GENETIC MAPPING OF QUANTITATIVE TRAIT LOCI FOR AND EXPLORING THE INTERACTION OF ROOT MICROBIOME WITH SOYBEAN CYST NEMATODE RESISTANCE

by

DUNG THUY TRAN

(Under the Direction of Zenglu Li)

ABSTRACT

Soybean cyst nematode (SCN) is the most damaging pest causing yield loss with more than $1.5 billion in soybean production in the U.S. However, the majority of resistant cultivars have their resistance derived from Plant introduction (PI) 88788 and 'Peking' carrying two major quantitative trait loci (QTLs) *Rhg1* and *Rhg4.* Using limited genetic sources of resistance has caused SCN to overcome their resistance, therefore, it is critical to identify additional major QTLs to slow down SCN adaptation. PI 567295 was found as resistant to multiple SCN populations and a QTL on Chr 10 was found associated with resistance to two *Heterodera glycines* (HG) type 0 and 1.2.-. Two minor QTLs on Chrs 6 and 8 were also mapped for HG type 1.2.- resistance. The result provides an additional resistant source and SNP markers for developing SCN resistant cultivars. Besides, the root microbiome plays an important role in plant growth and health, so understanding the composition of the root microbiome is necessary to discover beneficial bacteria related to significant pathogens and pests such as SCN. Two field tests in Plains, GA, and a greenhouse test with SCN pressure were conducted to describe the root

bacteria microbiome of bulked soil, rhizosphere, and root of soybean. The core of bacteria in soybean roots was identified and some beneficial bacteria might contribute to SCN management.

INDEX WORDS: *Glycine max*, *Heterodera glycines* (HG), Soybean cyst nematode (SCN), Resistance to *Heterodera glycines gene (Rhg)*, quantitative trait loci (QTL), soybean root microbiome.

GENETIC MAPPING OF QUANTITATIVE TRAIT LOCI FOR AND EXPLORING THE INTERACTION OF ROOT MICROBIOME WITH SOYBEAN CYST NEMATODE RESISTANCE

by

DUNG THUY TRAN

BS, Vietnam National University of Agriculture, Vietnam, 2013

MS, University of Georgia, 2017

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2022

GENETIC MAPPING OF QUANTITATIVE TRAIT LOCI FOR AND EXPLORING THE INTERACTION OF ROOT MICROBIOME WITH SOYBEAN CYST NEMATODE RESISTANCE

by

DUNG THUY TRAN

Major Professor:	Zenglu Li
Committee:	Melissa Mitchum
	Jason Wallace
	Peng Chee
	Cecilia McGregor

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2022

DEDICATION

Dedicated to my beloved mom, Vui Le – thank you for all your love and support me to do my best and pursue my dream.

ACKNOWLEDGEMENTS

I am deeply grateful to many different people that have supported me in my professional development and in completing my PhD research. I wish to thank my advisor, Dr. Zenglu Li for his support and mentorship for my both professional and personal development. I also want to thank Dr. Melissa Mitchum for guiding me more about soybean cyst nematode and supporting me throughout my PhD. Many thanks to Dr. Jason Wallace for his advice about statistical methods of the microbiome project; and his encouragement and support for me to learn more about coding skills. I am thankful to Dr. Peng Chee and Dr. Cecilia McGregor for great advice to complete my research. I want to extend my sincere gratitude to Steve Finnerty and Ben Averitt for spending much time setting up greenhouse tests with me. Also, I want to thank Tatyana Nienow and Nicole Bachleda for all their technical support in the soybean molecular breeding lab and their help in greenhouse experiments. I would also like to thank Dale Wood, Brice Wilson, Earl Baxter, Brian Little, and Greg Gokalp for helping me develop mapping populations and setting field tests. Thank you to Alexandra Ostezan, Renan Souza, Mark Miller, Sam McDonald, Liz Prenger, and Brooks Arnold for their help in my research and their friendship. They made the soybean breeding program as a great place to work. Importantly, I would like to acknowledge my great lab mate, classmate and friend, Ethan Menke for all his support not only in my research and also my life in Athens. Many thanks to my best friends, Claire Willis and Linh Le from whom I have received a lot of encouragement. Finally, I would like to thank all my friends at UGA and in my hometown for all their love and encouragement during my time in UGA.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Soybean introduction

Soybean [*Glycine max* (L.) Merrill], originating in China, was introduced to North America in 1765 by Samuel Bowen, a former sailor (Hymowitz and Harlan, 1983) but was not planted in the Midwest until 1851 (Hymowitz, 1987). Soybean was mainly a forage crop until the mid-20th century, then it shifted from a forage crop to a grain crop when the demand of soybean oil increased (Hart, 2017). Currently, soybean has been the second most frequently grown crop after maize in the U.S., covering approximately 33.6 million hectares and grossing $30.5 billion in 2019/ 2020 (Soystat, 2021). The average soybean yield in the U.S. for the year 2020 was 3.4 metric tons per hectare and soybean production in 2020 was 112.5 Million Metric Tons (MMTs), up 16% from 2019 (Soystat, 2021). The soybean has been grown in around 30 states in the U.S. and Illinois produced the most soybean in 2020 with 16.5 MMTs followed by Iowa and Minnesota with 13.4 and 9.8 MMTs, respectively (Soystat, 2021). In 2019, the U.S ranked the second in world soybean production with 96.7 MMTs, after Brazil with 128.5 MMTs.

