogenetic trees, results were visualized using Figtree software (Rambaut and Drummond, 2009). Based on phylogenetic tree outputs, haplotype groups with 100% similarity were placed in the ‘Peking’ group and the PI 88788 group at *Rhg1* and *Rhg4* loci.

Rhg1 and *Rhg4* are two major effect loci that provide resistance to SCN race 3. At the *Rhg1* locus on Chr 18 (LG G), three genes are known to contribute to SCN resistance (*Glyma.18g022400, Glyma.18g022500* and *Glyma.18g022700*) (Cook et al., 2012). Based on the soybean reference genome sequence version 2.0 of 'Williams 82', approximately 500 kb flanking both sides of these three genes were selected for analysis. The selected 990-kb region, which included the *Rhg1* locus at Chr18 (LG G), consisted of 103 SNPs.

At the *Rhg4* locus on Chr 8 (LG A2), the *serine hydroxymethylytransferase* (*SHMT*) gene was attributed with conveying SCN resistance (Liu et al., 2012). Two SNPs (ss715602757 and ss715602764) are situated close to this gene, thus a 0.5 Mb region (based on the soybean reference genome sequence version 2.0 of 'Williams 82') flanking both sides of these SNPs were selected for analysis, which consisted of 64 SNPs.

Genome-wide association analysis

The phenotyping data from the cyst counts were subjected to analysis of variance (ANOVA). A split-plot analysis for a mixed linear model was applied with blocks and sets being treated as random effects, while accession was considered as a fixed effect. R packages (cran.r-project.org) and JMP software (SAS, 2016) were used to conduct ANOVA (Table 2.6). No transformation of phenotype data was used prior to GWAS analysis.

The genetic diversity of the 461 plant accessions was analyzed using a principal component analysis (PCA) in the GAPIT R package (Lipka et al., 2012) and neighbor-joining (NJ) tree using TASSEL software (Bradbury et al., 2007). The phylogenetic trees from TASSEL software were visualized with Figtree software (Rambaut and Drummond, 2009). The LD analysis was estimated using squared allele frequency correlation for pairs of SNPs from TASSEL software (Bradbury et al., 2007).

Two statistical analysis packages including GAPIT (Lipka et al., 2012) and FarmCPU (Liu et al., 2016) were used to conduct GWAS. In GAPIT, a compressed Mixed Linear Model (cMLM) (Zhang et al., 2010) using the first five PCs and a kinship matrix to control Type I errors (false positives) were based off population structures. For the cMLM model, the equation used was: $y = \mu + X\alpha + P\beta + Zu + e$, where y is the phenotypic genetic value; μ is the grand mean; X is the matrix coefficient to the fixed marker effects α; P is the matrix coefficient related to fixed principal component (PC) effects β, and Z is the matrix coefficient related to random group effect u received from compression algorithm. The threshold of significant value of association was False Discovery Rate-adjusted *P* value ($P < 0.001$). In the FarmCPU package, the fixed and random model Circulating Probability Unification (FarmCPU) has two parts: a fixed effect model for markers and a random effect model for kinship. The first five PCs from GAPIT were used as covariates. The threshold *P* value (5.773946e-07) was calculated using the parameter "p.threshold =0.05/number of markers" after 1,000 permutations. The quantile-quantile (Q-Q) plot from all models was used for to evaluate how fit models explained population structure.

Candidate genes of SCN resistance

Significant SNPs at the *Rhg1* and *Rhg4* loci are situated on Chrs 18 and 8, respectively, were not included for candidate gene prediction in this study because resistance genes at those two genomic regions were previously cloned (Cook et al., 2012 ; Liu et al., 2012). Here, we focus on genes at the QTLs on Chrs 2, 7 and 10 for a haplotype analysis using Haploview software (Figure 2.7) (Barret et al., 2005). Genes located within each haplotype block included a significant SNP associated with SCN resistance were selected as possible candidate genes. If a given SNP did not locate within a haplotype block, then gene models located within a 50 kb

upstream or downstream segment of that SNP were considered. The protein sequences encoded by the predicted genes were retrieved from the 'Williams 82' soybean reference genome on SoyBase (Soybase.com). Two criteria were used to predict candidate genes responsible for SCN resistance: 1) if a gene was implicated as a resistance gene providing disease resistance for nematodes or other pathogens in previous studies; and 2) if genes were located at genomic regions where the peak SNPs were placed as a result of the GWAS analysis. The gene models without functional annotations or belonging to unknown functional families were excluded.

Results

Greenhouse screening for SCN resistance

Of the 535 soybean accessions that were screened, 462 accessions were reported for phenotyping results (Table 2.2 and 2.4) based on the selection criteria described above, which precluded the use of 73 accessions due to low germination and insufficient root system. The standard susceptible check 'Lee 74' had a constant mean number of cysts in the first (41 cysts) and third (44 cysts) sets, but had a low mean number in the second set (23 cysts). We speculate that the low mean cyst number in the second set was probably due to higher than normal greenhouse temperatures (>30^{o}C), which caused low SCN cyst counts due to poor soybean development. Similarly, the overall mean number of cysts counted on the second set were low as well.

Of the 462 accessions, seven indicator lines, seven checks, and 90 accessions were rated as resistant (R) with a calculated FI less than 10 %. Additionally, 56 accessions were rated as moderately resistant (MR) with a calculated FI between 10-30% and 170 accessions were rated as moderately susceptible (MS) with FI between 30-60%. Finally, 146 soybean accessions were rated as highly susceptible (S) with calculated FI greater than 60% (Table 2.2). Of the 146

resistant and moderately resistant accessions, 82 accessions were from China, 31 breeding lines and cultivars were from the U.S., and 32 accessions were from six other countries.

SNP marker genotyping and haplotype analysis

Based on the genotype results of KASP marker GSM381 at the *Rhg1* locus (G: resistance allele) (Figure 2.1), 90 accessions were predicted to be resistant to SCN race 3. According to genotype calls for marker GSM383 at *Rhg1* locus (G: 'Peking' type; C: PI 88788 type), 54 of 90 accessions were grouped as carrying 'Peking-type resistance, whereas 36 accessions were grouped as carrying PI 88788-type resistance. Combined with the phenotyping results, 88 of 90 accessions were rated as resistant or moderately resistant using the FI index except two accessions (PI 578376 and PI 398823), which were rated as susceptible. We speculate that this discrepancy is due to PI 398823 belonging to 'Peking' group at *Rhg1* with low copy number and not carrying a resistance allele at *Rhg4*. The previous studies demonstrated that without the *Rhg4* resistance allele, the 'Peking' type *rhg1* resistance might not function efficiently for SCN resistance (Cook et al., 2014; Jiao et al., 2015). The PI 578376 had female index 31.8 % which is very close to moderately resistance.

At the *Rhg4* locus, 59 soybean accessions carried resistance alleles based on GSM191 (G: resistance allele) genotyping results but three of the soybean accessions were not resistant based upon greenhouse screening tests. No resistance allele at *Rhg1* locus was found in these three accessions based on genotyping results of the GSM381 marker. Interestingly, 58 soybean accessions classified as being resistant and moderately resistant from greenhouse screening results did not carry either the *rhg1* or the *Rhg4* resistance alleles, suggesting that they might contain novel alleles for SCN resistance (Table 2.3). A large number (n = 316) of accessions rated moderately susceptible and susceptible in the greenhouse screening assay (including the six

susceptible checks) matched the expected genotyping results by carrying susceptible alleles at both the *Rhg1* and *Rhg4* loci.

Furthermore, 462 soybean accessions were genotyped using a functional SNP marker GSM039A (Pham et al. 2013) for the southern root-knot nematode (*Meloidogyne incognita*) resistance, one of the most damaging pests in Georgia. The results indicated that 135 accessions were carrying the desirable allele providing resistance to this nematode. Of those 135 accessions, 58 accessions were rated as resistant to SCN based on the greenhouse screening assay and of those 58 accessions, 18 accessions appear to be carrying novel genes or alleles conveying SCN resistance. These 58 accessions will be further subjected to a greenhouse screening for southern root-knot nematode resistance.

Using SNP data from SoySNP50K Infinium Chip for these 461 soybean accessions, the cluster analysis based on 103 SNPs from a 990-kb region including the *Rhg1* locus on Chr 18 (LG G) separated soybean accessions carrying PI 88788-type (23 accessions) and 'Peking-type (50 accessions) alleles into two groups. Two indicator lines, PI 209332 and PI 548316, were located in the same group as PI 88788, whereas PI 437654, PI 89772, and PI 90783 were grouped with 'Peking', which was consistent with a previous study (Cook et al., 2012). Compared to the genotyping results generated from SNP markers GSM381 and GSM383 at the *Rhg1* locus, 82% of the resistant accessions were located in two groups that included 'Peking' and PI 88788 from SoySNP50K data. A discrepancy was observed between our *rhg1* markers and haplotype analysis for 16 accessions. According to haplotype analysis, three accessions (PI 438489B, PI 567378, and PI 531068) were clustered in the 'Peking' group but were classified as carrying the PI 88788-resistance allele type according to genotyping results with two *rhg1* markers (GSM381 and GSM383). Eleven and two accessions belonged to PI 88788 and 'Peking'

groups, respectively, based on the genotyping results, but they were not included in the PI 88788 and 'Peking' groups using the haplotype analysis.

Cluster analysis based on the SNPs from a 997-kb region at *Rhg4* locus on Chr 8 placed 26 accessions into the 'Peking' group, which was lower than the indicated genotyping results using SNP marker GSM191 at the *Rhg1* locus (59 accessions). Of these 59 accessions, 30 accessions were placed in one cluster group and three were placed in a different cluster group, which were separated from the 'Peking' cluster. Three accessions, PI 438489B, PI 417092, and PI 416762, were grouped with 'Peking', but based on *Rhg4* SNP marker genotyping (GSM 191), they did not carry the 'Peking'-type resistance allele at *Rhg4*. The difference between haplotype analysis and functional markers might be that no informative SNPs at *Rhg1* and *Rhg4* loci were present in SoySNP50K Chip data.

Interestingly, 58 accessions that were identified in greenhouse screening assays for SCN resistance and predicted not carrying resistance alleles using three functional SNP markers: GSM381; GSM383 and GSM191, were grouped separately from PI 88788 and 'Peking' at both *Rhg1* and *Rhg4* loci. Although there were three different clusters of soybean accessions having the *Rhg4*-resistance allele based on phylogenetic tree analysis at the *Rhg4* region, these 58 unique accessions were placed in different clusters (Figure 2.2). Based on SNP markers and haplotype analysis at *Rhg1* and *Rhg4* loci, these 58 accessions might possess different resistance alleles or different backgrounds conferring SCN resistance than 'Peking' or PI 88788.

LD analysis and population structure

To understand the genetic architecture for SCN resistance and genetic relationship of 461 soybean accessions, more than 40,000 SNP data for 461 soybean accessions were retrieved from SoySNP50K Infinium Chip data (source: soybase.org). Three functional SCN markers:

GSM381; GSM383 and GSM191 at *Rhg1* and *Rhg4* loci were included. After filtering SNPs with MAF > 0.05 and missing data < 20%, a total of 35,817 SNPs were used for further analyses. The selected SNP markers ranged from 1,301 on Chr 12 (LG H) to 2,890 on Chr 18 (LG G), with an average count of 1,790 SNPs per chromosome. The SNPs occupied a 19.7 kb range on Chr 13 (LG F) and a 38.7 kb range on Chr 1 (LG D1a), with an average of one SNP per 27.20 kb. The recombination rate impacts the resolution of association mapping, and is estimated by LD decay rates. The LD decay rate was measured by the distance of the average pairwise correlation coefficient dropped to half of its maximum value. Here, the LD decay distance was estimated to be about 125 kb (Figure 2.3). LD decay showed a decreasing trend, suggesting that high recombination occurred among these soybean accessions.

The resulting tree showed seven groups based on the distribution of origins around the globe. Most of U.S. cultivars were placed in the same group, demonstrating that they were developed from limited ancestries. Two separate clusters were generated that represented accessions from Japan and South Korea. Due to large number of accessions from China (60%), Chinese accessions were widely spaced. Accessions from other origins did not locate to unique clusters. The results indicate that place of origin and population structure were in general correlated. In addition, plant maturity as dictated by MG was included in the population structure analysis, and the accessions did not appear to be clustered by maturity group. PCA showed similar results using the neighbor joining tree-accessions test, revealing that the accessions did not tend to cluster based on MG (Figure 2.5). Our results agree with a previous GWAS study, which indicated no correlation based on MG (Zhang et al., 2016). A heatmap showing the low level of relatedness among the 461 accessions is depicted in Figure 2.4.

GWAS for SCN resistance

Based on the analysis using both the cMLM in the GAPIT R package and the FarmCPU model, as determined by interpreting the quantile-quantile (Q-Q) plots, two different models adequately controlled genomic inflation, so both models were incorporated into the GWAS results (Figure 2.6). A total of 14 SNPs on Chrs 2, 7, 8, 10, and 18 were found to be associated with HG Type 0 (race 3) SCN resistance using the GWAS analysis. Of the 14 SNPs identified, 12 SNPs were identified using cMLM and two SNP was detected using the FarmCPU model (Figure 2.6; Table 2.7). Both models were able to detect a SNP with SNP marker GSM381 at the *Rhg1* locus on Chr 18 with the highest peak level of significance (GAPIT = 1.24E-17; FarmCPU = 3.18E-25). The other 12 SNPs identified using cMLM were as follows: one SNP on Chr 2, eight SNPs on Chr 7, one SNP on Chr 8, one SNP on Chr 10, and one SNP on Chr 18. SNP markers GSM383 and GSM191, located on Chrs 18 and 8 at known *Rhg1* and *Rhg4* regions, explained 34.1% and 26.2% of the phenotypic variation, respectively. Additionally, the genomic regions indicated by 10 SNPs on Chrs 2, 7, and 10 were associated with HG Type 0 resistance, which have not been reported in QTL analyses utilizing bi-parental mapping populations. On Chr 2, SNP marker ss715583938 could be considered as a minor effect QTL, due to the average FI value being 49.3%, which was slightly lower than average FI value calculated for the whole panel (49.8%) and also this SNP was barely above significant threshold p-value. Based on haplotype analysis, SNP marker ss715583938 is linked to 15 other SNPs, and associated with 31 resistant and moderately resistant accessions evaluated in our greenhouse assay (including 'Peking') that shared the same alleles in this linkage block on Chr 2. The eight significant SNPs identified on Chr 7 occupied 36.4-36.9 Mb, explained 25-27% of phenotypic variation. This genomic region for SCN resistance on Chr 7 overlapped with a previously recognized region

(rs36423980: 36.4 Mb) associated with HG Type 2.5.7 (SCN race 1) resistance reported in another GWAS study (Zhao et al., 2017). The ss715606985 SNP on Chr 10 (resistance allele = GG) accounted for 25.6% of the phenotypic variance with the an average FI index of 3.8%, which was significantly lower than the average FI index of the whole panel (49.76%). Fifty-five resistant and moderately resistant accessions identified using the greenhouse screening assay carried this resistance allele on Chr 10 based on GWAS. Interestingly, all of these 55 accessions, including known resistant sources 'Peking', PI 89772, PI 437654, and PI 90763, carried resistance alleles at the *Rhg1* locus based on GSM381 SNP marker allele calls. To date, no QTLs for SCN resistance at this region on Chr 10 have been reported using the populations derived from these known resistant sources.

Candidate genes and ontologies for SCN resistance

A total of 33 gene models were predicted using 10 significant SNPs on Chrs 2, 7, and 10 that were significantly associated with HG Type 0 resistance (Table 2.8). Based on haplotype analysis, seven of 10 significant SNPs located within four haplotype blocks associated with SCN resistance: one block on Chr 2 occupying a 144 kb region and three blocks (168 kb, 109 kb, and 6 kb) on Chr 7. The three significant SNPs on Chr 7 and 10 did not fall in any haplotype blocks but after scanning a 50 kb region flanking the significant SNPs were included in the 12 gene models. Based on information of gene ontology from SoyBase (Soybase.com), possible candidate genes were as follows: (1) a leucine rich repeat (LRR) protein kinase family protein gene, (2) a cytochrome P450 family protein gene, (3) a RING/ U-box protein gene, (4) a DNA synthesis gene, (5) a transcription regulation gene and (6) some in miscellaneous groups. Of these listed candidate genes and ontologies, some of them occupy domains on Chr 7 where R

genes have previously been categorized, such as a (LRR) receptor gene (*Glyma.07g199500*), a cytochrome P450 gene (*Glyma.07g194400*), and a RING gene (*Glyma.07g196000*).