There are many uses of soybean, which can be classified into three main categories: oil, meal, and whole bean. Soybean oils are used for cooking oil, margarine, biodiesel, and industry products, accounting for 59% of oilseed production worldwide (Soystat, 2021). Soybean is a rich source of protein, accounting for nearly 40% of dry soybean weight (Wilson, 2004). Besides, the protein from soybean is unique because it is the only plant-based food with all eight essential amino acids (Ali, 2010). About 70% of world protein meal consumption was from soybean in

2020, followed by rapeseed at 12% and sunflower at 6% (Soystat, 2021). Poultry is the main consumer of soybean meal, accounting for 55% of soybean meal consumption (Soystat, 2021). Besides being used in the animal feed industry, soybean is used in human food as well, including tofu, soymilk, edamame, and sprouts and providing numerous health benefits including a high protein content, being cholesterol free and a rich source of antioxidants (Rizzo and Baroni, 2018).

Soybean is an annual short-day plant, which is classified into 13 different maturity groups. The maturity groups are strongly affected by the environmental conditions in each growing region in the U.S. including day length and temperature (Zhang et al., 2007). There are three types of growth habits: determinate, semi-indeterminate, and indeterminate (Bernard, 1972). Soybeans in early maturity groups 000 to IV are typically indeterminate (Carter Jr. et al., 2004) which indicates that the vegetative part continues to grow after flowering (Carlson and Lersten, 2004). Later maturity groups from MG V to X are typically determinate, which means vegetative growth stops at flowering (Carlson and Lersten, 2004). The semi-indeterminate soybean cultivars have shorter main stems than indeterminate cultivars but longer stems than determinate (Kato et al., 2019). The vegetative growth stage begins with the emergence of two cotyledons, followed by the growth of primary unifoliate leaves that are ovary-shaped and located oppositely. Finally, trifoliate leaves emerge. The reproductive stage begins with flowering and proceeds until pod maturity. The soybean flower is a perfect papilionaceous flower with four complete parts: five sepals, four petals, nine stamens, and one pistil, ensuring highly self-pollinated. The flowers develop into pods with two pods or more per inflorescence. Each plant generally produces up to 80 pods with the number of seeds per pod ranging from 1 to 5 (Singh, 2017).

Soybean is a species in the *Fabaceae* family, which contains 650 genera and 18,000 species including other beans and alfalfa (Zhu et al., 2005). It is a member of the *Glycine* genus consisting of two sub genera: Glycine Willd and Soja (Moench) F. J. Hermann, with a total of 28 species (Chang et al., 2014). The subgenus Soja includes the domesticated soybean, *G. max* which is a diploid [2n=40] with 20 pairs of chromosomes and has an estimated genome size of 1,115 Mb with 46,430 predicted protein codes based on the reference sequence generated from Williams 82 and Lee cultivars (Schmutz et al., 2010; Valliyodan et al., 2019).

Soybean cyst nematode introduction

Soybean cyst nematode history and damage

Improving yield is the ultimate goal in soybean breeding and soybean cyst nematode (SCN, *Heterodera glycines*) is the most damaging pest in soybean production that caused yield losses of 3.4 million tonnes in 2018 (Crop protection network, 2021), equaling a loss of one billion dollars. SCN attacks soybeans and establishes a feeding site acting as a metabolic sink that takes all the nutrients necessary from plants for SCN development (Mitchum, 2016). Yield loss can be up to 40% even when above ground symptoms such as stunting and yellowing are not clearly present due to a low SCN population, so growers are slow to respond to SCN infection in fields (Wang et al., 2003; Winsor, 2020). These symptoms appear when SCN population densities are high, but symptoms may be confused with other nutrient deficiency problems (Smith et al., 2001). Therefore, these symptoms may not be an obvious and reliable method to detect SCN. In order to identify an SCN infection, plants should be removed, or soils should be collected to observe below ground symptoms such as cysts (Davis and Tylka, 2000). Cysts are dead females that are lemon-shaped and less than 1mm in diameter (Davis and Tylka, 2000). Moreover, SCN can interact and even increase the damage by other soilborne pathogens

including sudden death syndrome and brown stem rots (Tabor et al., 2003; Xing and Westphal, 2006)

The SCN was first discovered in China in 1899 (XongHong et al., 1997), then reported in Japan in 1915 (Davis and Tylka, 2000), and later was first discovered in the U.S. in 1954 in North Carolina (Riggs, 1977). The SCN spread to Tennessee and Missouri in 1956, and Arkansas and Kentucky in 1957 (Riggs, 1977). It was thought that the SCN was introduced into the U.S. with *Rhizobium* (Noel, 1986); then SCN damage became noticeable after large scale soybean production occurred. The SCN cannot move further by itself, but it can be transferred along with soil, so anything that moves infected soil, can move SCN (such as through water, wind, erosion, farm equipment, and even animals) (Davis and Tylka, 2000). Up to now, the SCN is found in every soybean-producing state except West Virginia (Tylka and Marett, 2021) (Fig. 1.1).

Life cycle of SCN

The SCN is an obligate endoparasite, feeding on living plants with a worm-like shape and an unsegmented invertebrate body. The morphology of SCN changes during its life cycle. There are five stages and four molts in the SCN's life cycle, ranging from 25 to 40 days depending on environmental conditions and the availability of hosts (Fig. 1.2). SCN enters inside the root at the second juvenile stage (J2), 375-520 µm in length and 18 µm in diameter while the first juvenile stage (J1) starts inside the egg, and molts into J2 prior to hatching (Davis and Tylka, 2000). At the J2 stage, SCN is worm-shaped with an offset head and tapering tail. The J2 stage develops a stylet to pierce the eggshell, then travels through soil in search of plant roots through chemoreception (Perry, 1996). Using its stylet to penetrate the root and intracellularly to the root vascular tissue, the J2 also uses its stylet to thrust and release secreted proteins from its secretory