Discussion

Screening for SCN resistant soybeans began as early as 1957 (Ross, 1957). However, to date soybean breeders have relied on just two primary resistant sources ('Peking' and PI 88788) to develop SCN resistant cultivars. In some cases, other SCN resistant lines have been identified, but they offer defense mechanisms that fall into 'Peking'-type or PI88788-type resistance based on major resistance QTL that they possess (Concibido et al., 2004). The 'Peking'-type of SCN resistance requires resistance alleles at both the *Rhg1* and *Rhg4* loci but PI 88788-type resistance requires only one preferable allele at *Rhg1* (Concibido et al., 2004). In the past 10 years, some of SCN populations have overcome the PI 88788-type resistance (Niblack et al., 2008). In Georgia, according to the HG Type test, a new population was found that was designated as HG Type 5 (the indicator line #5: PI 209332 has susceptible reaction) and was more aggressive than HG Type 0 (race 3). Based on greenhouse screening of 462 soybean accessions, our study was able to identify 58 accessions that were rated as resistant and moderately resistant to HG Type 0 (race 3). Of those, 24 accessions were rated as resistant or moderately resistant to the aggressive HG Type 5 race in our confirmation screening test. Haplotype analysis of these 58 accessions based upon SNPs at *Rhg1* and *Rhg4* regions indicated that that these 58 accessions were not clustered into 'Peking'-type or PI 88788-type clusters.

The origin of the top five of these 58 accessions (PI 574484, PI 567403B, PI 561334, PI 603529, and PI 561329) was China, and each was characterized as being highly resistant to aggressive SCN HG Type 5. Four of them except PI 603529 possess desirable agronomic and consumer traits, such as yellow seed coat, low lodging score (≤ 3 in scale 5), and low shattering

(< 2 in scale 5) (USDA-GRIN). Based on genotyping results using SNP marker GSM039A, PI 561329 is also predicted to have resistance to southern root-knot nematode, as were 18 out of the 58 SCN resistant accessions we identified. Therefore, these accessions could be valuable sources of resistance to two different nematode species, SCN (*Heterodera glycines* Ichinohe) and southern root-knot nematode (*Meloidogyne incognita*), respectively.

To understand genetic basis of SCN resistance in soybean, QTL mapping efforts have identified more than 30 QTLs on 17 of 20 soybean chromosomes (Concibidio et al., 2004; Guo et al., 2006; Kim et al., 2016). Among them, two major effect resistance QTLs on Chr 8 (*Rhg4*) and 18 (*Rhg1*) have been reported (Concibido et al., 2004). Recently, association mapping was conducted to explore the genetic architecture of soybean accessions to several traits including SCN resistance because of its advantages over linkage mapping, one of which is increased mapping resolution (Korte and Farlow, 2013). However, one major problem using GWAS is spurious association caused by population structure and family relatedness when allele frequency differences are caused by different origins or maturity groups sharing the same ancestry (Korte and Farlow, 2013). In this GWAS study, a total of 461 soybean accessions from a wide range of maturity groups from 28 different countries were used. Based on population structure analysis, geographical diversification was slightly correlated, but no correlation was detected by MG. There are a number of models used in GWAS that include a PCA and kinship matrix to effectively control population stratification. However, depending on the statistical model used the results may differ, thus a Q-Q plot analysis is conducted to identify the model that best suits the data. In our analysis, multiple statistical models (GLM, cMLM, ECMLM and FarmCPU) were utilized to detect and verify genomic regions controlling SCN resistance (Figure 2.6). Based on Q-Q plot, all models were suitable to analyze. Four significant SNPs (GSM381 on Chr

18 and ss715597494, ss71559749, and ss715597431 on Chr 7) were identified using all GWAS models with the exception of FarmCPU. It indicated the importance of these four significant SNPs for HG Type 0 resistance. Additionally, three other QTLs on Chrs 2, 8, and 10, respectively, were detected using two statistical models (GLM and cMLM). The genomic region (GSM191) on Chr 8 is associated with major resistance locus *Rhg4*. The genomic region at ss715606985 on Chr 10 might overlap with another known QTL (*qSCN10*, identified from PI 567516C), although this SNP was located 1 Mb from the previously mapped QTL (Vuong et al.,2010; Vuong et al., 2015). At this SNP, it was revealed that 55 of the resistant and moderately resistant accessions, including PI 567516C, carried the same resistance allele (A). Interestingly, some known sources of SCN resistance, including 'Peking', PI 89772, PI 90763, and PI 437654, also carried the same resistance alleles as PI 567516C at this significant SNP: ss715606985 although no QTL controlling SCN resistance on Chr 10 was reported in bi-parental mapping populations derived from 'Peking', PI 89772, PI 90763, and PI 437654. Furthermore, the results of the greenhouse screening indicated that the average FI of accessions carrying the resistance allele at the ss715606985 on Chr 10 was significantly low (3.8%), suggesting that this SNP marker could be located in the genomic regions responsible for SCN resistance. Additionally, the region on Chr 2 with significant SNP ss715583938 has not been previously reported. The result showed the limitation of few recombination events from two parents in QTL mapping while GWAS captures more allelic diversity from a larger panel of accessions, so can detect more genomic regions than bi-parental population.

To date, the two primary candidate genes controlling SCN resistance at the loci *Rhg1* and *Rhg4* have been cloned. However, it is necessary to discover more candidate genes conferring SCN resistance in order to enhance resistance as well as have alternative solutions when *rhg1*

and *Rhg4* might lose their effectiveness. Based on a single marker or haplotype, GWAS has identified new significant genetic regions associated with SCN resistance, and then predicted new candidate genes within these regions besides *Rhg1* and *Rhg4*. After all, due to the large number of gene models that occupy the haplotype blocks surrounding the significant SNPs that were identified using GWAS, it was difficult to specify the candidate genes conferring SCN resistance. However, based on the gene ontologies in SoyBase near the significant SNPs identified using GWAS, we propose that 33 different gene models on Chrs 2, 7, and 10 may play a role in providing SCN resistance. Of them, three belong to chromosomal domains with previously reported plant disease resistance (R) genes, such as LRR, cytochrome 450 and RING/U box. (Han et al., 2015), but further work is needed to confirm if these genes might be responsible for conveying resistance to SCN

Conclusions

With overuse of resistance genes from 'Peking' and PI 88788 for development of SCN resistant cultivars, it has led to a shift of SCN race populations that have overcome the resistance mechanisms associated with the *rhg1* and *Rhg4* alleles. Therefore, it is of paramount importance to identify new sources of SCN resistance with different genetic backgrounds than these two overused sources. In this study, 461 soybean accessions from various origins and maturity groups were evaluated in a greenhouse screening assay using a HG Type 0 (SCN race 3) inoculant, which is the predominant SCN race limiting soybean production. It was found that 146 soybean accessions tare resistant to HG Type 0. Using SNP markers for *rhg1* (GSM381 and GSM383) and *Rhg4* alleles loci (GSM191) and haplotype analysis, 58 resistant soybean accessions were found that did not possess the same resistance alleles as 'Peking' and PI 88788

at *Rhg1* and *Rhg4* loci, suggesting these accessions could carry novel alleles conferring SCN resistance.

Using these 461 accessions, a GWAS analysis was conducted using a high density of 35,817 SNPs generated from SoySNP50K Infinium Chips coupled with the phenotypes from greenhouse screening assays to detect genomic regions associated with SCN resistance. The results identified 13 significant SNPs at five genomic regions on Chrs 2, 7, 8, 10, and 18 were highly significantly associated with HG Type 0 (SCN race 3) resistance. Two of those genomic regions correspond to *Rhg1* and *Rhg4* regions on Chrs 8 and 18, respectively and two genomic regions are situated on Chrs 7 and 10 that explained 26% of phenotypic variation and overlapped or were close to regions previously identified for SCN resistance from other GWAS studies. The QTL on Chr 2 have not been reported and located ~5 Mb away from previous region reported in GWAS mapping (Vuong et al., 2015). Further studies are needed to confirm these QTLs. This GWAS study has shed light on the genetic architecture underlying the quantitative basis for SCN resistance and will allow soybean breeders to use these resistant sources in their breeding programs.

References

Bao, Y., T. Vuong, C. Meinhardt, P. Tiffin, R. Denny, S. Chen, et al. 2014. Potential of Association Mapping and Genomic Selection to Explore PI 88788 Derived Soybean Cyst Nematode Resistance. The Plant Genome 7: 0. doi:10.3835/plantgenome2013.11.0039.

Barret, J.C., F.B. Maller, M.J. Daly. 2005 Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21 (2): 263-265.

Bastien, M., H. Sonah and F. Belzile. 2014. Genome Wide Association Mapping of Resistance in Soybean with a Genotyping-by-Sequencing Approach. The Plant Genome 7: 0. doi:10.3835/plantgenome2013.10.0030

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss and E.S. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633-2635. doi:10.1093/bioinformatics/btm308.

Concibido, V.C., B.W.Diers and P.R.Arelli. 2004. A decade of QTL mapping for cyst nematode resistance in soybean. Crop Sci. 44: 1121-1131.

Cook, D.E., A.M. Bayless, K. Wang, X. Guo, Q. Song, J. Jiang, et al. 2014. Distinct Copy Number, Coding Sequence, and Locus Methylation Patterns Underlie *Rhg1*-Mediated Soybean Resistance to Soybean Cyst Nematode. Plant Physiol. 165: 630-647. doi:10.1104/pp.114.235952.

Cook, D.E., T. G. Lee et al. 2012. Copy number variation of multiple genes at rhg1 mediates nematode resistance in soybean. Science 338: 1206-1209.

Glover, K.D. 2004. Near isogenic lines confirm a soybean cyst nematode resistance gene from PI 88788 on linkage group J. Crop Sci. 44: 936-941.

Guo, B., D.A. Sleper, H.T. Nguyen, P.R. Arelli and J.G. Shannon. 2006. Quantitative Trait Loci underlying Resistance to Three Soybean Cyst Nematode Populations in Soybean PI 404198A. Crop Sci. 46: 224. doi:10.2135/cropsci2004.0757.

Han, Y., X. Zhao, G. Cao, Y. Wang, Y. Li, D. Liu, et al. 2015. Genetic characteristics of soybean resistance to HG type 0 and HG type 1.2.3.5.7 of the cyst nematode analyzed by genome-wide association mapping. BMC Genomics 16: 598. doi:10.1186/s12864-015-1800-1.

Hassan, S.M. 2013. Soybean, Nutrition and Health. Licensee InTech. doi:10.5772/54545.

Hymowitz, T. 1970. On the domestication of the Soybean. Econ. Bot 24: 408-421.

Ichinohe, M. 1959. Studies on the soybean cyst nematode *Heterodera glycines* and its injury to soybean plants in Japan. Plant Pis. Reptr. Suppl. 260: 238.

Jackson, G. 2014. Soybean cyst nematode. Mississippi State University Extension Service Publication 1293.

Jiao, Y., T.D. Vuong, Y. Liu, Z. Li, J. Noe, R.T. Robbins, et al. 2015. Identification of quantitative trait loci underlying resistance to southern root-knot and reniform nematodes in soybean accession PI 567516C. Mol Breed 35: 131. doi:10.1007/s11032-015-0330-5.

Jun, T. H., K. Van, M.Y. Kim, S.-H. Lee and D.R. Walker. 2008. Association analysis using SSR markers to find QTL for seed protein content in soybean. Euphytica 162: 179-191. doi:10.1007/s10681-007-9491-6.

Kazi, S., J. Shultz, J. Afzal, R. Hashmi, M. Jasim, J. Bond, et al. 2010. Iso-lines and inbred-lines confirmed loci that underlie resistance from cultivar 'Hartwig' to three soybean cyst nematode populations. Theor Appl Genet. 120: 633-644. doi:10.1007/s00122-009-1181-4.

Keim, P., T. Olson and R. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. Soybean genetics newsletter-United States Department of Agriculture, Agricultural Research Service (USA).

Kim, M. and B.W. Diers. 2013. Fine Mapping of the SCN Resistance QTL and from PI 468916. Crop Sci. 53: 775. doi:10.2135/cropsci2012.07.0425.

Kim, M., D.L. Hyten, A.F. Bent and B.W. Diers. 2010. Fine Mapping of the SCN Resistance Locus *rhg1-b* from PI 88788. The Plant Genome. 3: 81. doi:10.3835/plantgenome2010.02.0001.

Kim, M.S., M.K. Sung, M.W. Kim, H.J. Seo, D.G. Kim and J.I. Chung. 2013. Investigation of Soybean Cyst Nematode *Heterodera Glycines* Type and Evaluation of Resistance on Soybean Varieties and Germplasms in Korea. Korean J. Crop Sci. 58: 161-168. doi:10.7740/kjcs.2013.58.2.161.

Koenning, S.R. and J.A. Wrather. 2010. Suppression of Soybean Yield Potential in the Continental United States by Plant Diseases from 2006 to 2009. Plant Health Prog. doi:10.1094/php-2010-1122-01-rs.

Korte, A., and A. Farlow. 2013. The advanatges and limitations of trait analysis with GWAS: a review. Plant Methods 9: 29.

Li, Y.H., M.J.M. Smulders, R.Z. Chang and L.J. Qiu. 2011. Genetic diversity and association mapping in a collection of selected Chinese soybean accessions based on SSR marker analysis. Conserv. Genet. 12: 1145-1157. doi:10.1007/s10592-011-0216-y.

Lipka, A.E., F. Tian, Q. Wang, J. Peiffer, M. Li, P.J. Bradbury, et al. 2012. GAPIT: genome association and prediction integrated tool. Bioinformatics 28: 2397-2399. doi:10.1093/bioinformatics/bts444.

Liu, S., P.K. Kandoth, S.D. Warren, G. Yeckel, R. Heinz, J. Alden, et al. 2012. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. Nature 492: 256-260. doi:10.1038/nature11651.

Liu, X. 1997. History and status of soybean cyst nematode in China. Int. J. Nematol. 7: 18-25.

Liu, X., M. Huang, B.Fan, E. S. Bucker, Z. Zhang. 2016. Iterative usage of fixed and random effect models for powerful and efficient genome-wide assocociation studies. PLoS Genet. 12(2): e1005767. doi: 10.1371/journal.pgen.1005767

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

Melito, S., A.L. Heuberger, D. Cook, B.W. Diers, A.E. MacGuidwin and A.F. Bent. 2010. A nematode demographics assay in transgenic roots reveals no significant impacts of the *Rhg1* locus LRR-Kinase on soybean cyst nematode resistance. BMC Plant Biol. 10: 104. doi:10.1186/1471-2229-10-104.

Mitchum, M.G., J.A. Wrather, R.D. Heinz, J.G. Shannon and D. Gene. 2007. Variability in distribution and virulence phenotypes of *Heterodera glycines* in Missouri during 2005. Plant Dis. 91: 1473-1476.

Moore, W.F et al. 1981. Soybean cyst nematode. [United States] : Prepared and distributed under the auspices of the Soybean Industry Resource Committee, [1981].

NASS. 2017. Crop Prod 2015 Summary. USDA.

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

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

Niblack, T.L., K. N. Lambert, and G. L. Tylka. 2006. A Model Plant Pathogen from the Kingdom Animalia: *Heterodera glycines*, the Soybean Cyst Nematode. Annu. Rev. Phytopathol. 44: 283-303. doi:10.1146/.

Pham, A.T., K. McNally, H. Abdel-Haleem, H. Roger Boerma and Z. Li. 2013. Fine mapping and identification of candidate genes controlling the resistance to southern root-knot nematode in PI 96354. Theor Appl Genet. 126: 1825-1838. doi:10.1007/s00122-013-2095-8.

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

Rambaut, A. and A. Drummond. 2009. FigTree v1. 3.1[computer program] http://tree.bio.ed.ac.uk.

Riggs, R.D. 1975. Worldwide distribution of soybean-cyst nematode and its economic importance. Ann. Meet. Nematol. & Am. Phytopathol. Soc. Houston, Texas, U.S.A.

Riggs, R.D. and D.P.Schmitt. 1988. Complete characterization of the race scheme for *Heterodera glycines*. J. Nematol. 20: 392-395.

Ross, J.P. 1957. Resistance of soybeans to the soybean cyst nematode as determined by a double-row method. Plant Dis. Rep. 41: 923-924.

Schmitt, D.P. and G. Shannon. 1992. Differentiating Soybean Responses to *Heterodera Glycines* Races. Crop Sci.32: 275-277. doi:10.2135/cropsci1992.0011183X003200010056x.

Shi, Z., S. Liu, J. Noe, P. Arelli, K. Meksem and Z. Li. 2015. SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. BMC Genomics 16: 314. doi:10.1186/s12864-015-1531-3.

Tylka, G.L. and C.C. Marett. 2014. Distribution of the Soybean Cyst Nematode, *Heterodera glycines*, in the United States and Canada: 1954 to 2014. Plant Health Progress. doi:10.1094/php-br-14-0006.

Tylka, G. L., and M. P. Mullaney. 2016. Soybean cyst nematode-resistant soybeans for Iowa. Ext. Punl. PM 1649. Iowa State Univ., Ames.

Vaughn, Justin N., Randall L. Nelson, Qijian Song, Perry B. Cregan, and Zenglu Li. 2014. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. G3: Gene, Genomes, Genetics. 4:2283-2294. doi:10.1534/g3.114.013433

Vuong, T.D., D.A. Sleper, J.G. Shannon and H.T. Nguyen. 2010. Novel quantitative trait loci for broad-based resistance to soybean cyst nematode (*Heterodera glycines*) in soybean PI 567516C. Theor Appl Genet. 121: 1253-1266. doi:10.1007/s00122-010-1385-7.