glands that modify the plant nucleus and other organelles (Niblack et al., 2006). Cell walls of neighboring cells will dissolve and then fuse into a large cytoplasmic feeding space called the syncytium to feed the plants. Once the feeding site within the host has been established, the J2 migrates and grows to the J3 and J4 as a sausage-like shape and continues feeding on nutrients from plants. The J2 will not continue to grow and develop if the syncytium is not maintained (Davis and Tylka, 2000). The J3 and J4 undergoes sex differentiation: males and females having different shapes. Sex determination is not genetic; it is affected by environmental factors including: temperature, infection density, and host nutrition (Colgrove and Niblack, 2005). Adult males have a vermiform shape and do not feed on the root; a male exists in the root to search for a female. The female will remain sedentary within the root; and will become larger as a pyriform shape so that the body including the vulva is displayed on the root exterior where it emits pheromones to attract the male. After fertilization, the female dies and its body, called a cyst, hardens to protect the eggs. Based on maturity level, the cyst varies in color from white, yellow to brown (Niblack et al., 2006). Besides dormant eggs forming into cysts, some eggs are left outside in a gelatinous matrix known as an egg sac (Niblack et al., 2006). The female can produce up to 600 eggs depending on the host (Davis and Tylka, 2000). Dormant eggs (which are protected inside cysts) can survive for up to 10 years or more in the soil, making it very difficult to control this disease (Winsor, 2020).

SCN classification system

Virulence of SCN is variable with 16 possible races which was firstly classified in 1970 by using a race test based on SCN response of four indicator lines including ‘Peking’, ‘Pickett’, PI 88788, and PI 90763 (Table 1.1) (Golden and Et, 1970). Races are determined by measuring and comparing the number of cysts on these four lines with the susceptible check ‘Lee’ which is

defined as a female index (Riggs and Schmitt, 1988). Female index (FI) is calculated as the ratio of number of cysts on a soybean line to the number of cysts on a susceptible check in percentage. If any indicator line has an FI of greater than 10%, then this line is considered to be susceptible, indicated with a "+" sign, while a line with an FI of lower than 10% is considered resistant, negative reaction (Riggs and Schmitt, 1988). However, this classification system was not suitable to capture proper variability of SCN because of its genetic diversity. Therefore, a new classification system named HG type was developed with HG as the abbreviation of the scientific name of SCN (*Heterodera glycines*) to classify SCN virulence based on the number of cysts on seven indicator lines instead of four in the old classification system (Niblack et al., 2002). Besides, the susceptibility check 'Lee' was replaced with 'Lee 74' or 'Williams 82' because the number of cysts on 'Lee' was variable. Seven indicator lines are designated as numbers from 1 to 7 based on the order of reported resistant reaction of these lines (Table 1.2). For example, an HG type 0 means no indicator lines have an FI greater than 10%; by contrast, HG type 1.2 indicates that both 'Peking' (No. 1) and PI 88788 (No. 2) have FI greater than 10%. The new classification system is considered to be more accurate because it can capture new SCN populations. New indicator lines can be included to characterize new SCN populations. Normally, 'Pickett' (PI 548988), an indicator line from the old system is usually included in SCN population designation with the new classification system to be more informative with HG type and SCN race names. For example, HG type 0 or HG type 7 can be SCN race 3. A modification of the HG type test including three lines instead of seven lines: 'Peking' (No. 1), PI 88788 (No. 2) and PI 437654 (No. 4) can be used because previous studies reported that there is a high correlation of SCN reaction among 'Peking' (No.1), PI 90763 (No.3) and PI 89772 (No.

6); and, among PI 88788 (No. 2), PI 209332 (No. 5), and PI 548316 (No. 7) (Niblack et al., 2002).

Sources of resistance to SCN

Utilization of SCN-resistant cultivars and rotation with non-host crops are the most effective methods to control SCN. Screening soybean germplasms for SCN resistance has been performed in the U.S. since the 1950s. The initial sources of resistance to SCN identified were PI 90763, PI 84751, and 'Peking' using double row planting method by Ross and Brim (1957) and more than 158 soybean accessions have been reported as resistant to SCN in National Plant Germplasm System (Rincker et al., 2017). However, the sources of SCN resistance generally have poor agronomic performance and low yield and the major resistance QTL, *Rhg4* closely linked with the black seed coat gene, so it became difficult to introduce the trait into elite germplasms (Matson and Williams, 1965). The first SCN resistant cultivars were 'Pickett', and 'Dyer', which were released in 1966 and 1967, respectively. 'Pickett' with yellow seed originated from a selected F_4 progeny from NC55-1 x (D49-24916 × 'Dorman'). The NC55-1, the resistant parent, was derived from the backcross of 'Lee' and 'Peking' by Brim and Ross (1966). 'Dyer' with the light green seed was developed from a cross of 'Hill' × ['Lee' × ('Lee' × 'Peking')] (Epps and Hartwig, 1967).

In 1972, PI 88788 was reported with a high level of resistance and has been used extensively to develop commercial resistant cultivars for nearly 30 years because it was introgressed successfully into high yielding varieties and carrying one major gene, *Rhg1* for resistance. In the report published by Iowa State Extension publication in 2020 (Tylka, 2020), 810 of 849 varieties were derived from PI 88788. In general, SCN-resistant soybean varieties produced higher yields than the susceptible varieties.

SCN resistance genomic regions reported

Rhg1 locus

The most important QTL for SCN resistance, known as *Rhg1* (resistance to *H. glycine*) located on Chromosome (Chr) 18 (Linkage Group G) (Table 1.3, Fig. 1.3). This locus has been designated by the Soybean Genetics Committee as *cqSCN-001*. It has been mapped in many SCN resistant sources including all seven indicator lines in the HG type test. Two unique alleles were reported at *Rhg1* locus: *rhg1-a* and *rhg1-b* which carry copy number variation and allelic difference (Cook et al., 2012b; Shi et al., 2015; Liu et al., 2017). Seven indicator lines which are used in HG type designation carry *Rhg1* locus with two different resistance alleles. Of them, 'Peking', PI 90763, PI 89772, and PI 437654 carry *rhg1- a* resistance allele while PI 88788, PI 209332, and PI 548316 carry *rhg1-b* resistance allele.

These two types of resistance led to the difference in cellular and timing response during SCN infection (Mitchum, 2016). J2 SCN established a feeding site, called a syncytium in both SCN-resistant and susceptible plants; but for the genotypes carrying the *rhg1-a* allele derived from the 'Peking' source, necrosis of syncytium and neighboring cells happened earlier [by 5 days after inoculation (DAI)] than those carrying the *rhg1-b* allele derived from PI 88788 source (by 8-10 DAI) (Bayless et al., 2016; Mitchum, 2016). Besides, the necrosis began inside the syncytium in plants having the *rhg1-a* resistant allele while it happened outside of the cell wall of the syncytium in the plants carrying the *rhg1-b* resistance allele (Kim et al., 2012; Mitchum, 2016).

Genes for the *rhg1-b* allele at the *Rhg1* locus were cloned, and it was found that three of four distinct genes contributed to SCN resistance: *Glyma18g02580, Glyma18g02590,* and *Glyma18g02610* (Cook et al., 2012b). 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*) (Cook et al., 2012b). Silencing any of these three genes reduced SCN resistance in the 'Fayette' soybean line which carries a *rhg1-b* allele. While overexpression of all three genes together improved resistance (Cook et al., 2012b). *GmSNAP1*8 had amino acid polymorphism at the highly conserved C-terminus (Bayless et al., 2016). Bayless (2016) found *GmSNAP18* in resistant soybean plants could not bind N-ethylmaleimide sensitive factor (NSF), which affected the disassembly of SNARE bundles and inhibited vesicular trafficking within the syncytia. Based on the expression patterns and gene sequence, *GmAAT* was predicted to encode an amino acid transporter which was mostly expressed in vasculature. Overexpression of this gene might enhance the accumulation of glutamate in roots by affecting the transportation of glutamate from shoots to roots (Guo et al., 2019). The accumulation of glutamate might relate to the Jasmonic acid (JA) pathway which could activate the pathogen defense response (Guo et al., 2019). The *GmWI12* was found to have a higher expression level in 'Peking', the resistant cultivar than in Essex, the susceptible cultivar (Dong and Hudson, 2021). Knocking out the *GmWI12* in 'Peking' by CRISPR-Cas9, the SCN susceptibility was significantly increased because of decreasing SA levels. *GmWI12* was found in the protein- protein interaction with DELLA on Chr 18 which regulated both JA and SA (Salicylic acid) levels that are associated with plant growth and defense (Dong and Hudson, 2021).

Based on the number of repeats of these genes, cultivars can be classified into three types: single copy (*rhg1-WT*) which is susceptible; low copy (*rhg1-LC*) with 2 to 4 repeats which is *rhg1-a* resistance type; and high copy (*rhg1-HC*) with more than four copies which is *rhg1-b* resistance type (Cook et al., 2014a). For example, PI 88788, PI 209332, and PI 548316 ('Cloud')

carry 9, 10 and 7 copies, respectively, which are classified as *rhg1-b* or *rhg1-HC*. 'Peking', PI 90763, PI 89772, and PI 437654 carry three copies which are classified as *rhg1-a* or *rhg1-LC*. Susceptible cultivars such as 'Williams 82' carry only one single copy at *Rhg1* locus. In *rhg1-b* (*rhg1-HC*) lines also carry at least one copy similar to susceptible single copy while *rhg1-a* (*rhg1-LC*) lines did not carry any copy that was similar to susceptible lines (Cook et al., 2012; Cook et al., 2014; Lee et al., 2015). Based on the gene sequences, functional markers were designed which were named as GSM381 and GSM383 (Shi et al., 2015). GSM383 is able to distinguish between *rhg1-a* and *rhg1-b* resistance alleles while GSM381 can be used to differentiate *rhg1-a/b* resistance alleles from the wild-type allele (*Rhg1)* (Shi et al., 2015).

Rhg2 locus

A 3 cM from two RFLP markers: php02275a and php02301a on Chr 7, considered as *Rhg2* locus was first mapped in PI 437654 for SCN race 3 resistance in 1995 (Table 1.3) (Webb et al., 1995). Although this QTL has not been confirmed in other mapping studies, it has been reported in multiple Genome-wide Association Studies (GWAS) (Vuong et al., 2015; Tran et al., 2019). Recent publications indicated a SNP marker, ss715597431 on Chr 7 (36.4 Mb based on Wm82 a2 version) was significantly associated with HG type 0 resistance (Bayless et al., 2018; Tran et al., 2019). This SNP was present at *Glyma.07g196900* [*N-ethylmaleimide sensitive factor (NSF)* gene]. Both α-SNAP and NSF are housekeeping proteins and the interaction of these genes might intervene with normal vesicle trafficking, leading to cell death (Bayless et al., 2016). Based on a vitro binding experiment, the *NSF* gene on Chr 7 bond with the C-terminal polymorphism of the *Rhg1* resistance type *α- SNAP* gene on Chr 18. This SNP might be co-presented with *Rhg1* locus on Chr 18. Eight hundred and fifty-five soybean accessions from USDA germplasm which carried *α-SNAP* gene at *Rhg1* locus was also homozygous for this SNP

based on haplotype analysis using SoySNP50K iSelect BeadChip (Bayless et al., 2018). The study also examined F_5-derived RILs developed from the crosses of the IA3023 (susceptible parent) and eight different resistant parents carrying either *rhg1-a* or *rhg1-b.* The co-inheritance was also observed among the RILs that carried a resistance allele at *Rhg1* locus.