Vuong, T.D., H. Sonah, C.G. Meinhardt, R. Deshmukh, S. Kadam, R.L. Nelson, et al. 2015. Genetic architecture of cyst nematode resistance revealed by genome-wide association study in soybean. BMC Genomics 16: 593. doi:10.1186/s12864-015-1811-y

Wang, D., P. R. Arelli, R. C. Shoemaker, and B. W. Diers. 2001. Loci underlying resistance to race 3 of SCN in *Glyine soja* plant introduction 468916. Theor Appl Genet. 103.

Wang, J., P.E. McClean, R. Lee, R.J. Goos and T. Helms. 2008. Association mapping of iron deficiency chlorosis loci in soybean (*Glycine max* L. Merr.) advanced breeding lines. Theor Appl Genet. 116: 777-787. doi:10.1007/s00122-008-0710-x.

Wang, X., G. Bergstrom, S. Chen, D. M. Thurston, J. Cummings, Z. Handoo, et al. 2017. First Report of the Soybean Cyst Nematode, *Heterodera glycines*, in New York. The Am. Phytopathol. Soc. 101 (11): 1957

Wrather, A., G. Shannon, R. Balardin, L. Carregal, R. Escobar, G.K. Gupta, et al. 2010. Effect of Diseases on Soybean Yield in the Top Eight Producing Countries in 2006. Plant Health Prog. doi:10.1094/php-2010-0125-01-rs.

Zhang, H., C. Li, E.L. Davis, J. Wang, J.D. Griffin, J. Kofsky, et al. 2016. Genome-Wide Association Study of Resistance to Soybean Cyst Nematode (*Heterodera glycines)* HG Type 2.5.7 in Wild Soybean (*Glycine soja*). Front Plant Sci. 7: 1214. doi:10.3389/fpls.2016.01214.

Zhang, H., Q. Song, J.D. Griffin and B.H. Song. 2017. Genetic architecture of wild soybean (*Glycine soja*) response to soybean cyst nematode (*Heterodera glycines*). Mol Genet Genomics. doi:10.1007/s00438-017-1345-x.

Zhang, Z., E. Ersoz, C.-Q. Lai, R.J. Todhunter, H.K. Tiwari, M.A. Gore, et al. 2010. Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42: 355-360.

Zhao, X., W. Teng, Y. Li, D. Liu, G. Cao, D. Li, et al. 2017. Loci and candidate genes conferring resistance to soybean cyst nematode HG type 2.5.7. BMC Genomics. 18: 462. doi:10.1186/s12864-017-3843-y.

Zheng, J., Y. Li., and S. Chen. 2006. Characterization of the virulence phenotypes of *Heterodera glycines* in Minnesota. J. Nematol. 38 (3): 383-390.

Zhu, C., M. Gore, E.S. Buckler and J. Yu. 2008. Status and Prospects of Association Mapping in Plants. The Plant Genome. 1: 5. doi:10.3835/plantgenome2008.02.0089.

Table 2.1 Soybean accessions (n = 535) evaluated with SCN HG Type 0 (race 3) for SCN resistance in greenhouse.

Set[a]	No of accessions	Planting date	Inoculation date
1	82	3rd Feb, 2016	10th Feb, 2016
2	204	6th June, 2016	15th June, 2016
3	249	22nd Aug, 2016	2nd Sep, 2016
Confirmation	106	12th Dec, 2016	22nd Dec, 2016

[a] The soybean accessions were divided into three sets for greenhouse screening. The confirmation set (n = 106) were selected based on the previous screening results.

Table 2.2 Summary of greenhouse screening results for SCN resistance.

Rating	Female index[a] (%)	No. of accessions
Resistant	0-10	90
Moderately resistant	10-30	56
Moderately susceptible	30-60	170
Susceptible	>60	146

[a] The level of resistance was defined by the female index (FI) that was calculated based on the ratio between the mean numbers of cysts on a given line and 'Lee 74', reported as a percentage (Niblack et al., 2002).

Table 2.3 Summary of genotyping and phenotyping results for 462 soybean accessions that were screened for SCN race 3 resistance.

Resistance alleles	Reaction to SCN race 3	No of lines	GSM381 *rhg1*	GSM383 *rhg1*	GSM191 *Rhg4*
rhg1('Peking' type) + *Rhg4*	RES	50	GG [a]	GG	GG
rhg1 only ('Peking' type)	RES/SUS	4	GG	GG	CC
rhg1 (PI 88788 type) + *Rhg4*	RES	6	GG	CC	GG
rhg1 only (PI 88788 type)	RES	30	GG	CC	CC
Rhg4 only	SUS	3	TT	CC/GG	GG
Absence of *rhg1* and *Rhg4*	RES	58	TT	CC	CC
Absence of *rhg1* and *Rhg4*	SUS	316	TT	CC	CC

[a] A genotype call of GG denotes the soybean accession carries resistance allele.

Table 2.4 Country of origin, maturity group (MG), greenhouse phenotypes and SNP genotyping results of 455 soybean accessions that were screened with greenhouse screening assays for SCN race 3 resistance.

Accessions	Origin	MG	Female index (%)[a]	Greenhouse phenotype[b]	GSM381[c] (Chr 18)	GSM383 (Chr 18)	GSM191 (Chr 8)	Predicted phenotype
PI 531068	USA	VII	0	R	GG	CC	GG	R
PI 556949	China	IV	0	R	GG	GG	GG	R
PI 567386	China	VI	0	R	GG	GG	CC	R
PI 494182	Japan	0	0	R	GG	GG	GG	R
PI 461509	China	I	0	R	GG	CC	CC	R
PI 567491A	China	III	0	R	GG	GG	GG	R
PI 404198A	China	IV	0	R	GG	GG	GG	R
PI 548402	China	IV	0	R	GG	GG	GG	R
PI 084751	South Korea	IV	0	R	GG	GG	GG	R
PI 437725	China	IV	0	R	GG	GG	GG	R
PI 548974	USA	V	0	R	GG	CC	GG	R
PI 507423	Japan	VI	0	R	GG	GG	GG	R
PI 548988	USA	VI	0	R	GG	GG	GG	R
PI 591825	USA	VI	0	R	GG	CC	CC	R
PI 522236	USA	VII	0	R	GG	GG	GG	R
Woodruff	USA	VII	0	R	GG	GG	GG	R
PI 548665	USA	VIII	0	R	GG	GG	GG	R
PI 417091	Japan	II	0	R	GG	CC	CC	R

PI 507476	Japan	VI	0	R	GG	GG	GG	R
PI 468915	China	II	0	R	GG	GG	GG	R
PI 555453	USA	VII	0	R	GG	GG	GG	R
PI 468903	China	II	0	R	GG	GG	GG	R
PI 087631_1	Japan	III	0	R	GG	CC	CC	R
PI 553047	USA	VII	0	R	GG	GG	GG	R
PI 437690	China	III	0.6	R	GG	GG	GG	R
PI 437770	China	III	0.6	R	GG	CC	CC	R
PI 438498	USA	IV	0.6	R	GG	GG	GG	R
PI 404198B	China	IV	0.6	R	GG	GG	GG	R
PI 518772	USA	V	0.6	R	GG	GG	GG	R
PI 548970	USA	VIII	0.6	R	GG	GG	GG	R
PI 548982	USA	VI	0.8	R	GG	GG	GG	R
PI 670017	USA	V	0.9	R	GG	GG	GG	R
PI 595645	USA	VII	1.2	R	GG	GG	GG	R
PI 416762	Japan	II	1.2	R	GG	CC	CC	R
PI 438503A	USA	II	1.2	R	GG	CC	CC	R
PI 232993	Japan	II	1.2	R	GG	CC	CC	R
PI 404166	China	III	1.2	R	GG	GG	GG	R
PI 438496B	USA	III	1.2	R	GG	GG	GG	R
PI 438497	USA	III	1.2	R	GG	GG	GG	R
PI 495017C	China	IV	1.2	R	GG	CC	CC	R
PI 398682	South	IV	1.2	R	GG	CC	CC	R

	Korea							
PI 423927	Japan	IV	1.2	R	GG	GG	GG	R
PI 303652	China	V	1.2	R	GG	GG	GG	R
PI 548402S	China	V	1.2	R	GG	GG	GG	R
PI 507422	Japan	VI	1.2	R	GG	GG	GG	R
PI 548981	USA	VI	1.2	R	GG	CC	GG	R
PI 553040	USA	VI	1.8	R	GG	CC	GG	R
	South							
PI 509095	Korea	VII	1.8	R	GG	GG	GG	R
PI 574484	China	IV	2.0	R	TT	CC	CC	S
	South							
PI 398276	Korea	IV	2.4	R	GG	CC	CC	R
PI 602597	USA	VI	2.4	R	GG	GG	GG	R
PI 599333	USA	VI	2.4	R	GG	CC	CC	R
PI 612157	USA	VIII	2.4	R	GG	CC	GG	R
PI 540556	USA	II	2.5	R	GG	CC	CC	R
PI 467312	China	II	2.5	R	GG	CC	CC	R
PI 091102	China	II	2.5	R	GG	CC	CC	R
PI 089783	China	III	2.5	R	GG	CC	CC	R
PI 464912	China	IV	2.5	R	GG	CC	CC	R
PI 507470	Japan	VI	2.5	R	GG	GG	GG	R
PI 543855	USA	II	2.5	R	GG	GG	GG	R
PI 507475	Japan	V	3.1	R	GG	GG	GG	R

PI 567516C	China	IV	3.7	R	GG	GG	CC	R
PI 339868B	South Korea	IV	3.7	R	GG	GG	GG	R
PI 511813	USA	VI	3.7	R	GG	GG	GG	R
PI 437655	China	III	4.9	R	GG	CC	CC	R
PI 509100	South Korea	VII	4.9	R	GG	GG	GG	R
PI 507354	China	I	4.9	R	GG	GG	GG	R
PI 467327	China	II	5.5	R	GG	CC	CC	R
PI 559370	USA	V	5.5	R	GG	GG	GG	R
PI 458520	China	II	6.1	R	GG	CC	CC	R
PI 438489B	USA	IV	6.1	R	GG	CC	CC	R
PI 533605	USA	V	6.7	R	GG	GG	GG	R
PI 567403B	China	VII	6.9	R	TT	CC	CC	S
PI 567378	China	VI	7.2	R	GG	CG	CC	R
PI 507443	Japan	IV	7.4	R	GG	GG	GG	R
PI 467332	China	II	7.4	R	GG	CC	CC	R
PI 063468	China	IV	7.4	R	GG	CC	CC	R
PI 548655	USA	V	7.4	R	GG	GG	GG	R
PI 561334	China	IV	9.2	R	TT	CC	CC	S
PI 603529	China	VIII	9.2	R	TT	CC	CC	S
PI 561329	China	II	9.6	R	TT	CC	CC	S
PI 437679	China	IV	9.8	R	GG	GG	GG	R

PI 567285	China	IV	9.8	R	TT	CC	CC	S
PI 549077	China	0	10.0	MR	TT	CC	CC	S
PI 549075	China	00	10.3	MR	TT	CC	CC	S
PI 221717	South Africa	VI	10.8	MR	TT	CC	CC	S
PI 341241A	Israel	IX	11.2	MR	TT	CC	CC	S
PI 561285A	China	I	11.5	MR	TT	CC	CC	S
PI 417260A	Japan	VIII	11.9	MR	TT	CC	CC	S
PI 549029	China	III	12.8	MR	TT	CC	CC	S
PI 603588	China	V	13.3	MR	TT	CC	CC	S
PI 506846	Japan	V	18.2	MR	TT	CC	CC	S
PI 548975	USA	VI	18.4	MR	GG	CG	GG	R
PI 512322D	Georgia	II	18.5	MR	TT	CC	CC	S
PI 574476A	China	IV	19.3	MR	TT	CC	CC	S
PI 522186	Ukraine	0	19.5	MR	TT	CC	CC	S
PI 561310	China	III	19.9	MR	TT	CC	CC	S
PI 548980	USA	VI	20.4	MR	TT	CC	CC	S
PI 171441	China	VI	21.6	MR	GG	GG	CC	R
PI 587660A	China	VII	22.0	MR	TT	CC	CC	S
PI 567175A	China	000	22.0	MR	TT	CC	CC	S
PI 578484	China	V	22.2	MR	TT	CC	CC	S
PI 578432A	China	0	22.7	MR	TT	CC	CC	S
PI 417141	Japan	V	22.7	MR	TT	CC	CC	S

PI 567156A	China	0	22.7	MR	TT	CC	CC	S
PI 561367	China	I	22.7	MR	TT	CC	CC	S
PI 157430	South Korea	V	23.1	MR	TT	CC	CC	S
PI 567206	Georgia	VI	23.1	MR	TT	CC	CC	S
PI 417129A	Japan	IX	23.5	MR	TT	CC	CC	S
PI 567488B	China	IV	23.8	MR	TT	CC	CC	S
PI 274582	Japan	VIII	23.9	MR	TT	CC	CC	S
PI 561242	China	0	24.4	MR	TT	CC	CC	S
PI 561341A	China	I	24.4	MR	TT	CC	CC	S
PI 561229	China	I	25.0	MR	TT	CC	CC	S
PI 561233A	China	I	25.0	MR	TT	CC	CC	S
PI 561330A	China	III	25.0	MR	TT	CC	CC	S
PI 578480	China	V	25.8	MR	TT	CC	CC	S
PI 417320	Japan	VII	25.8	MR	TT	CC	CC	S
PI 561299A	China	0	26.2	MR	TT	CC	CC	S
PI 549019	China	V	26.5	MR	TT	CC	CC	S
PI 578479	China	III	26.7	MR	TT	CC	CC	S
PI 578494A	China	IV	26.9	MR	TG	CC	CC	S
PI 561323	China	III	27.3	MR	GG	CC	CC	R
PI 567163	China	I	27.3	MR	TT	CC	CC	S
PI 561343	China	I	27.3	MR	TT	CC	CC	S
PI 561231	China	I	27.8	MR	TT	CC	CC	S

PI 561344	China	0	28.0	MR	TT	CC	CC	S
PI 181558	Japan	V	28.1	MR	TT	CC	CC	S
PI 587672	China	VII	28.4	MR	TT	CC	CC	S
PI 561313	China	III	28.8	MR	GG	CC	CC	R
PI 567166	China	II	29.0	MR	TT	CC	CC	S
PI 561350A	China	II	29.0	MR	TT	CC	CC	S
PI 549043	China	IV	29.5	MR	TT	CC	CC	S
PI 578377	China	0	29.5	MR	TT	CC	CC	S
PI 561230	China	II	29.5	MR	TT	CC	CC	S
PI 561353	China	I	29.5	MR	TG	CC	CC	S
PI 567161	China	II	29.5	MR	TT	CC	CC	S
PI 561235	China	I	29.5	MR	TT	CC	CC	S
PI 567295	China	VIII	29.6	MR	TT	CC	CC	S
PI 561237	China	I	30.0	MS	TT	CC	NN	S
PI 578358	China	V	30.7	MS	TT	CC	CC	S
PI 561234	China	II	31.5	MS	TT	CC	CC	S
PI 578376	China	II	31.8	MS	GG	CC	CC	R
PI 200459	Japan	VIII	31.8	MS	TT	CC	CC	S
PI 603517A	China	VI	32.7	MS	TT	CC	CC	S
PI 647085	USA	VII	32.7	MS	TT	CC	CC	S
PI 587671	China	VII	33.0	MS	TT	CC	CC	S
PI 561325	China	IV	33.0	MS	TT	CC	CC	S
PI 549030A	China	III	33.3	MS	TT	CC	CC	S

PI 578481	China	I	33.5	MS	TT	CC	CC	S
PI 587668A	China	VI	33.5	MS	TT	CC	CC	S
PI 567316B	China	VI	33.7	MS	TT	CC	CC	S
Jindou19	China	IV	33.7	MS	TT	CC	CC	S
PI 089775	China	VI	33.7	MS	TG	CC	CC	S
PI 341264	Liberia	VI	33.7	MS	TT	CC	CC	S
PI 506552	Japan	V	33.8	MS	TT	CC	CC	S
PI 341248	Tanzania	IX	34.2	MS	TT	CC	CC	S
PI 549078	China	00	34.6	MS	TT	CC	CC	S
PI 603528	China	VII	34.9	MS	TT	CC	CC	S
PI 567410A	China	VII	34.9	MS	TT	CC	CC	S
PI 381683	Uganda	VI	34.9	MS	TT	CC	CC	S
PI 430737	Zimbabwe	VII	35.2	MS	TT	CC	CC	S
PI 219698	Pakistan	VI	35.4	MS	TT	CC	CC	S
PI 587669	China	VI	35.4	MS	TT	CC	CC	S
PI 549079	China	00	35.4	MS	TT	CC	CC	S
PI 567237	China	IV	35.6	MS	TT	CC	CC	S
PI 506791	Japan	V	35.6	MS	TT	CC	CC	S
PI 556950	China	IV	35.6	MS	TT	CC	CC	S
PI 181569	Japan	VII	35.6	MS	TT	CC	CC	S
PI 561335	China	III	35.8	MS	TT	CC	CC	S
PI 341241B	Israel	IX	36.1	MS	TT	CC	CC	S
PI 341242	Tanzania	IX	36.9	MS	TT	CC	CC	S