Rhg3 locus

The third QTL, *cqSCN-005*, also *Rhg3* was initially mapped on Chr 17 (LG D2) from 'Hartwig', a SCN-resistant cultivar derived from PI 437654 using 92 $F_{5:11}$ RILs and 144 SSR markers for SCN race 14 (Table 1.3) (Schuster et al., 2001). The QTL was confirmed using a near isogenic line (NIL) population derived from the resistant RIL parent (Kazi et al., 2010). It was located between SSR markers Satt488 and Satt514 and explained 11% of phenotypic variation for HG type 1.2.5 (SCN race 2) (Kazi et al., 2010).

Rhg4 locus

The *Rhg4* locus, designated as *cqSCN-002* was mapped on Chr 8 (LG A) and found in 'Peking' type resistance including PI 89772, PI 90763, and PI 437654 (Table 1.3). This locus is located at 0.35 cM away from the I locus that controls black seed color (Matson and Williams, 1965). This locus was fine mapped within 8 kb comprising two genes (Liu et al., 2012), but only one gene, *Serine hydroxymethyltransferase* (*GmSHMT08*) was found to contribute to SCN resistance (Liu et al., 2012). Silencing this gene in the resistant $F_{2:6}$ RIL ExF67 derived from a cross between 'Forrest' and 'Essex' increased SCN susceptibility and transferring *GmSHMT08* gene into the SCN susceptible $F_{2:6}$ RIL ExF63 made it be resistant to SCN (Liu et al., 2012). Mutation of *GmSHMT08* gene affects one carbon folate metabolism which may trigger hypersensitive response (Liu et al., 2012). Sequencing this gene revealed two polymorphisms located in the first and second exons which caused missense mutation (Liu et al., 2012) . Based

on the finding, a functional marker, designated as GSM191 was developed for marker-assisted selection (Shi et al., 2015). Previous studies reported the interaction between *rhg1-a* and *Rhg4* alleles in both 'Peking' (Meksem et al., 2001) and PI 43654 (Wu et al., 2009) background. Without *Rhg4*, *rhg1-a* alone did not provide full resistance while *rhg1-b* was still resistant to SCN. By evaluating populations segregating for *Rhg1* copy number, type and *Rhg4* allele type for SCN resistance, Yu et al. (2016) reported that *Rhg4* only interacted with *rhg1-a* to increase resistance. Recently, whole genome re-sequencing of 106 soybean lines revealed copy number variation and allelic difference at *Rhg4* locus (Patil et al., 2019). Based on whole genome resequencing, there are three haplotype groups at the *Rhg4* locus: *Rhg4-a* 'Peking' type, *Rhg4-b* susceptible type, *Rhg4-c* PI 437654 type related to amino acid change. Copy number variation was also reported at *Rhg4* locus: 'Peking' was found to have 2.3 copies and PI 437654 was found to have 4.3 copies (Patil et al., 2019). Besides, polymorphism was found in the promoter region of *Rhg4* locus which regulated higher gene expression that might affect SCN resistance (Patil et al., 2019).

Rhg5 locus

The *Rhg5* locus, also *cqSCN003* was mapped on Chr 16 (LG J) from PI 88788 between SSR markers Satt547 and Satt432 for HG type 0 and HG type 1.3.5.6.7 (SCN race 14) resistance with minor effect (explaining less than 10% of phenotypic variation) (Table 1.3) (Glover et al., 2004). Based on the physical position of 'William82' version 2, this QTL was located from 34.0 to 36.2 Mb. This QTL was also mapped in PI 90763 for HG type 0 resistance using RFLP markers and F_2 population. The QTL had a smaller effect than *Rhg1* locus on Chr 18 in PI 90763 in this study with a percentage of variation of 18.8 and 44.8% respectively (Concibido et al.,

1997). Similarly, using RFLP markers, this locus, located near to B032V-1 marker, was mapped in PI 290332 for HG type 0 resistance (Concibido et al., 1996).

Other loci for SCN resistance

More than 200 QTLs have been mapped for SCN resistance, but most of them have not been confirmed. There are six QTLs which were confirmed and designated as *cqSCN* by the Soybean Genetic Committee (Table 1.3). Of the six, two are major QTLs: *Rhg1* and *Rhg4* loci were designated as *cqSCN-001* and *cqSCN-002* for multiple SCN populations. *cqSCN003* was designated for *Rhg5* locus on Chr 16. In 2001, two QTLs, named as *cqSCN-006* and *cqSCN007* were mapped on Chr 15 (LG E) and Chr 18, (LG G), respectively, from a wild soybean accession PI 468916 for HG type 0 resistance (Wang et al., 2001). In 2017, these QTLs were fine mapped for HG type 2.5.7 resistance using an RIL population derived from a cross between cultivar LD00-2817P derived from PI 437654 and the germplasm line LDX01-1-65 which was derived from *Glycine soja* PI 468916 (Yu and Diers, 2017). The *cqSCN-006* is located in the 212.1 kb interval on Chr 15, flanking between two SNPs: ss715621232 and ss715621239 and explaining 23% of phenotypic variation for HG type 2.5.7 resistance. The *cqSCN-007* resided in a 103.2 kb interval on Chr 18, flanking between SSR marker BARC18_1669 and SNP marker ss715631888, which was located 61 cM away from the *Rhg1* locus. It explained 27% of phenotypic variation and (Yu and Diers, 2017).

Minor QTLs on Chr 11 were mapped in multiple studies. The minor QTL on Chr 11 was mapped from PI 89772 in an interval of RFLP marker A006 to SSR marker Satt583 (Wm82.a2: 10.8 to 25.5 Mb) for SCN race 1 and 2 resistance, accounting for 6.6 and 6.8% of phenotypic variation, respectively (Yue et al., 2001b). Overlapping with this QTL, the genomic region flanking between RFLP markers: A118 to A006 (Wm82.a2: 10.8 to 24.5 Mb) was mapped in PI

89772 for race 5 resistance, explained 9.5% of phenotypic variation (Yue et al., 2001b). Another QTL was also mapped in PI 90763 for race 2 and 5 resistance, explaining 6.7 and 11.2 % of phenotypic variation, respectively. The QTL was located within Satt453 to Satt359 (32.4 to 34.2 Mb). Recently, this QTL was fine mapped between SSR markers Sat_123 and BARCSOYSSR11_1420 (32.1 to 33.0 Mb), in the interval of 821 kb from PI 84751 for race 1 resistance (Suzuki et al., 2020). The study used a RIL population derived from a cross between susceptible line Tokei 758 and resistant line To-8E, derived from a cross between PI 84751 and Gedenshirazu. The resistant parent, To-8E is resistant to both SCN race 1 and race 3 (Suzuki et al., 2020). Interestingly, a *SNAP* gene, *Glyma.11G234500 (GmSNAP11)* (32.9 Mb) within this QTL has been considered as a candidate gene for SCN resistance because it is closely related and shared similar structure with a *SNAP* gene at the *Rhg1* locus (Tian et al., 2019). The level of similarity between two gene sequences was 80.3% (Tian et al., 2019); and the level of identity of amino acid sequence was 92.4% (Lakhssassi et al., 2017). Based on an expression study using qRT-PCR, both *GmSNAP18* and *GmSNAP11* were upregulated (Lakhssassi et al., 2017). Based on the phenotyping score of a RIL population from a cross between 'Essex' and 'Forrest' ('Peking' type resistance), the average female index was 4% if lines carried both genes (Lakhssassi et al., 2017). With only *GmSNAP11*, the female index was 20.3% but if the lines carried *rhg1*, *Rhg4*, and *GmSNAP11* resistance alleles, the female index dropped to 1.9% (Tian et al., 2019).

PI 567516 C, originated from China was reported as a novel source of resistance to multiple SCN populations (race 1, 2, 3 and LY1- a synthetic SCN population) because it did not carry *rhg1* and *Rhg4* resistance allele (Arelli et al., 2010; Vuong et al., 2010). *Rhg1* locus in PI 567516 C was detected for reniform nematode resistance (Jiao et al., 2015a). Based on whole

genome sequencing, this PI carried the *Rhg1-a* allele, same to resistance type from 'Peking' (Jiao et al., 2015a), which was confirmed by genotyping with KASP SNP marker (GSM381 and GSM383) (Tran et al., 2019). Based on KASP SNP marker genotyping, PI 567516 C did not carry the *Rhg4* resistance allele (Tran et al., 2019). For SCN resistance, two QTLs on Chr 10 (LG O) and 18 (LG G) were mapped using a $F_{2:3}$ population derived from Magellan × PI 567516 C (Vuong et al., 2010). The QTL on Chr 18 located between two SSR markers: Satt612 and Satt191, 80 cM away from *Rhg1* locus (Vuong et al., 2010) but partially overlapped with *cqSCN-007* from a wild soybean accession, PI 468916 (Yu and Diers, 2017). This QTL was fine mapped to a 166 kb, containing seven candidate genes for SCN resistance (Usovsky et al., 2021b). The major QTL, *qSCN10* on Chr 10 in PI 567516 C is associated with resistance to five SCN populations, explaining 7.9 to 21.7% of phenotypic variation (Vuong et al., 2010). This QTL resided 15.3 cM between two SSR markers: Satt592 and Sat_038 (Vuong et al., 2010). Recently, this QTL was fine mapped to a 379 kb (42.4 to 42.8 Mb) region on Chr 10and 51 genes were found within the 379 kb interval (Zhou et al., 2021b).

Besides QTL mapping, GWAS has also been used for identification of QTL for SCN resistance. Two major QTLs, *Rhg1* and *Rhg4* loci have been repeatedly reported in GWAS studies (Han et al., 2015; Vuong et al., 2015; Tran et al., 2019). *qSCN-10*, the novel locus from PI 567516 C was confirmed in GWAS study (Vuong et al., 2015), and *Rhg2* locus on Chr 7, mapped in PI 437654 was also identified in a GWAS study for HG type 0 resistance (Tran et al., 2019). There is no QTL found on Chr 2 based on QTL mapping, but from GWAS studies, a QTL for SCN resistance on Chr 2 has been identified (Han et al., 2015; Vuong et al., 2015).

SCN resistance breakdown due to limited resistance sources

There has been limited diversity of SCN resistance source in commercial cultivars with more than 95% of U.S soybean cultivars found to carry the *rhg1-b* allele from PI 88788 (Mitchum, 2016). In Iowa, of 849 reported varieties, only 35 varieties were derived from 'Peking'-type resistant sources (Tylka et al., 2020). Due to exploiting a single source of resistance from PI 88788, SCN populations are gradually able to overcome its resistance. McCarville et al. (2017) reported virulence of SCN increased 1.6% per year on PI 88788 derived resistant varieties since 2001 in the field in Iowa. Similarly, more than 70% of collected SCN populations overcame PI 88788 derived resistance in Illinois, Missouri, and Wisconsin (Mitchum et al., 2007; Niblack et al., 2008; MacGuidwin, 2012). 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. At the University of Georgia Soybean Breeding Program, we have screened more than 700 soybean accessions and breeding lines originating from various origins against HG type 0 (SCN race 3), a predominant population in Georgia. All these soybean accessions were also genotyped using three functional markers that were developed at *Rhg1* and *Rhg4* loci (Shi et al., 2015). These markers were used to identify resistant sources that did not have the same resistance alleles as 'Peking' and PI 88788. Combining phenotyping and genotyping results, 58 resistant and moderately resistant lines were identified that did not possess the same resistance alleles as 'Peking' and PI 88788 at *Rhg1* and *Rhg4* loci. Based on haplotype analysis at these two loci assembled with SoySNP50K SNP Infinium Chip data, these 58 lines were also grouped separately from PI 88788 and 'Peking' (Tran et al., 2019). The highly resistant line was selected to develop a mapping population to identify a novel

genomic region for SCN resistance which enhances SCN resistance through pyramiding genes and provides an additional resistant source to slow down SCN adaptation.