	South							
PI 417562	Africa	VI	36.9	MS	TT	CC	CC	S
PI 567173	China	00	36.9	MS	TT	CC	CC	S
PI 561296A	China	I	36.9	MS	TT	CC	CC	S
PI 561297	China	II	36.9	MS	TT	CC	CC	S
PI 561370	China	III	36.9	MS	TT	CC	CC	S
PI 549021A	China	III	36.9	MS	TT	CC	CC	S
PI 574478A	China	II	37.1	MS	TT	CC	CC	S
PI 561238	China	I	37.5	MS	TT	CC	CC	S
PI 561332	China	0	37.7	MS	TT	CC	CC	S
FC003659	China	VI	38.5	MS	TT	CC	CC	S
PI 587665	China	VII	38.5	MS	TT	CC	CC	S
PI 549022	China	IV	38.6	MS	TT	CC	CC	S
PI 561304A	China	I	38.6	MS	TT	CC	CC	S
PI 587663	China	VII	38.6	MS	TT	CC	CC	S
PI 578428A	China	0	38.6	MS	TT	CG	CC	S
PI 561302A	China	0	38.6	MS	TT	CC	CC	S
PI 603517B	China	VI	39.7	MS	TT	CC	CC	S
PI 417184A	Japan	VIII	39.8	MS	TT	CC	CC	S
PI 578475	China	VII	40.2	MS	TT	GG	GG	S
PI 578387	China	000	40.2	MS	TT	CC	CC	S
PI 578365	China	II	40.2	MS	TT	CC	CC	S
PI 561368	China	I	40.3	MS	TT	CC	CC	S

PI 476905A	China	V	40.3	MS	TT	CC	CC	S
PI 567158	China	0	40.8	MS	TT	CC	CC	S
PI 603536	China	VIII	40.9	MS	TT	CC	CC	S
Fendou56	China	IV	40.9	MS	TT	CC	CC	S
PI 549041A	China	III	40.9	MS	TT	CC	CC	S
PI 561309A	China	II	40.9	MS	TT	CC	CC	S
PI 567036	Morocco	IX	41.0	MS	TT	CC	CC	S
PI 567406B	China	VI	41.2	MS	TT	CC	CC	S
PI 561354	China	I	41.5	MS	TT	CC	CC	S
PI 578361	China	X	41.5	MS	TT	CC	CC	S
PI 578476	China	IV	41.5	MS	TT	CC	CC	S
PI 322692	Australia	IX	41.7	MS	TT	CC	CC	S
PI 567377B	China	VI	41.7	MS	TT	CC	CC	S
PI 567157A	China	0	42.0	MS	TT	CC	CC	S
PI 574480A	China	II	42.6	MS	TT	CC	CC	S
PI 578495	China	IV	43.1	MS	TT	CC	CC	S
PI 592939	China	IV	43.1	MS	TT	CC	CC	S
PI 279081	South Africa	VII	43.1	MS	TT	CC	CC	S
PI 574482	China	III	43.2	MS	TT	CC	CC	S
PI 171443	China	VI	43.3	MS	TT	CC	CC	S
PI 322689	Angola	VII	43.3	MS	TT	CC	CC	S
PI 416937	Japan	VI	43.3	MS	TT	CC	CC	S

PI 578374	China	I	43.8	MS	TT	CC	CC	S
PI 561318A	China	I	43.8	MS	TT	CC	CC	S
N06-7543	USA	VII	44.2	MS	TT	CC	CC	S
PI 561311A	China	II	44.3	MS	TT	CC	CC	S
PI 423960A	Japan	IX	44.3	MS	TT	CC	CC	S
PI 482601	Zimbabwe	IX	44.5	MS	TT	CC	CC	S
PI 592756	USA	VI	44.5	MS	TT	CC	GG	S
PI 090768	China	VI	44.6	MS	TT	CC	CC	S
PI 549017	China	IV	44.6	MS	TT	CC	CC	S
PI 561240	China	II	44.7	MS	TT	CC	CC	S
PI 603514	China	VI	44.9	MS	TT	CC	CC	S
PI 417389A	Japan	VIII	44.9	MS	TT	CC	CC	S
PI 567349B	China	VI	45.4	MS	TT	CC	CC	S
PI 556948	China	III	45.5	MS	TT	CC	CC	S
PI 603540A	China	VII	45.6	MS	TT	CC	CC	S
PI 553045	USA	VIII	45.6	MS	TT	CC	CC	S
PI 567493	China	IV	45.7	MS	TT	CC	CC	S
PI 434981	Central African Republic	VIII	45.7	MS	TT	CC	CC	S
PI 371610	Pakistan	V	46.0	MS	TT	CC	CC	S
PI 506522	Japan	V	46.2	MS	TT	CC	CC	S
PI 561236	China	II	46.2	MS	TT	CC	CC	S
PI 424131	Zimbabwe	VII	46.5	MS	TT	CC	CC	S

PI 376069	Cameroon	VIII	46.5	MS	TT	CC	CC	S
PI 561336	China	II	46.6	MS	TT	CC	CC	S
PI 578419A	China	II	47.2	MS	TT	CC	CC	S
PI 567405	China	VI	47.2	MS	TT	CC	CC	S
PI 416874A	Japan	IX	47.7	MS	TT	CC	CC	S
PI 578417B	China	I	47.7	MS	TT	CC	CC	S
PI 567325B	China	V	47.8	MS	TT	CC	CC	S
PI 578485A	China	0	48.3	MS	TT	CC	CC	S
PI 561308	China	I	48.3	MS	TT	CC	CC	S
PI 561371	China	IV	48.9	MS	TT	CC	CC	S
PI 506934	Japan	V	48.9	MS	TT	CC	CC	S
PI 578380A	China	I	49.4	MS	TT	CC	CC	S
PI 549020	China	V	49.4	MS	TT	CC	CC	S
PI 416873A	Japan	VIII	49.4	MS	TT	CC	CC	S
PI 417561	South Africa	VI	49.7	MS	TT	CC	CC	S
PI 587657	China	VI	50.0	MS	TT	CC	CC	S
PI 567159A	China	I	50.0	MS	TT	CC	CC	S
PI 578478A	China	IV	50.3	MS	TT	CC	CC	S
PI 603520	China	VI	50.5	MS	TT	CC	CC	S
PI 561324	China	II	51.1	MS	TT	CC	CC	S
PI 578472	China	VI	51.1	MS	TT	CC	CC	S
PI 561348	China	I	51.1	MS	TT	CC	CC	S

PI 578477A	China	IV	51.1	MS	TT	CC	CC	S
PI 430736	Zimbabwe	VI	51.3	MS	TT	CC	CC	S
PI 561346	China	I	51.5	MS	TT	CC	CC	S
PI 506583	Japan	V	52.3	MS	TT	CC	CC	S
PI 417401	Japan	V	52.3	MS	TT	CC	CC	S
PI 423961A	Japan	IX	52.3	MS	TT	CC	NN	S
PI 574477	China	IV	52.3	MS	TT	CC	CC	S
PI 549042A	China	IV	52.3	MS	TT	CC	CC	S
PI 381680	Uganda	VII	52.3	MS	TT	CC	CC	S
PI 639573	Burundi	VIII	52.3	MS	TT	CC	CC	S
PI 567162	China	II	53.0	MS	TT	CC	CC	S
PI 549080	China	00	53.0	MS	TT	CC	CC	S
PI 567171	China	00	53.4	MS	TT	CC	CC	S
PI 561331	China	I	53.8	MS	TT	CC	CC	S
PI 603521	China	VIII	53.8	MS	TT	CC	CC	S
PI 639575	Burundi	VIII	54.5	MS	TT	CC	CC	S
PI 549076A	China	00	54.5	MS	TT	CC	CC	S
PI 561283	China	I	54.5	MS	TT	CC	CC	S
PI 548983	USA	VI	54.6	MS	TT	CC	CC	S
Fendou65	China	IV	54.6	MS	TT	CC	CC	S
Fendou78	China	IV	55.3	MS	TT	CC	CC	S
PI 549028	China	V	55.3	MS	TT	CC	CC	S
PI 561327A	China	I	55.7	MS	TT	CC	CC	S

PI 549045A	China	IV	55.7	MS	TT	CC	CC	S
PI 561337	China	I	56.3	MS	TT	CC	CC	S
PI 561307	China	I	56.3	MS	TT	CC	CC	S
PI 561351	China	I	56.3	MS	TT	CC	CC	S
PI 561320	China	II	56.3	MS	TT	CC	CC	S
PI 561339	China	III	56.4	MS	TT	CC	CC	S
PI 567174A	China	00	56.8	MS	TT	CC	CC	S
PI 578482A	China	0	56.8	MS	TT	CC	NN	S
PI 639576	Burundi	VIII	57.3	MS	TT	CC	CC	S
PI 561347	China	II	57.4	MS	TT	CC	CC	S
PI 561333	China	I	57.4	MS	TT	CC	CC	S
PI 561317	China	I	57.4	MS	TT	CC	CC	S
PI 578363	China	II	57.6	MS	TT	CC	CC	S
PI 221715	South Africa	VII	57.7	MS	TT	CC	CC	S
PI 587666	China	VI	57.7	MS	TT	CC	CC	S
PI 322695	Angola	VI	57.7	MS	TT	CC	CC	S
PI 549073	China	0	58.0	MS	TG	CC	CC	S
PI 567726	China	IV	58.2	MS	TT	CC	CC	S
PI 567683B	China	VI	58.5	MS	TT	CC	CC	S
PI 404192C	China	I	58.8	MS	TT	CC	CC	S
PI 578392A	China	I	58.8	MS	TT	CC	CC	S
PI 567164	China	0	59.1	MS	TT	CC	CC	S

PI 201422	China	VI	59.1	MS	TT	CC	NN	S
PI 587676	China	VI	59.7	MS	TT	CC	CC	S
PI 567170A	China	II	59.7	MS	TT	CC	CC	S
PI 587662A	China	VII	59.9	MS	TG	CC	CC	S
PI 603506	China	VI	60.0	S	TT	CC	CC	S
PI 221716	South Africa	VII	60.0	S	TT	CC	CC	S
PI 159093	South Africa	VII	60.1	S	TT	CC	CC	S
PI 247678	Zaire	VIII	60.8	S	TT	CC	CC	S
PI 587661A	China	VI	61.2	S	TT	CC	CC	S
PI 603535	China	VIII	61.3	S	TT	CC	CC	S
PI 230972	Japan	VIII	61.4	S	TT	CC	CC	S
PI 587670A	China	VI	61.4	S	TT	CC	CC	S
PI 603539A	China	VI	62.5	S	TT	CC	CC	S
PI 423971A	Japan	IX	62.5	S	TT	CC	CC	S
PI 587675	China	VII	62.6	S	TT	CC	CC	S
NC06-1090	USA	VI	62.7	S	TT	CC	CC	S
PI 578385	China	I	62.9	S	TT	CC	CC	S
PI 567410C	China	VII	63.7	S	TT	CC	CC	S
PI 561227	China	II	64.4	S	TT	CC	CC	S
PI 603538C	China	VIII	64.9	S	TT	CC	CC	S
PI 210349	Mozambique	VIII	64.9	S	TT	CC	CC	S

PI 567412	China	VI	64.9	S	TT	CC	CC	S
PI 341253	Sudan	X	64.9	S	TT	CC	CC	S
PI 462312	India	VIII	64.9	S	TT	CC	CC	S
PI 330635	South Africa	VIII	65.7	S	TT	CC	CC	S
PI 578388A	China	0	65.9	S	TT	CC	CC	S
PI 561319A	China	II	66.5	S	TT	CC	CC	S
PI 587659A	China	VI	67.0	S	TT	CC	CC	S
PI 603532	China	VI	67.3	S	TT	CC	CC	S
PI 548657	USA	VII	67.3	S	TT	CC	CC	S
PI 347544A	Hungary	00	67.4	S	TT	CC	CC	S
PI 506730	Japan	V	67.4	S	TT	CC	CC	S
PI 578366	China	III	67.6	S	TT	CC	CC	S
PI 578474	China	I	67.7	S	TT	CC	CC	S
PI 549031	China	III	67.7	S	TT	CC	CC	S
PI 159095	South Africa	VII	68.7	S	TT	CC	CC	S
PI 578401A	China	I	68.8	S	TT	CC	CC	S
PI 561298	China	II	68.9	S	TT	CC	CC	S
PI 123577A	China	V	68.9	S	TT	CC	CC	S
PI 407738	China	VI	69.7	S	TT	CC	CC	S
PI 561342	China	0	69.9	S	TT	CC	CC	S
PI 561303	China	IV	69.9	S	TT	CC	CC	S

PI 567773	China	IV	70.5	S	TT	CC	CC	S
PI 561301	China	I	71.2	S	TT	CC	CC	S
NTCPR94-5157	USA	VI	72.3	S	TT	CC	CC	S
PI 561316	China	II	73.3	S	TT	CC	CC	S
PI 398823	South Korea	IV	73.7	S	GG	GG	CC	R
PI 587667	China	V	73.9	S	TT	CC	CC	S
PI 561305	China	I	73.9	S	TT	CC	CC	S
PI 578375A	China	0	74.2	S	TT	CC	CC	S
PI 587655	China	VII	74.2	S	TT	CC	CC	S
PI 381662	Uganda	VI	74.5	S	TT	CC	CC	S
PI 374220	South Africa	VI	74.6	S	TT	CC	CC	S
PI 567160	China	II	75.0	S	TT	CC	CC	S
PI 549074	China	00	75.0	S	TT	CC	CC	S
PI 561314A	China	II	75.6	S	TT	CC	GG	S
PI 567334	China	VI	75.7	S	TT	CC	CC	S
PI 561315	China	I	75.8	S	TT	CC	CC	S
PI 578362	China	I	76.5	S	TT	CC	CC	S
PI 306702A	Tanzania	IX	76.5	S	TT	CC	CC	S
PI 494851	Zambia	VI	76.9	S	TT	CC	CC	S
PI 323570	India	VII	76.9	S	TT	CC	CC	S
PI 603534B	China	VIII	78.5	S	TT	CC	CC	S

PI 567350B	China	VI	80.1	S	TT	CC	CC	S
PI 578409A	China	I	80.7	S	TT	CC	CC	S
PI 567167	China	00	80.7	S	TT	CC	CC	S
PI 434980B	Central African Republic	IX	81.7	S	TT	CC	CC	S
PI 416880A	Japan	IX	81.8	S	TT	CC	CC	S
PI 567205	Georgia	VI	82.7	S	TT	CC	CC	S
PI 561306	China	I	84.1	S	TT	CC	CC	S
PI 417365A	Japan	VIII	84.1	S	TT	CC	CC	S
PI 587656	China	VI	84.7	S	TT	CC	CC	S
PI 587674A	China	IV	84.8	S	TT	CC	CC	S
PI 553046	USA	VII	84.9	S	TT	CC	CC	S
PI 341245	Tanzania	IX	85.3	S	TT	CC	CC	S
PI 159094	South Africa	VII	85.6	S	TT	CC	CC	S
PI 561295	China	I	85.8	S	TT	CC	CC	S
PI 374180	Burundi	VIII	86.6	S	TT	CC	CC	S
PI 561241	China	I	87.9	S	TT	CC	CC	S
PI 578359	China	V	87.9	S	TT	CC	CC	S
PI 639572	Ghana	VIII	88.1	S	TT	CC	CC	S
PI 495016	Srilanka	X	88.1	S	TT	CC	CC	S
PI 567737	China	IV	88.2	S	TT	CC	CC	S
PI 437126B	Georgia	VI	88.9	S	TT	CC	CC	S

PI 561349	China	II	90.9	S	TT	CC	NN	S
PI 507017	Japan	VII	92.0	S	TT	CC	CC	S
N05-7432	USA	VIII	92.3	S	TT	CC	CC	S
PI 603519	China	VI	92.9	S	TT	CC	CC	S
PI 603513B	China	VIII	93.5	S	TT	CC	CC	S
PI 561340	China	II	93.5	S	TT	CC	CC	S
PI 221714	South Africa	VI	93.8	S	TT	CC	CC	S
PI 578473A	China	III	94.3	S	TT	CC	CC	S
PI 416825A	Japan	IX	94.9	S	TT	CC	CC	S
PI 603512	China	VI	95.0	S	TT	CC	CC	S
PI 417014A	Japan	IX	95.4	S	TT	CC	CC	S
PI 578417A	China	I	95.4	S	TT	CC	CC	S
PI 578364	China	II	96.2	S	TT	CC	CC	S
PI 578360	China	II	96.2	S	TT	CC	CC	S
PI 603534A	China	VII	99.4	S	TT	CC	CC	S
PI 549040	China	IV	100.0	S	TT	CC	CC	S
PI 567315	China	VII	100.6	S	TT	CC	CC	S
PI 374221	South Africa	VI	101.0	S	TT	CC	CC	S
PI 548660	USA	VII	101.5	S	TT	CC	CC	S
PI 416894A	Japan	IX	103.4	S	TT	CC	CC	S
PI 203406	South	VIII	103.8	S	TT	CC	CC	S