Plant microbiome introduction

Impact of plant microbiome

It is well known that plants and microbes possess a mutual relationship. In this symbiotic association, bacteria play a fundamental role in plant nutrition acquisition by breaking down mineral substances from soil into absorbable forms for plants. The nitrogen-fixation bacteria are a well-known example of beneficial bacteria that convert nitrogen to useful ammonia (Igiehon and Babalola, 2018). The bacteria: *Bradyrhizobium, Rhizobium* and *Achromobacter* were reported as nitrogen fixators in cowpea (Azarias Guimarães et al., 2012). Recent studies indicated that bacteria are able to synthesize phytohormones such as: auxin, idole 3-acetic acid (IAA), and gibberellins that support plant growth and development (Egamberdieva et al., 2017). For example, IAA-producing *Mycobacteria* and *Sphingomonas* genera were found in the rhizosphere of the tropical orchid species *Dendrobium moschatum* (Tsavkelova et al., 2005). With these bacteria cultures, the orchid seeds could germinate, and seedlings developed with height of root increased 9.4 and 13.4 times than plants without inoculation (Tsavkelova et al., 2005). In other reports, genera *Arthrobacter, Pseudomonas, Rhizobium* and *Bacillus* that were found in the rhizosphere of soybeans could produce IAA (Naz et al., 2008). Also, some IAA-producing bacteria species belonging to the genera *Bacillus* and *Enterobacter* that could improve maize growth in Cadmium-contaminated soil were reported (Ahmad et al., 2016). Beyond that, the plant microbiome is capable of preventing pathogen's colonization by competing nutrients and producing antibiotic compounds, which boost the plant immune response upon pathogen attack (Berendsen et al., 2018). The bacterium *Bacillus subtilis* produces non-ribosomal

antibiotics that play a role in competition with other microorganisms during spore germination (Patel and N.Amaresan, 2014). Various bacteria species including *Pseudomonas trivialis*, *P. fluorescens*, *Stenotrophomnas maltophilia* have the ability to produce volatile organic compounds that inhibit mycelial growth of fungal plant pathogens in vi-tro condition (Schulz-Bohm et al., 2017). Pyrolnitrin produced by the *P. fluorescens* BL915 strain had a potential positive effect on *Rhizoctonia solani* in cotton (Qiao et al., 2017). Phenazine, produced by *Pseudomonads* is able to prevent *F. oxysporum* (Chin-A-Woeng et al., 2003).

Bacteria has been studied as a trigger to plant immune resistance with two forms: induced systematic resistance (ISR) and salicylic acid-dependent (SAR). SAR leads to induction of a hypersensitive response that is associated with programmed cell death in the infected area, while induced systematic resistance requires jasmonate and ethylene instead of salicylic acid (Beneduzi et al., 2012). Several specific *Pseudomonas* strains have been reported to initiate ISR in various plants including cucumber, radish, tobacco, and *Arabidopsis* (Beneduzi et al., 2012). In addition to deploying resistant plants, crop rotation and utilization of beneficial bacteria are able to control diseases in crop production. In terms of plant breeding, besides exploiting plant genetic traits to improve plant growth and resistance to biotic and abiotic stresses, plant breeders can evaluate plant cultivars for their broad interaction with beneficial bacteria.

Potential application of plant microbiome

The biological management which exploits a single strain begins with a screening of a collection of microbial inoculums of plant growth promoting beneficial bacteria for disease prevention. Using this approach, it showed success in the lab and greenhouse conditions but failed in the field. For example, *Azospirillum brasilense* increased growth of maize and wheat under controlled environment but had no significant effect on their growth in the field (Hungria

et al., 2010). Similarly, *Rhizobium legyminosarum bv. Trifoli*i inoculation increased biomass of rice plants in the greenhouse but not in the field (Kecskés et al., 2016). The reason might be that the tested microorganism was not able to compete with microorganisms in field conditions; and the variation of environmental conditions could affect microbial activity. The plant microbiome can be selected together with the plant genome to develop a new plant breeding approach. Understanding what makes a plant a suitable host to select essential microbiota is potential in improving crops because in nature, plants exit by interacting with diverse microbes (Corbin et al., 2020). Therefore, identifying the genes that control plants to assemble microbiota in the root is considered as a new breeding approach to improve both quality and quantity.

Plant genotypes recruit different microbiome members conferring resistance to abiotic and biotic stresses. GWAS has been used to characterize the effects of plant genotypes on associated microbiomes. Two important aspects of GWAS are genotypes and phenotypes. Phenotypes can be used to group genes (ex: 16sRNA) into operational taxonomic units (OTUs) or amplicon sequence variants (ASVs). OTUs were generated by clustering sequence reads based on similarity thresholds (Westcott and Schloss, 2015) while ASVs were created by modeling and correcting sequencing errors (Callahan et al., 2016). Wallace et al. (2018) conducted a GWAS using 300 diverse maize genotypes to map genomic regions associated with leaf microbiome. The study found two predominant taxa: *Sphingonomads* and *Methylobacteria* in maize leaf. Based on GWAS analysis, most genomic regions affecting the leaf microbial community might have a small effect with a lower LOD score than 7 (Wallace et al., 2018). Horton et al. (2014) conducted a GWAS of 196 *Arabidopsis thaliana* genotypes to map genomic regions for leaf microbiome by sequencing taxonomic marker genes for both bacteria and fungi. In the analysis of individual OTU of 100 most represented OTUs, they identified five genomic regions on Chr 1