	Africa							
PI 603513A	China	VIII	105.4	S	TT	CC	CC	S
PI 561322	China	II	105.8	S	TT	CC	CC	S
PI 360846	Japan	IV	107.4	S	TT	CC	CC	S
PI 417234	Japan	VIII	108.1	S	TT	CC	CC	S
PI 429330	Nigeria	VIII	109.0	S	TT	CC	CC	S
PI 574485	China	IV	109.0	S	TT	CC	CC	S
PI 561326	China	II	110.2	S	TT	CC	CC	S
PI 561338A	China	II	111.4	S	TT	CC	CC	S
PI 090495	China	VI	111.8	S	TT	CC	CC	S
PI 265498	Zaire	VIII	111.8	S	TT	CC	CC	S
N06-7194	USA	VIII	111.8	S	TT	CC	CC	S
PI 615694	USA	VII	112.2	S	TT	CC	CC	S
PI 086736	Japan	VII	112.5	S	TT	CC	CC	S
PI 549023B	China	V	112.5	S	TT	CC	CC	S
PI 322694	Zimbabwe	VI	113.8	S	TT	CC	CC	S
PI 157444	South Korea	V	116.4	S	TT	CC	CC	S
PI 603509	China	VIII	119.5	S	TT	CC	CC	S
PI 567394A	China	VI	120.2	S	TT	CC	CC	S
PI 103419A	China	V	126.1	S	TT	CC	CC	S
PI 322691	Mozambique	IX	126.6	S	TT	CC	CC	S
PI 505649B	Zambia	IX	129.8	S	TT	CC	CC	S

PI 269518B	Pakistan	VI	132.2	S	TT	CC	CC	S
PI 603566	China	III	132.8	S	TT	CC	CC	S
PI 486329	India	VIII	134.6	S	TT	CC	CC	S
PI 486328	India	VIII	135.9	S	TT	CC	CC	S
PI 603537D	China	VII	137.4	S	TT	CC	CC	S
PI 561284	China	I	140.5	S	TT	CC	CC	S
PI 639574	Burundi	VIII	142.6	S	TT	CC	CC	S
PI 567393	China	VII	145.1	S	TT	CC	CC	S
PI 574486	China	III	145.8	S	TT	CC	CC	S
PI 090499	China	VI	148.2	S	TT	CC	CC	S
PI 532458	China	VIII	153.8	S	TT	CC	CC	S
PI 578498A	China	III	157.6	S	TT	CC	CC	S
PI 428691	India	VIII	175.5	S	TT	CC	CC	S
PI 587673	China	VI	176.2	S	TT	CC	CC	S
PI 341244B	Tanzania	IX	183.8	S	TT	CC	CC	S
PI 567771C	China	IV	191.2	S	TT	CC	CC	S
PI 424387	South Korea	IV	199.8	S	TT	CC	CC	S
PI 497967	India	VII	200.0	S	TT	CC	CC	S
Susceptible checks								
PI 518664 'Hutcheson'	USA	V	74.0	S	TT	CC	CC	S
PI 548445 'CNS'	USA	VII	92.5	S	TT	CC	CC	S

PI 548656 'Lee'	USA	VI	94.0	S	TT	CC	CC	S
PI 548658 'Lee 74'	USA	VI	100.0	S	TT	CC	CC	S
PI 617045 'NC Roy'	USA	VI	73.0	S	TT	CC	CC	S
PI 641156 'NC Raleigh'	USA	VII	61.9	S	TT	CC	CC	S

[a] The level of resistance was defined by the female index (FI) that was calculated based on the ratio between the mean numbers of cysts on a given line and 'Lee 74', reported here as a percentage (Niblack et al., 2002).
[b] Rating scale of SCN was based on Schmitt and Shannon (1992): FI < 10% (resistant, R); 10% < FI < 30% (moderately resistant, MR); 30% < FI < 60% (moderately susceptible, MS), and FI > 60% (susceptible, S).
[c] All soybean accessions were genotyped using KASP markers at the *Rhg1* and *Rhg4* loci that were reported by Shi et al. (2015).

Table 2.5: The HG Type determination included with each set.
HG Type 0 and HG Type 5 are SCN race 3 based on the previous SCN race determination.

No	Indicator lines	Set 1	Set 2	Set 3	Confirmation test
			Female index [a]		
1	Peking	0*	0	0	6.5
2	PI 88788	7.3	0	6.2	6.9
3	PI 90763	0	0	0	0
4	PI 437654	0	0	0	0
5	PI 209332	0.4	0	9.9	49.0
6	PI 89772	0	0	0	0
7	PI 548316	0	0	6.8	7.7
	HG Type [b]	HG Type 0	HG Type 0	HG Type 0	HG Type 5

[a] The number is female index in percentage. Female index is the ratio between number of cysts on the given line and number of cysts on 'Lee 74', the standard susceptible check.
[b] HG Types were designated by number based on which indicator line FI > 10 %.

Table 2.6. Analysis of variance of greenhosue phenotping ratings among 462 accesssions.

Source	DF number	F-ratio	Prob > F
Accessions (Fixed effect)	461	1.651	<.0001*
Test (Random effect)	2		<.0001*
Block (Random effect)	3		0.34

Table 2.7. SNPs that were identified for SCN resistance using multiple GWAS models.

SNP	Ch[a]	Position	P value	Rsq	R allele	S allele	Avg FI for the R allele (%)	Avg FI for the S allele (%)	Model
ss715583938	2	8,011,984	8.94E-08	-[b]	T	G	49.2	51.0	FarmCPU
ss715597408	7	36,368,238	4.42E-08	0.26	G	A	18.5	60.2	cMLM
ss715597409	7	36,371,468	2.59E-08	0.26	G	A	18.3	60.5	cMLM
ss715597410	7	36,376,909	2.23E-08	0.26	A	C	17.5	60.4	cMLM
ss715597413	7	36,399,537	5.65E-09	0.26	G	T	18.0	60.5	cMLM
ss715597431	7	36,449,014	1.38E-09	0.27	T	C	17.2	60.8	cMLM/ECMLM
ss715597474	7	36,745,679	9.54E-08	0.25	A	G	28.3	61.5	cMLM
ss715597497	7	36,907,461	4.59E-09	0.27	G	T	27.6	62.3	cMLM/ECMLM
ss715597494	7	36,894,266	4.62E-09	0.27	T	C	29.4	62.3	cMLM/ECMLM
GSM191[c]	8	8,361,148	1.26E-08	0.26	G	C	4.6	56.2	cMLM
ss715606985	10	40,672,699	6.15E-08	0.26	A	G	3.8	56.1	cMLM
GSM381[c]	18	1,645,407	1.24E-17	0.34	G	T	4.2	60.7	cMLM/ECMLM/FarmCPU
GSM383[c]	18	1,643,660	6.81E-08	0.26	G	C	4.1	55.8	cMLM

[a] Abbreviation: Ch: Chromosome; Rsq: R-square; R.allele: resistance allele; S.allele: susceptible allele; AvgFI: average of female index.
[b] FarmCPU did not provide R-square
[c] GSM381, GSM383 and GSM191 were KASP SNP markers at *Rhg1* and *Rhg4* loci

Table 2.8. Predicted genes for SCN resistance located in three genomic regions on chromosomes 2, 7, and 10.

Gene model	Chr	Position (bp)	Function
Glyma.02g089800	2	7,896,985-7,897,303	RNA binding
Glyma.02g089900	2	7,903,639-7,908,241	transmembrane transport
Glyma.02g090100	2	7,925,640-7,933,611	transmembrane transport
Glyma.02g090300	2	7,939,823-7,944,358	DNA ligase (ATP) activity
Glyma.02g090400	2	7,957,508-7,978,145	DNA ligase (ATP) activity
Glyma.02g090500	2	7,980,214-7,986,739	metabolic process
Glyma.02g090600	2	7,997,792-8,001,669	protein folding
Glyma.02g090800	2	8,023,897-8,026,772	protein binding
Glyma.02g090900	2	8,027,717-8,029,908	zinc ion binding
Glyma.07g194200	7	36,276,988-36,282,713	DNA binding
Glyma.07g194400	7	36,294,553-36,297,818	Cytochrome P450 CYP2 subfamily
Glyma.07g194800	7	36,326,338-36,337,611	Protein binding
Glyma.07g195000	7	36,346,694-36,349,527	Intracellular protein transport
Glyma.07g195100	7	36,354,678-36,359,664	LRR-Protein kinase activity
Glyma.07g195300	7	36,383,490-36,386,028	Metabolic process
Glyma.07g195400	7	36,395,763-36,399,680	RING/ U-box superfamily protein
Glyma.07g195500	7	36,411,199-36,419,729	Nucleotide-excision repair
Glyma.07g195700	7	36,428,858-36,434,430	ATP binding\mismatch repair
Glyma.07g195900	7	36,437,193-36,449,264	ATP binding
Glyma.07g196000	7	36,452,344-36,455,168	Protein binding
Glyma.07g196500	7	36,483,372-36,490,967	Ligase activity
Glyma.07g199000	7	36,759,783-36,761,919	DNA binding
Glyma.07g200100	7	36,873,150-36,873,836	Protein binding
Glyma.07g199500	7	36,809,934-36,815,777	LRR-like protein kinase
Glyma.07g199700	7	36,835,550-36,836,275	Regulation of transcription
Glyma.07g199900	7	36,854,378-36,857,292	Protein binding
Glyma.10g172700	10	40,646,539-40,649,573	Metabolic process
	10	40,658,709-40,668,066	Protein import into mitochondrial outer membrane
Glyma.10g172800			
Glyma.10g172900	10	40,666,359-40,670,242	Transmembrane transport
Glyma.10g173000	10	40,686,435-40,696,440	Protein kinase activity
Glyma.10g173100	10	40,699,391-40,701,312	Electron carrier activity
Glyma.10g173300	10	40,706,365-40,713,418	Protein binding
Glyma.10g173400	10	40,720,918-40,723,699	Plasma membrane

Source: soybase.org

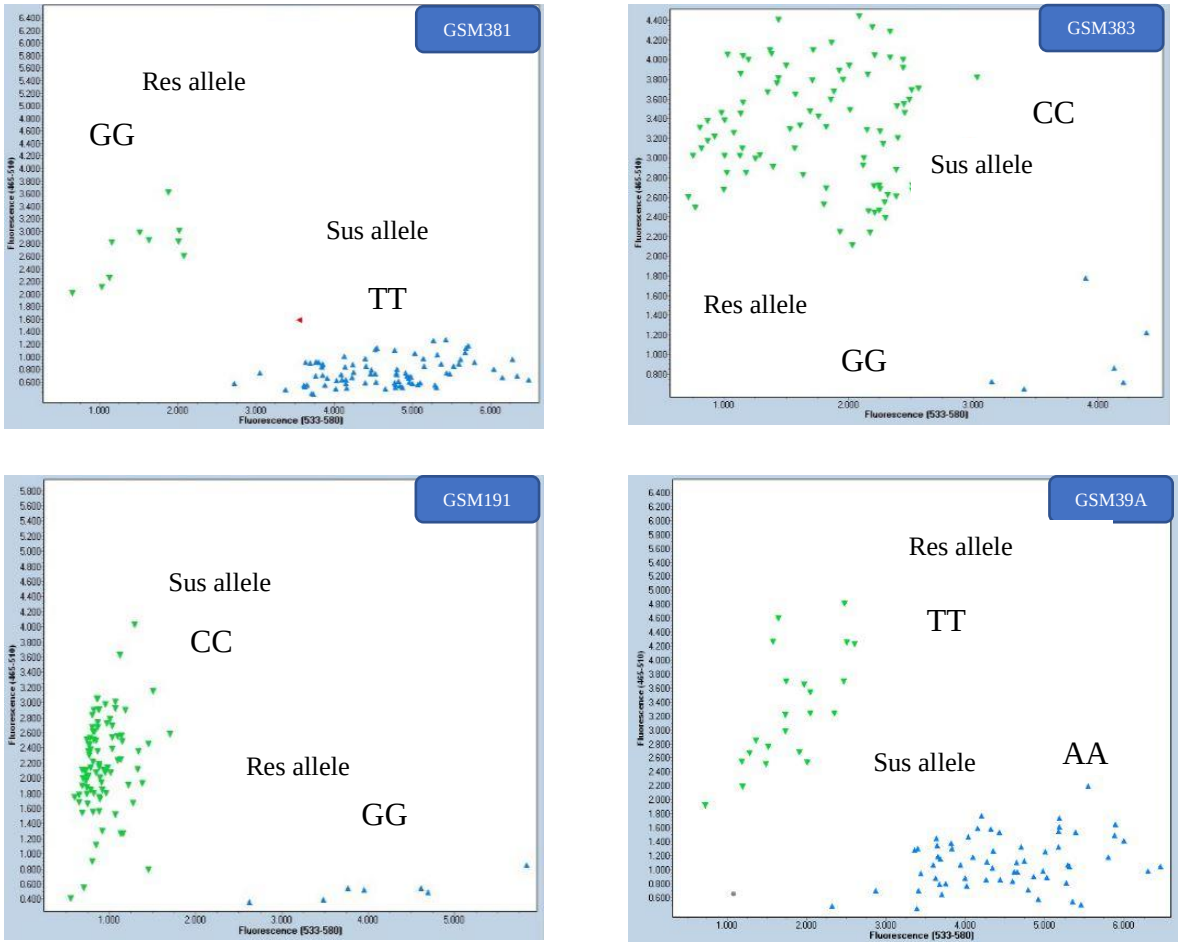

Figure 2.1 Examples of SNP graphs of 462 soybean accessions that were genotyped with three SNP markers.

They included GSM381; GSM383 and GSM191 at *Rhg1* and *Rhg4* loci reported by Shi et al. (2015) and one SNP marker GSM39A for southern RKN reported by Pham et al. (2013).

Res = Resistance; Sus =Succeptible.

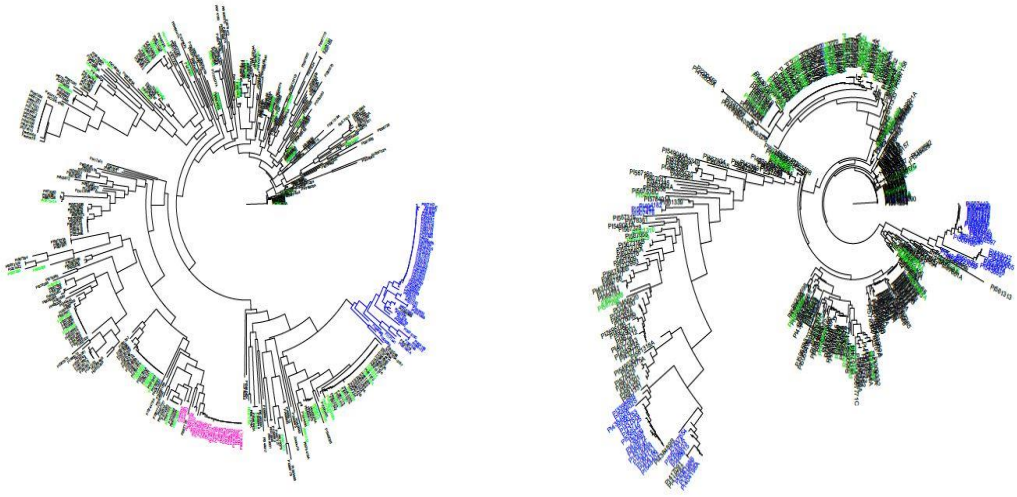

Figure 2.2. Dendogram of 461 soybean accessions generated using the haplotype SNPs at *Rhg1* and *Rhg4* loci.

Left: *Rhg1* locus: 990kb region on chr. 18 consisting of 103 SNPs. **Right:** *Rhg4* locus: A 997kb on chr.8 consisting of 64 SNPs. 58 unique lines (green) were grouped separately from PI 88788 and 'Peking' (pink & blue) at both *Rhg1* and *Rhg4* regions.

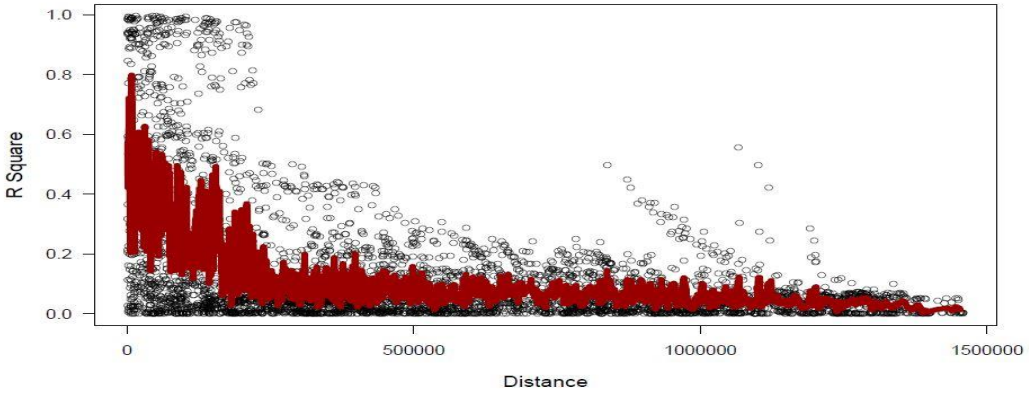

Figure 2.3 Genome-wide linkage disequilibrium decay of the 461 soybean accessions.

The LD was estimated as the chromosome distance (base pair values on the x-axis) where the pairwise correlation coefficient r^2 values (on the y-axis) dropped to half of its maximum value (0.4).

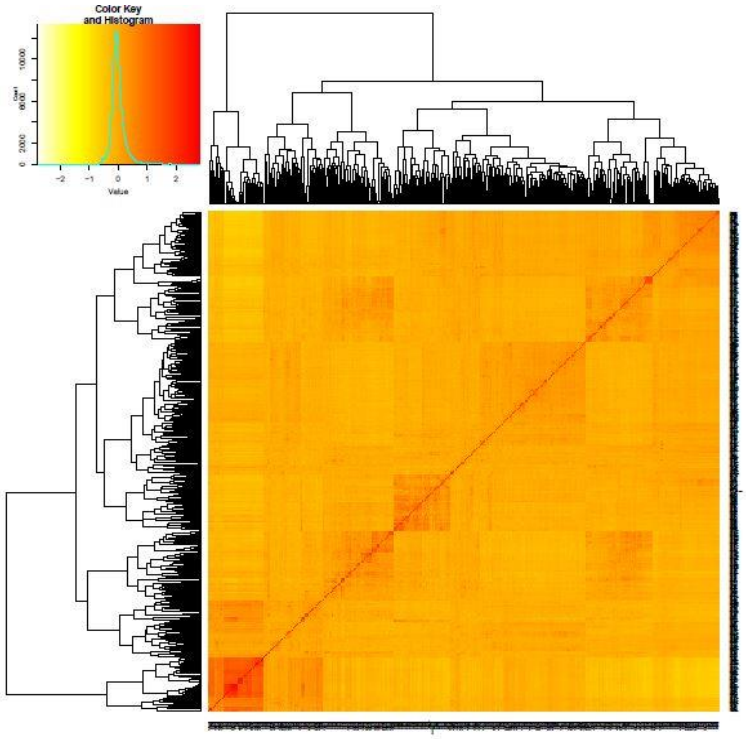

Figure 2.4 Kinship matrix among 461 soybean accessions using 35,817 SNPs.

Kinship matrix were calculated using GAPIT package. The color (yellow to red) indicated the level of kinship.

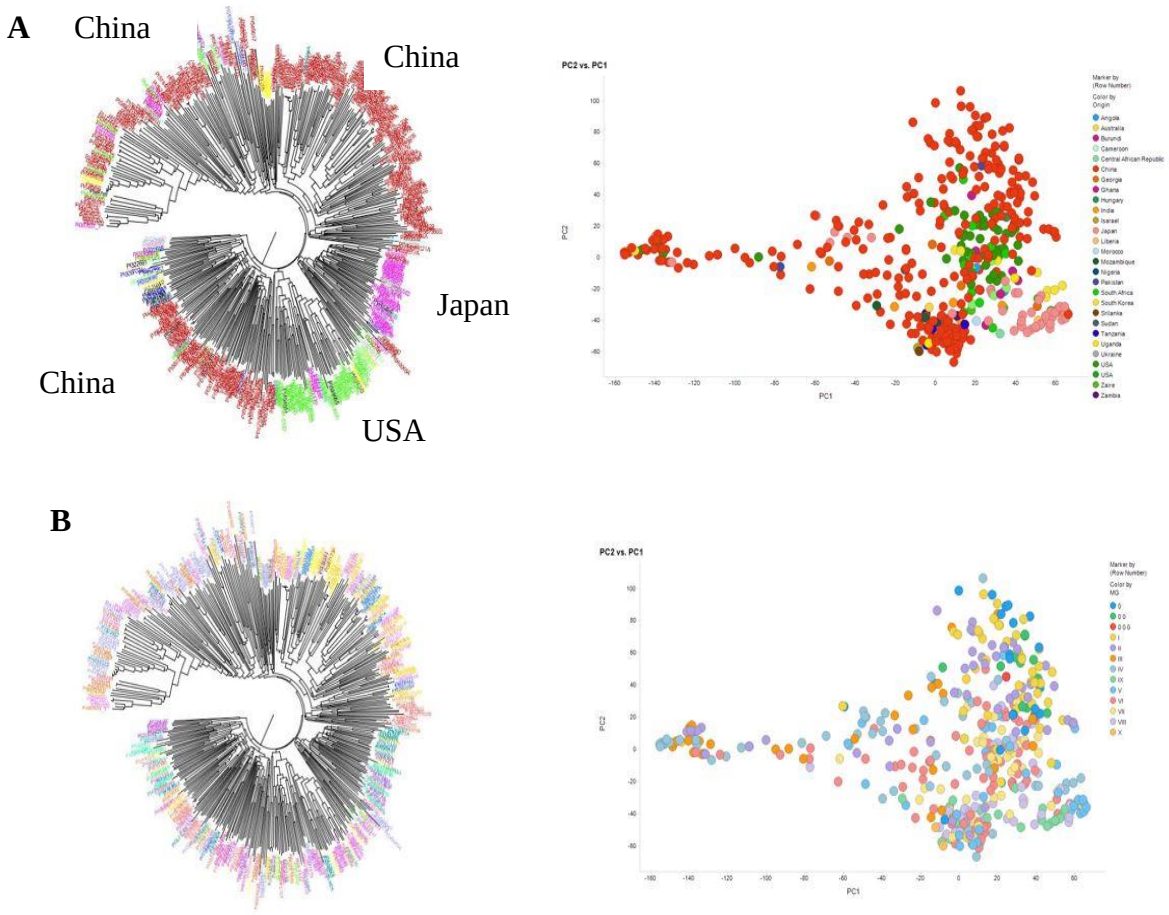

Figure 2.5 Neighbor-joining tree and principal component analysis depicted the clusters that were formed among 461 accessions.

A: colored by origin: China: red, Japan: Pink, USA: green, other colors: other countries); B. colored by maturity. The result of neighbor joining tree was similar with Principle Component Analysis using first two components.

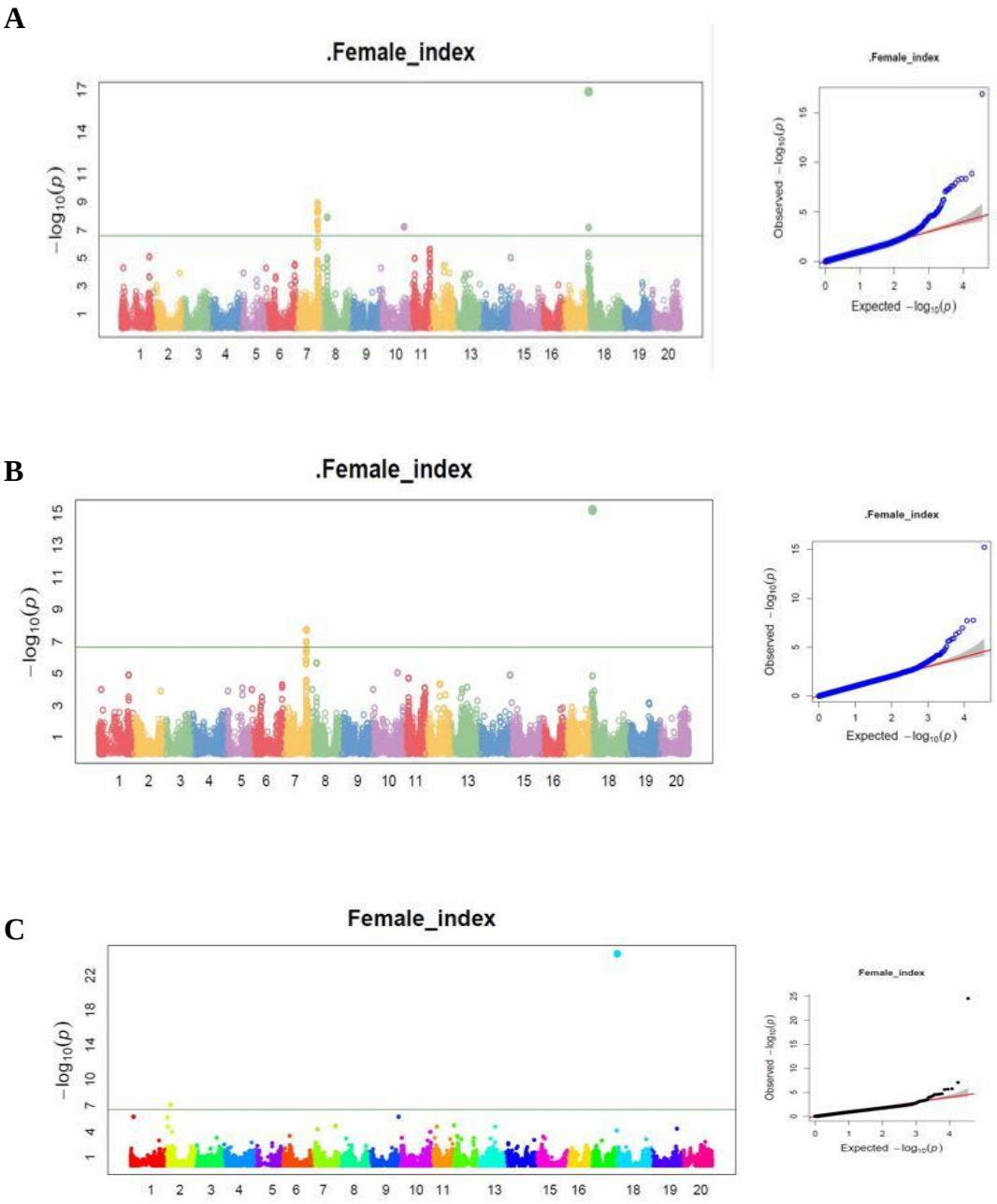

Figure 2.6 Manhattan plots generated from genome-wide analysis using GAPIT and FarmCPU packages.

A) compressed mix linear model (cMLM); B) enriched compressed mix linear model (ECMLM) and C) fixed and random model circulating probability unification (FarmCPU). The -log10 P values from a genome-wide scan are plotted against the position on each of the 20 chromosomes. The horizontal green lines indicated the genome-wide significant. **Right:** Q-Q plot showing the expected P value compared to the observed value. The red line is X= Y line indicated the null hypothesis: no true association. Q-Q plot curved at the tail which implied the small number of true SNPs association.

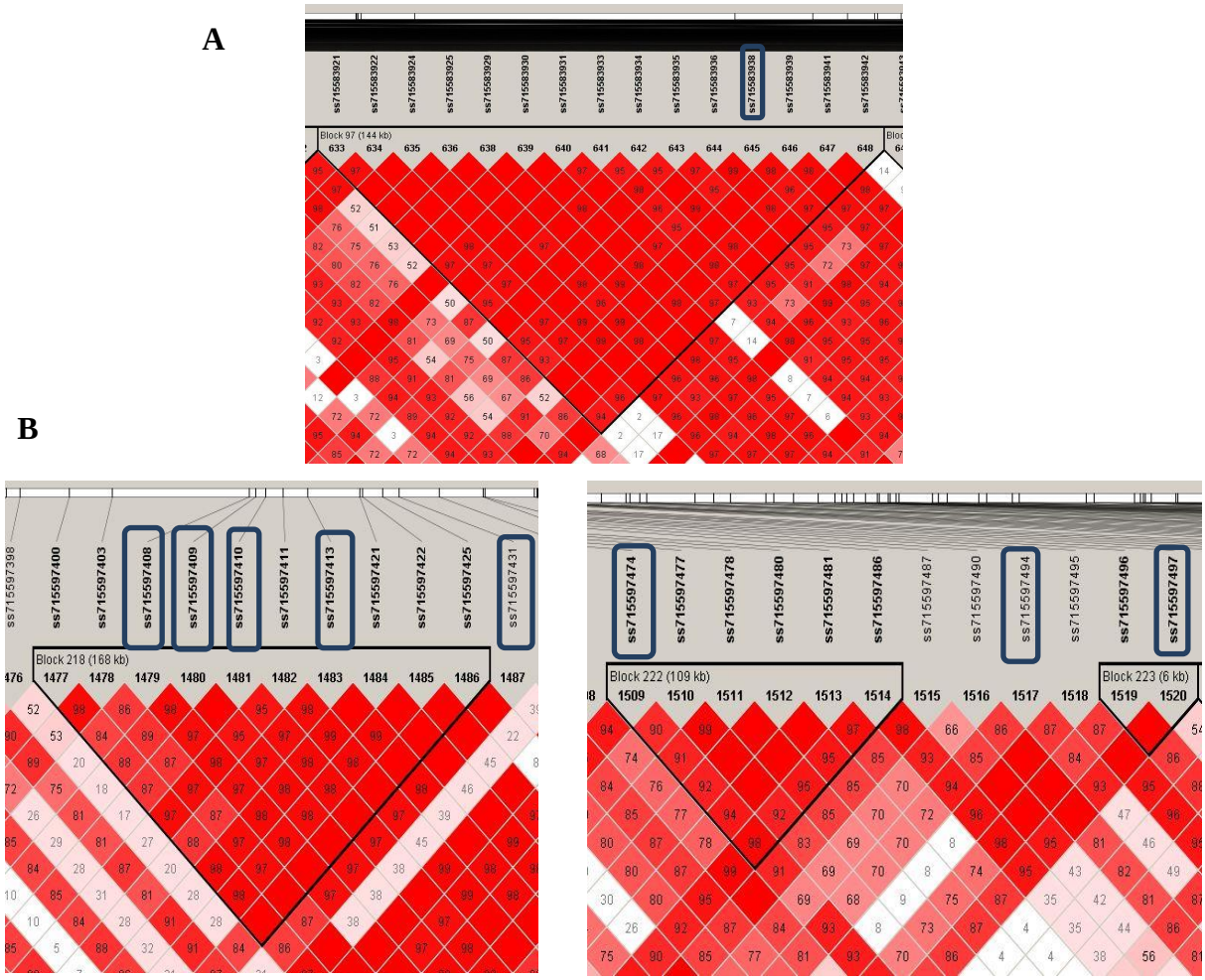

Figure 2.7 Linkage disequilibrium (LD) plot in two genomic regions on Chr 2 (A); and Chr 7 (B).

Nine SNPs associated with HG Type 0 resistance were indicated in blue boxes. The r^2 value in the LD triangles were expressed with color intensity by Haploview software (Barrett et al., 2005).

CHAPTER 3

IDENTIFICATION OF QUANTITATIVE TRAIT LOCI UNDERLYING RESISTANCE TO SOYBEAN CYST NEMATODE (*HETERODERA GLYCINES*) IN SOYBEAN PI 567488B[2]

[2] Tran, D. T., Noe, J., Arelli, P. R., Li. Z. Identification of quantitative trait loci underlying resistance to soybean cyst nematode (*Heterodera glycines*) in soybean PI 567488B. To be submitted to Molecular Breeding.

Abstract

Soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is one of the most damaging pests of soybean production in the U.S. To date, breeders have mainly relied on SCN resistance alleles at the *Rhg1* and *Rhg4* loci present in 'Peking' and PI 88788, respectively to develop SCN resistant cultivars, However, overuse of these two sources has led to a shift that some of SCN populations have overcome *rhg1* and *Rhg4* derived resistance. Therefore, it is critical to identify new germplasm sources with SCN resistance beyond the defense mechanisms found at the *Rhg1* and *Rhg4* loci. Our SCN greenhouse screening studies revealed that PI 567488B was resistant to HG Type 5 (SCN race 3) but did not carry either *rhg1* or *Rhg4* resistance alleles. Thus, the objective of this study was to identify quantitative trait loci underlying HG Type 5 (SCN race 3) resistance in PI 567488B. One hundred fortyone-$F_{2:3}$ families derived from the cross of 'Lee' × PI 567488B were phenotyped with HG Type 5 in greenhouse and genotyped using the SoySNP6K Infinium Chips, which identified 1,618 polymorphic SNPs that were used to construct a linkage map. Subsequent composite interval mapping identified two significant QTLs on chromosomes (Chr) 16 and 20, respectively, conferring HG Type 5 (SCN race 3) resistance. The QTL on Chr 20 overlapped with a previously reported QTL found in a wild soybean accession (PI 468925B), whereas the QTL on Chr 16 was closely mapped to the location of another previously reported QTL identified in PI 438489B. The results of this research will allow for the introgression of PI 567488B-type SCN resistance into elite soybean germplasm to develop SCN resistant cultivars.

Keywords: Resistance to *Heterodera glycines* (*Rhg*) gene; Single nucleotide polymorphism (SNP); Soybean cyst nematode (SCN); Linkage mapping; Linkage group (LG), Quantitative trait loci (QTL), SoySNP6K Infinium Chip data.

Introduction

Soybean [*Glycine max* (L.) Merr.) is an important source of protein and oil for both animal feed and human consumption. The U.S. ranks as the top producer, accounting for more than 30% of the world production (fao.org/faostat). However, soybean production in the U.S. is affected considerably by soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe), the most devastating pest thwarting production with an estimated loss of nearly 30% of yield annually (extension.cropscience.illinois.edu). SCN was originally detected in Japan (Ichinohe, 1959) and first reported in the U.S. in 1954 in North Carolina (Riggs, 1975). To date, it has spread to most soybean producing states except West Virginia (Tylka and Marett, 2014; Wang et al., 2017).

Quantitative trait loci (QTL) mapping for SCN resistance has been conducted for nearly 20 years (Concibido et al., 2004; Kim et al., 2016), which identified many genomic regions conferring SCN resistance on almost every soybean chromosome. However, only two major QTLs conferring SCN resistance have been reported among many germplasm sources Concibido et al., 2004; Kim et al., 2016). The most important major resistance QTL resides at the *Rhg1* locus on chromosome (Chr) 18 [linkage group (LG) G], which is associated with resistance to multiple SCN races including SCN race 3 (Concibido et al., 2004). This QTL was detected in a majority of resistant germplasm sources including 'Peking', PI 88788, PI 437654, PI 89772, PI 90763, PI 209332 and PI 548316, accounting for up to 50% of the phenotypic variance observed for SCN race 3 resistance (Mudge et al., 1997; Concibido et al., 2004). The second major resistance QTL for HG Type 0 (race 3) and HG Type 2.5.7 (race 1) is the *Rhg4* locus located on Chr 8 (LG A2), which was detected in 'Peking', PI 90763, PI 437654 and PI 89772 and explained ~28% of phenotypic variance (Weisemam et al., 1992; Concibido et al., 2004). However, 'Peking', PI 437654, PI 89772 and PI 90763 require both *rhg4 and Rhg4* alleles to

provide resistance to SCN. Recently, resistance alleles at these two loci were cloned (Cook et al., 2012; Liu et al., 2012). At *Rhg1* locus, three genes together significantly contribute to SCN resistance: *Glyma18g02580* that encodes a predicted amino-acid transporter; *Glyma18g02590* that encoded a SNAP protein, and *Glyma18g02610* that encoded wound-inducible protein 12. Three types of copy number variation of these three genes were found in the susceptible cultivar 'William 82' (1 copy), 'Peking' group (2 – 4 copies), and PI 88788 group (more than 6 copies). At the *Rhg4* locus, a *serine hydroxymethyltransferase* (*SHMT*) gene was found in 'Peking' group, which requires *rhg1* allele on Chr 18 for SCN resistance. Based on these gene cloning information reported by Cook et al. (2012) and Liu et al. (2012), Shi et al. (2015) developed three SNP markers (GSM381, GSM383 and GSM191) for marker-assisted selections of the *rhg1* and *Rhg4* alleles, respectively using the Kompetitive Allele Specific PCR (KASP) assay. Two of these KASP markers (GSM381 and GSM383) were designed for selection of the *rhg1* resistance alleles while GSM383 can be used to distinguish 'Peking' and PI 88788-type resistance at the *Rhg1* locus. The other KASP marker, GSM191, was designed for selection of the resistance allele at the *Rhg4* locus.

To date, both types of resistance have been extensively utilized to develop SCN resistant cultivars by soybean breeders (Tylka et al., 2016). It was reported that more than 95% of SCN resistant cultivars planted in Illinois are derived from PI 88788 sources (Melito et al., 2010). Due to the fact that soybean breeders have primarily used 'Peking' and PI 88788 sources to develop soybean cultivars with SCN resistance, some of SCN populations have shifted, rendering the 'Peking' and PI 88788 SCN resistance no longer effective. In fact, some SCN populations detected in Missouri and Illinois have been categorized as virulent to some indicator lines including PI 88788 and 'Peking' (Niblack et al., 2008; Zheng and Chen, 2011). In Georgia, a

new aggressive SCN population designated as HG Type 5 has been found based on greenhouse screening studies using the HG Type determination system (Table 3.1). Therefore, it is critical to find new sources of SCN resistance beyond *Rhg1* and *Rhg4* loci to support SCN resistant breeding efforts. In addition to *Rhg1* and *Rhg4* loci, other genomic regions for SCN resistance have been reported: (1) *cqSCN-003* located on Chr 16 (LG J) conferring HG Type 7 (race 3) and HG Type 1.3.5.6.7 (race 14) (Glover, 2004); (2) *cqSCN-005* on Chr 17 (LG D2) detected in PI 437654 (Kazi et al., 2010); (3) *cqSCN-006* and *cqSCN-007* on Chrs 15 (LG E) and 18 (LG G), respectively, conferring for SCN race 3 detected in wild soybean PI 468916 (Wang et al., 2001); and (4) *qSCN10* detected in PI 567516C for HG Type 0 resistance on Chr 10 (LG O) (Vuong et al., 2010).

Herein, in an effort to identify novel QTL for SCN resistance, 462 soybean accessions with various origins were screened in 2016 for resistance to HG Type 0 (SCN race 3) using an inoculation assay conducted at the University of Georgia greenhouse. Later, 106 potential resistant accessions were screened with new aggressive HG Type 5 (SCN race 3). Of these accessions screened, PI 567488B from China was identified as being resistant to both HG Type 0 and HG Type 5 (SCN race 3) (Tran et al., unpublished data). A subsequent genotyping and haplotype analysis of the *rhg1* and *Rhg4* resistance alleles using KASP markers and SoySNP50K Infinium Chip data at these two genomic regions on Chrs 8 and 18, respectively, indicated that PI 567488B did not carry the *rhg1* or *Rhg4* resistance alleles, suggesting that this PI may carry novel genes or alleles conferring SCN resistance. In previous SCN resistance screening studies, it was reported that PI 567488B was characterized as being resistant to HG Type 1.2.5.7 (SCN race 2) (Arelli et al., 2015). Additionally, PI 567488B also carries desirable agronomic traits, such as yellow seed color, and low lodging and shattering scores (Arelli et al., 2015).

Thus, the aims of the present study were to map QTL associated with HG Type 5 (SCN race 3) resistance in PI 567488B using an $F_{2:3}$ population derived from the cross of 'Lee' (susceptible) × PI 567488B (resistant), and to develop breeder-friendly robust SNP markers for marker-assisted selection.

Materials and Methods

Population development

A $F_{2:3}$ population was developed by crossing the susceptible line 'Lee' with PI 567488B that possesses the HG Type 5 resistance. 'Lee' (PI 548655) is a maturity group VI cultivar derived from the cross of 'S-100' and 'CNS' (USDA-GRIN). PI 567488B is the Chinese plant introduction in maturity group IV, which is resistant to multiple SCN races: 1, 2, 3, and 5 (USDA-GRIN). The cross of 'Lee' × PI 567488B was made in summer 2015 at Plant Science Farm of the University of Georgia, Athens, GA. The F_1 plants were grown in the winter nursery in Puerto Rico during 2015-2016. F_2 plants were grown in summer 2016 at the University of Georgia Iron Horse Farm near Athens and harvested individually to form 141 $F_{2:3}$ families.

Phenotyping the population for HG Type 5 (race 3) resistance

One hundred and forty-one $F_{2:3}$ families were evaluated for HG Type 5 (race 3) resistance in the Plant Pathology greenhouse at the University of Georgia. Two parents ('Lee' and PI 567488B), five susceptible checks ('NC Roy', 'NC Raleigh', 'CNS', 'Hutcheson' and 'Lee 74') and seven indicator lines ('Peking', PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, PI 548316) were included for evaluation. All plants were grown in 10.2-cm clay pots that were filled with fumigated sandy loam soil. Pots were arranged into a randomized completed block design (RCBD).

Two experiments were then conducted with six replications each, so each family had 12 replicates, corresponding to 12 plants. The HG Type 5 (SCN race 3) population was originally collected from Collins, GA, and then cultured on susceptible cultivar 'Haskell'. Each seedling, representing a single replicate, was inoculated with 2,000 eggs HG Type 5 in 3-4 mL of water with a dispenser machine at 7 d after planting. Approximately 60 d after inoculation, when the cyst counts on 'Lee 74' was around 100 or greater, soybean roots were individually washed free of soil, and then flushed with water to collect cysts through 20 and 60 µm aperture sieves. The nematode cysts for individual plants were counted under a stereoscope. The level of resistance was classified based on the female index (FI) using the formula: FI = [(average number of cysts on a given individual/ average number of cysts on 'Lee 74')] × 100% (Niblack et al., 2002).

The FIs among $F_{2:3}$ families were tested for normality using JMP software (SAS Institute Inc., Cary, NC). The Shapiro-Wilk statistic was used to test for normality with null hypothesis holding that FI is normally distributed; if the calculated *P* value is greater than 0.05, then the null hypothesis is accepted and the FI index is deemed normally distributed. Additionally, a mixed model analysis was performed using R package (cran-rproject.org) with three variables: genetic effect (fixed), and experiment and block (random effects). Broad sense heritability was calculated using ANOVA results with the following formula:

$$H^2 = \frac{\sigma_G^2}{(\sigma_G^2)+(\sigma_e^2)/r}$$

where H^2 = broad sense heritability, σ_G^2 = genotypic variance of $F_{2:3}$ families, σ_e^2= error variance, and r = number of replicates. To minimize the effects of environment, best linear unbiased estimator (BLUE) values of individual family were calculated using the R package lme4 (cran-rproject.org/lme4) and then BLUE values were subsequently used for QTL analysis.

DNA extraction and genotyping

Young leaves from 10-12 plants per family as well as leaves from each parent were collected as a bulk from greenhouse. The bulked leaf tissue sample from each family was lyophilized for 48 h and then ground into a powder using a GenoGrinder (SPEX Sample Prep, Metuchen, NJ, USA). DNA extractions were performed using a modified CTAB method (Keim et al., 1988), and then diluted with TE buffer to obtain a final concentration of ~50 ng μL^{-1}. Genotyping of the 141 $F_{2:3}$ families and their parents 'Lee' and PI 567488B was performed using the SoySNP6K iSelect Bead Chip at the Soybean Genomics and Improvement lab, USDA ARS, Maryland. Genotypes were called using GenomeStudio 2.0. (Illumina, San Diego, USA).

QTL analysis

Linkage map was created using JoinMap 4.1 (Van Ooijen, 2006) with LOD (likelihood-odds) = 5 as a threshold using the maximum likelihood mapping algorithm. Recombination fractions were converted into genetic distance (cM) using the Kosambi mapping function. The linkage groups were each assigned a chromosome number based on the soybean reference genome of cultivar 'Williams 82'(soybase.org).

The BLUE values of FI for each family was used for QTL detection of SCN resistance using a composite interval mapping (CIM) method in Windows QTL Cartographer 2.5 (Wang et al., 2007). The Model 6 was selected with control marker numbers of five (cofactor) and a window size of 10 cM. A permutation test was performed with 1,000 runs, a 0.5 cM walking speed, and significance set at $\alpha = 0.05$. The highest LOD score on a given chromosome was used to indicate the QTL position. A diagram displaying the SNP markers, their cM positions and respective chromosomes and the QTLs that were detected were generated using the MapChart option in JoinMap 4.1 (Van Ooijen, 2006).

Results and Discussion

Phenotypic variation

The reactions of soybean indicator lines and each parent to HG Type 5 (SCN race 3) are summarized in Table 3.1. The phenotypic variance for the $F_{2:3}$ families (Figure 3.1) showed large genetic variation for SCN resistance. The mean FI (%) for HG Type 5 among all $F_{2:3}$ families was 42.3% with a range of 3.0 -85.8%. The Shapiro-Wilk test for normality revealed that FI was normally distributed (*P* value = 0.36). The skewness value of 0.12 was also greater than 0, indicating that the distribution was slightly skewed to the right (the lower FI%). The broad sense heritability estimate for FI of HG Type 5 resistance was 26%, indicating that the error variance rather than genetic variance contributed to a larger part of phenotypic variation, which might be an explanation for the low phenotypic variance explained (8.0 – 12.7 %) by each significant QTL that were detected for HG Type 5 resistance in the PI 567488B-derived mapping population. Another explanation for the large error variance could be a large environmental effect, which might also suggest that SCN reaction was significantly affected by environmental conditions. Previous studies had similar results (Schuster, 2001; Ferreira, 2011), indicating that environmental factors need to be better controlled to attribute the variation observed to genotype.

Genetic linkage analysis

Of 6,000 SNPs used to genotype the $F_{2:3}$ mapping population, 1,817 SNPs were polymorphic and thus were utilized to construct the genetic map. The largest number of markers were detected on Chrs 8 (LG A2, 140) and 13 (LG F, 140) whereas the lowest number of markers was detected on Chr 12 (LG H, 33). In total, 20 linkage groups were assembled representing all 20 soybean chromosomes with a total map distance of approximately 3,249 cM. The average genetic distance between markers was 1.78 cM and generally the marker order was

in agreement with the physical mapping positions. However, some differences were observed in some regions including the region on Chr 20 where a significant QTL was detected. Possible explanations for the discrepancy between marker order on the linkage map vs the physical position of markers from the 'Williams 82' reference genome could be due to parental genetic background, population size, population type, or the accuracy of genotyping.

Detection of QTLs for HG Type 5 (SCN race 3) resistance

Based on a genome-wide permutation test, a LOD threshold of 3.4 (α= 0.05) was used to identify significant QTLs for HG Type 5 (race 3) resistance. CIM indicated two QTLs were significantly associated with HG Type 5 (race 3) resistance. The QTLs were mapped to Chrs 16 (LG J) and 20 (LG I), respectively. No significant QTLs were identified on Chrs 8 (LG A2) or 18 (LG G) where the *Rhg4* and *Rhg1* loci are located, which was consistent with the *rhg1* and *Rhg4* SNP genotyping results.

The first QTL detected was identified at marker Gm20_37746786 on Chr 20 (LG I), which explained 12.7% of the total phenotypic variation for resistance to HG Type 5 (race 3) (Figure 3.2; Table 3.2). Based on the 1-LOD drop off from the peak LOD score confidence interval which determines the cM region by the peak LOD minus 1 (Lander et al., 1989), the QTL position is located between SNPs Gm20_38578470 and Gm20_37857633 on Chr 20 (LG I). The total phenotypic variance explained at this QTL region varied from 10.8-12.7%. According to the information from Soybase (SoyBase.com), there are three previously mapped QTLs for SCN resistance on Chr 20 (LG I) that were reported in 'Peking', PI 437655, PI 437654, and PI 464925B. Qiu et al. (1999) detected one significant RFLP marker K011 on Chr 20 in 'Peking' for SCN race 5 resistance. Based on SoyBase data, this region is estimated to fall between markers Satt239 (24,593,095 bp) and BARC-027790-06672 (34,074,375 bp). Vuong et

al. (2015) also reported a SCN resistance QTL on Chr 20 through GWAS, which overlapped with the one previously detected in PI 437654 using linkage mapping (Wu et al., 2009) and located between markers BARC-044361-08677 (41,563,793 bp) and BARC-042685-08348 (43,959,781 bp). Winter et al. (2007) reported a different QTL on Chr 20, associated with resistance to HG Type 7 (SCN race 3) in a wild soybean accession (PI 464925B). They used a mapping population derived from two moderately resistant lines 'S08-80' and PI 464925B, with results revealing that the QTL on Chr 20 explained just 7.6% of the phenotypic variation (Winter et al., 2007). The QTL detected on Chr 20 by Winter et al. (2007) was estimated to be located between BARC-050455-09643 (36,575,544 bp) and Satt292 (40,623,814 bp) based on Soybase.org, which might overlap with the QTL detected on Chr 20 in the present study. Therefore, our result provides confirmation of the previously reported QTL and suggests that this SCN resistance QTL may not be unique to *G. soja*. Further analysis is required to determine if this QTL carries the same or different resistance alleles as the wild soybean accession PI 464925B.

The second QTL identified in the present study was located at the distal end of the long arm of Chr 16 (LG J) (Figure 3.2). SNP marker Gm16_35846 on this chromosome was the nearest marker to the peak LOD and was significantly associated with HG Type 5 resistance (SCN race 3) resistance. The LOD score for the QTL was 3.4, which explained 8% of phenotypic variance for SCN resistance (Table 3.2). Based on data obtained from SoyBase, there are seven QTLs associated with SCN resistance on Chr 16 (LG J); six of them overlapping, and are located between the 30 -37 Mb chromosomal region on Chr 16 (LG J) (soybase.com) while the QTL designated as *qSCN001-04* from PI 438489B on Chr 16 (LG J) was the closest to our

detected QTL, which was ~35 kb away. Due to the low LOD score and small phenotypic variance observed for this QTL, further study is required for confirmation.

Conclusions

The ability of SCN populations to evolve rapidly is problematic for soybean breeders as well as nematologists. In SCN greenhouse screening assays conducted at the University of Georgia, we identified a shift of the SCN race 3 population based upon indicator line PI 209332 (carrying *rhg1*) having a susceptible reaction to SCN race 3. Therefore, the new population was designated HG Type 5 based on Niblack's HG Type determination system. To our knowledge, this is the first SCN study documenting HG Type 5 for SCN race 3. Similarly, this study also represents the first QTL mapping study conducted for HG Type 5 (SCN race 3).

Our results further indicate that PI 567488B possesses resistance to multiple HG Types and carries desirable agronomic attributes, as noted above. Based on the mapping results using the present $F_{2:3}$ population, two minor effect QTLs are responsible for HG Type 5 resistance: Gm16_35846 and Gm20_37746786. Although the phenotypic variance observed for SCN resistance was relatively low, the results might be beneficial for the development of new KASP markers to deploy marker-assisted selections and will allow for the stacking of desirable resistance alleles into new cultivars to improve SCN resistance in soybean.

References

Arelli, P.R., A. Mengistu, R.L. Nelson, S.R. Cianzio and T. Vuong. 2015. New Soybean Accessions Evaluated for Reaction to Populations. Crop Sci. 55: 1236. doi:10.2135/cropsci2014.10.0696.

Concibido, V.C., B.W.Diers and P.R.Arelli. 2004. A decade of QTL mapping for cyst nematode resistance in soybean. Crop Sci. 44: 1121-1131.

Cook, D.E., T. G. Lee et al. 2012. Copy number variation of multiple genes at *rhg1* mediates nematode resistance in soybean. Science 338: 1206-1209.

Ferreira, M. F. S. et al. 2011. QTLs for resistance to soybean cyst nematode, races 3, 9, and 14 in cultivar Hartwig. Pesquisa Agropecuasria Brasileria. 46 (4): 420-428. https://dx.doi.org/10.1590/S0100-204X2011000400012

Glover, K.D. 2004. Near isogenic lines confirm a soybean cyst nematode resistance gene from PI 88788 on linkage group J. Crop Sci. 44: 936-941.

Guo, B., D.A. Sleper, H.T. Nguyen, P.R. Arelli and J.G. Shannon. 2006. Quantitative Trait Loci underlying Resistance to Three Soybean Cyst Nematode Populations in Soybean PI 404198A. Crop Sci. 46: 224. doi:10.2135/cropsci2004.0757.

Ichinohe, M. 1959. Studies on the soybean cyst nematode *Heterodera glycines* and its injury to soybean plants in Japan. Plant Pis Reptr Suppl. 260: 238.

Jackson, G. 2014. Soybean cyst nematode. Mississippi State University Extension Service Publication 1293.

Kazi, S., J. Shultz, J. Afzal, R. Hashmi, M. Jasim, J. Bond, et al. 2010. Iso-lines and inbred-lines confirmed loci that underlie resistance from cultivar 'Hartwig' to three soybean cyst nematode populations. Theor Appl Genet. 120: 633-644. doi:10.1007/s00122-009-1181-4.

Kim, K.S., T.D. Vuong, D. Qiu, R.T. Robbins, J. Grover Shannon, Z. Li, et al. 2016. Advancements in breeding, genetics, and genomics for resistance to three nematode species in soybean. Theor Appl Genet.129: 2295-2311. doi:10.1007/s00122-016-2816-x.

Keim P., Olson T.C., Shoemaker R. C. (1988) A rapid protocol for isolating soybean DNA. Soybean Genet Newsl 15:150–152

Koenning, S.R. and J.A. Wrather. 2010. Suppression of Soybean Yield Potential in the Continental United States by Plant Diseases from 2006 to 2009. Plant Health Prog. doi:10.1094/php-2010-1122-01-rs.

Lander, E. S., and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 121: 185-199.

Liu, S., P.K. Kandoth, S.D. Warren, G. Yeckel, R. Heinz, J. Alden, et al. 2012. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. Nature 492: 256-260. doi:10.1038/nature11651.

Melito, S., A.L. Heuberger, D. Cook, B.W. Diers, A.E. MacGuidwin and A.F. Bent. 2010. A nematode demographics assay in transgenic roots reveals no significant impacts of the *Rhg1* locus *LRR-Kinase* on soybean cyst nematode resistance. BMC Plant Biol 10: 104. doi:10.1186/1471-2229-10-104.

Mudge, J.P., B. Cregan, J.P. Kenworthy, W. J. Kenworthy, J. H. Orf, and N. D. Young. 1997. Two mucrosatellite markers that flank the major soybean cyst nematode resistance locus. Crop Sci. 37: 1611-1615.

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

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

Riggs, R.D. 1975. Worldwide distribution of soybean-cyst nematode and its economic importance. Joint Annu. Meet. Soc. Nematol. Am. Phytopathol. Houston, Texas, U.S.A.

Schuster, I. R. V. Abdelnoor, S. R. R. Marin, et al. 2001. Identifcation of a new major QTL associated with resistance to soybean cyst nematode (*Heterodera glycines*). Theor. Appl. Genet. 102: 91-96.

Shi, Z., S. Liu, J. Noe, P. Arelli, K. Meksem and Z. Li. 2015. SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. BMC Genomics 16: 314. doi:10.1186/s12864-015-1531-3.

Tylka, G.L. and C.C. Marett. 2014. Distribution of the Soybean Cyst Nematode, Heterodera glycines, in the United States and Canada: 1954 to 2014. Plant Health Prog. doi:10.1094/php-br-14-0006.

Tylka, G.L., C. C. Marett, and M. P. Mullaney. 2016. Evaluation of Soybean varieties resistant to soybean cyst nematode in Iowa-2016. Iowa State University Extension and Outreach.

Van Ooijen J.W. (2006) JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen

Vuong, T.D., D.A. Sleper, J.G. Shannon and H.T. Nguyen. 2010. Novel quantitative trait loci for broad-based resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe) in soybean PI 567516C. Theor Appl Genet. 121: 1253-1266. doi:10.1007/s00122-010-1385-7.

Qiu, B. X., P. R. Arelli, D. A. Sleper. 1999. RFLP markers associated with soybean cyst nematode resistance and seed composition in a 'Peking' x 'Essex' population. Theor Appl Genet. 98: 356-364.

Wang, D., P. R. Arelli, R. C. Shoemaker, and B. W. Diers. 2001. Loci underlying resistance to race 3 of SCN in glyine soja plant introduction 468916. Theor Appl Genet. 103.

Wang, S., C. J. Batsen, and Z. B. Zeng. 2007. Windows QTL Cartographer 2.5. Dep. of Statistics, North Carolina State Univ. Raleigh, NC.

Wang, X., G. Bergstrom, S. Chen, D. M. Thurston, J. Cummings, Z. Handoo, et al. 2017. First Report of the Soybean Cyst Nematode, *Heterodera glycines*, in New York. The Am. Phytopathol. Soc. 101 (11): 1957.

Weisemam, J.M., B. F. Matthews, and T. E. Devine. 1992. Molecular markers located proximal to the soybean cyst nematode resistance gene, *Rhg4*. Theor Appl Genet. 85: 136-138.

Winter, S. M., B. J. Shelp, T. R. Anderson, T. W. Welacky, I. Rajcan. QTL associated with horizontal resistance to soybean cyst nematode in *Glycine soja* PI 464925 B. Theor Appl Genet. 114 (3): 461-472.

Wu, X., S. Blake, D. A. Sleper, J. G. Shannon, P. Cregan, H. T. Nguyen. 2009. QTL, additive and epistatic effects for SCN resistance in PI 437654. Theor Appl Genet. 118: 1093-1105.

Zheng, J.W., and S.Y.Chen. 2011. Estimation of virulence type and level of soybean cyst nematode field populations in response to resistant cultivars. J. Nematol. 3: 37-43.

Tables and Figures

Table 3.1. Female index of seven indicator lines in HG Type determination test.

	'Peking'	PI 88788	PI 90763	PI 437654	PI 209332	PI 89772	PI 548316
	1	2	3	4	5	6	7
	Female index (%)[a]						
Experiment 1	5.2	7.5	1.4	1.0	61.1	0.9	9.1
Experiment 2	2.0	1.1	0	0	12.1	0	1.6
HG Type 5	-	-	-	-	+[b]	-	-

[a] Female index is the ratio between the number of cyst on each indicator line and the number of cyst on 'Lee 74' in percentage.
[b] Based on reactions of indicator lines, SCN population is classified as HG Type 5 (SCN race 3).

Table 3.2 QTLs associated with HG Type 5 resistance in the $F_{2:3}$ population derived from 'Lee' x PI 567488B.

Chr	**Interval**	**Significant SNPs**	**LOD**	**R^2**
16	Gm16_35846	Gm16_35846	3.4	8.0
20	Gm20_38578470-Gm20_37857663	Gm20_37746786	5.1	12.7

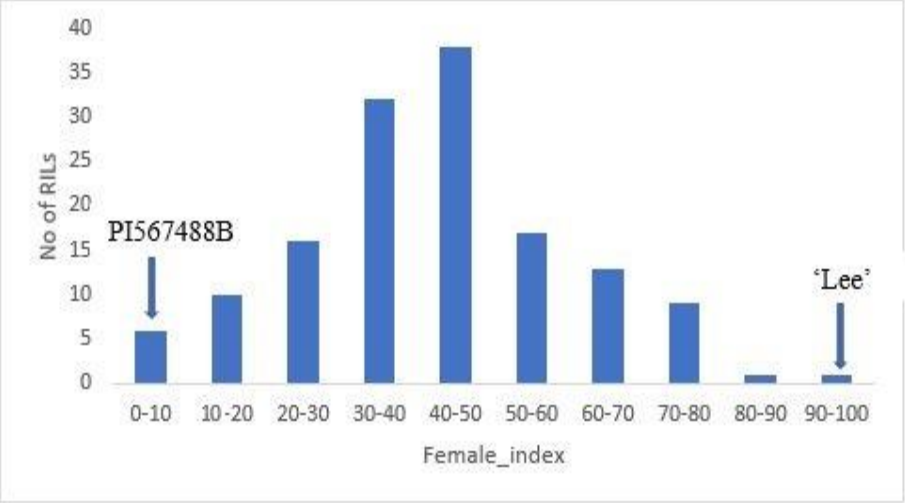

Figure 3.1. Distribution of female indices of 141 $F_{2:3}$ families derived from 'Lee' and PI 567488B

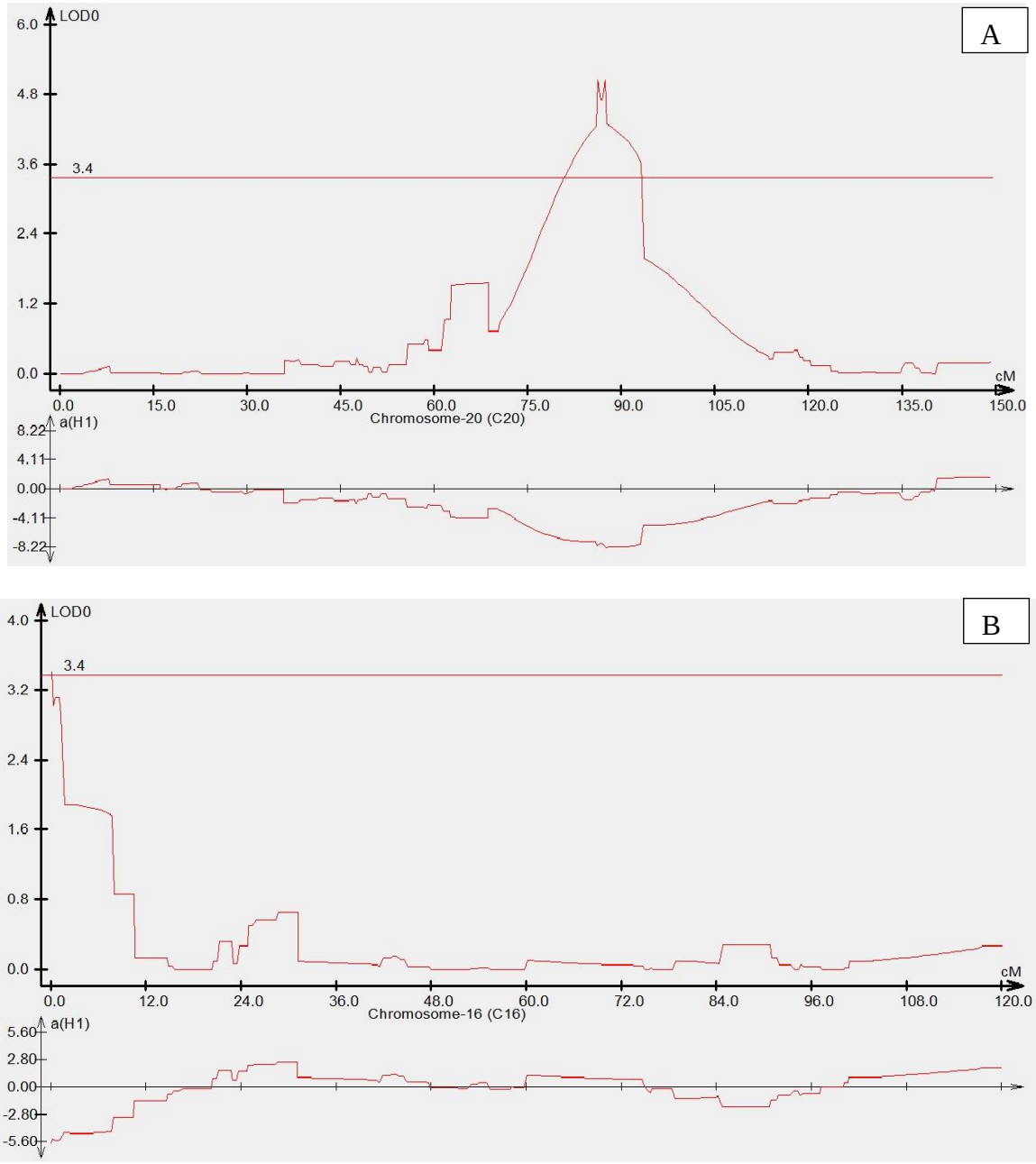

Figure 3.2. Likelihood of odds (LOD) plots for Chrs 20 (A) and 16 (B) indicating genomic location of QTL responsible for resistance to HG Type 5.

Note: X-axis: position in cM; Y-axis: LOD score.

CHAPTER 4

CONCLUSIONS

The soybean cyst nematode (SCN, *Heterodera glycines*) has been the most destructive pest for soybean production in the U. S. SCN was first identified in the U.S in 1954 in North Carolina (Riggs, 1975) and currently has spread to most of soybean producing states (Tylka and Marett, 2014; Wang et al., 2017). Utilizing resistant cultivars has been a primary method for SCN management. However, limited genetic resistant sources (Peking and PI 88788) have been extensively utilized to develop SCN resistant cultivars. Thus, the risk of SCN populations overcoming their resistance has occurred. The objectives of this research were to: 1) screen soybean germplasm to search for new sources of SCN resistance that are different from 'Peking' and PI 88788 types; and 2) to identify QTLs derived from these new sources of SCN resistance using GWAS and bi-parental mapping approaches.

To discover novel source of resistance to SCN race 3, the predominant race in U. S. soybean production (Jackson, 2014), 462 soybean accessions originating from various origins were screened with HG Type 0 (SCN race 3) in the greenhouse and genotyped using three functional markers that were developed at two major loci: *Rhg1* and *Rhg4* (Shi et al., 2015). Using these SNP markers helped us identify resistant accessions that do not carry same resistance alleles as 'Peking' and PI 88788. By combining both phenotyping and genotyping results, 58 soybean accessions were identified as putatively novel sources of resistance because they did not possess the same resistance alleles as 'Peking' and PI 88788 at *Rhg1* and *Rhg4* loci. They also have been confirmed with haplotype analysis at two these loci assembled with

SoySNP50k Infinium Chip data. In addition, all soybean accessions were also genotyped for southern root-knot nematode resistance with a functional marker GSM039A reported by Pham et al. (2013). Of 58 SCN resistant accessions identified, 18 soybean accessions were predicted to be resistant to southern root-knot nematode based on SNP marker GSM039A.

To elucidate the genetic basis of HG Type 0 resistance, a genome-wide association study (GWAS) was performed using a panel of 461 soybean accessions that were genotyped with SoySNP50k Infinium Chips. The GWAS identified 13 significant SNPs from five genomic regions located on Chrs 2, 7, 8, 10, and 18 that are significantly associated with SCN resistance. The genomic regions on Chrs 8 and 18 were known to be *Rhg4* and *Rhg1* loci. Favorable alleles and 30 candidate genes responsible for SCN resistance from these three genomic regions on Chrs 2, 7, and 10 were reported.

Based on germplasm screening results, bi-parental populations were formed from the resistant accessions identified. PI 567488B was chosen to develop a mapping population with 'Lee', an SCN susceptible cultivar. This PI originated from China and is resistant to multiple SCN races including HG Type 0 and HG Type 5 (SCN race 3). Genotyping results with functional markers at *Rhg1* and *Rhg4* indicated that PI 567488B does not carry resistance alleles at these two major loci. One hundred and forty-one $F_{2:3}$ families derived from 'Lee' x PI 567488B were evaluated using HG Type 5 (SCN race 3) with 12 replicates. The population was genotyped using SoySNP6k Infinium Chips. QTL analysis identified two genomic regions on Chrs 16 and 20, respectively, significantly associated with SCN resistance. Both QTLs were overlapped with other reported QTL: one on Chr 20 mapped in a wild soybean accession PI 468925B and other on Chr 16 identified in PI 438489B. The results will enable soybean breeders

to perform marker-assisted selections for SCN resistance and to introgress or stack these confirmed resistance alleles into elite lines to improve SCN resistance in soybean.