(2 regions) , 3 (2 regions) and 4 (1 region) which were associated with 34 different OTUs. Bergelson et al. (2019) showed candidate genes affecting the composition of *Arabidopsis* root microbiome through GWAS were related to cell wall integrity and immunity. A GWAS on 3,024 rice accessions were performed using 6.5 million SNPs. Overall, they identified 22 significant SNPs on six chromosomes associated with 12 bacterial genus abundance (Roman-Reyna et al., 2020). Wild rice also carries the beneficial allele at the OsCERK1 locus that enhances arbuscular mycorrhiza fungi (AMF) in roots which are beneficial fungi that promote the growth of host plants (Huang et al., 2020). Selection and breeding of plants for plant association with beneficial microbiota is a promising approach.

The factors affect plant microbiome structure

Bacteria colonize every part of a plant including above-ground parts, leaves, stems, and flowers (phyllosphere); interior parts, cells of leaves, stems, and roots (endosphere); root surface (rhizoplane) and rhizospheric region that is defined as the narrow region of adjacent soil that interact with the plant roots (Turner et al., 2013). Of those, the rhizosphere has been considered a microbial hot spot, so most studies have focused on this region. Research in *Arabidopsis* (Lundberg et al., 2012) and rice (Edwards et al., 2015) indicated that rhizosphere microbial communities were more diverse than other root compartments. It is commonly believed that rhizobacteria are influenced by root exudates. A study showed that plants could secrete up to 40% of their photosynthates into the rhizosphere (Bais et al., 2006). Exudates released from plant roots that may be specific to plant species or genotypes. The influence of plant genotypes on rhizobacteria was reported in *Arabidopsis thaliana* (Micallef et al., 2009).

In a study in potato, plant genotype effects were reported in rhizosphere communities. Differences in abundance on three potato cultivars were detected mainly for species belonging to

Pseudomonales, Streptomycetaceae, and *Micromonosporaceae,* which might be related to plant pathogen management (Weinert et al., 2011). A similar trend of plant genotypes was also reported in wheat research (Meyer et al., 2010). Along with plant genotypes, other factors including sampling sites, cultivation method, and plant developmental stages have been reported as main drivers to supplement the microbiome effects. For example, soil types and cultivation methods were reported in the rice microbiome (Edwards et al., 2015); and the wine grape microbiome (Bokulich et al., 2014). Indeed, while *Pseudomonas* is abundant under humid conditions (Mendes et al., 2011), *Bacillus* is abundant under arid conditions (Köberl et al., 2011).

For *Arabidopsis* rhizosphere microbial communities, early and late plant development stage influenced the diversity of bacteria including *Actinobacteria, Bacteroidetes* and *Cyanobacteria*, suggesting there was a selection in bacteria during plant development (Chaparro et al., 2014). Similarly, an assessment of the soybean rhizosphere microbial communities was influenced by plant development: beginning seed and full maturity stage clustered away from other stages in field condition based on principal component analysis and denaturing gradient gel electrophoresis (DGGE) (Xu et al., 2009). Early reproductive growth stages showed more complicated communities than later stages. One reason that explained the difference might be due to the difference in quantity and quality of root exudates. One hypothesis was that roots released sugar at the early stage but released specific substrates such as antimicrobial compounds that might help the plant to select for beneficial microbes for pathogen suppression (Chaparro et al., 2014). However, some bacteria belonging to phyla *Chloroflexi, Firmicutes, Planctomycetes* and *Verrucomicrobia* did not demonstrate detectable changes through different stages of growth (Chaparro et al., 2014).

Soybean microbiome

Soybean microbiome and its function could be applied in soybean production by utilizing beneficial microbes in biofertilizer and selecting and developing new varieties through plant breeding that enhance a symbiotic relationship with benefit microbes. Thus, one must first understand the diversity of microbial communities in soybean and factors which affect its diversity. In 2009, rhizosphere bacterial communities of three soybean genotypes were evaluated in field and greenhouse conditions by sequencing denaturing gradient gel electrophoresis bands (Xu et al., 2009). The result demonstrated that different soil types and reproductive stages of soybean differed in the profile of bacterial communities in the rhizosphere at both conditions. Although they could not detect a significant difference in bacterial communities among three soybean genotypes, they found culturable bacterial communities showed significant differences between the two soybean genotypes.

Rascovan et al. (2016) applied 16s rRNA pyrosequencing to study root-associated microbiomes of soybean 'Williams 82' in eight field conditions and found that soil properties with different soil pH from eight fields significantly affected the richness of soybean microbiome with $P \leq 0.05$ using a Manel test. Besides, based on Pearson correlation, there was a positive correlation with pH and class *Bacilli* within root samples (Rascovan et al., 2016). *Proteobacteria* was the most abundant phylum in soybean root (Rascovan et al., 2016).

Soybean production is challenged by diseases, insects and abiotic factors including unfavorable weather. SCN was ranked as the most destructive pest affecting soybean yield in the U.S. It even has been postulated that plants actively recruit beneficial bacteria in soil to beat pathogen attack. Various bacteria strains that can be used as biocontrol agents including *Trichoderma* isolates for inhibiting sudden death syndrome pathogen growth and reducing root

rot (Pimentel et al., 2020); *Rhizobium japonicum* coated seeds could improve seed germination and reduce root rot diseases caused by *Fussarium solani* and *Macrophimina phaseolina* (Al-ani et al., 2018). Hence, we hypothesized that soybean roots might recruit beneficial bacteria from the soil to suppress the SCN. The knowledge about bacterial microbiota related to SCN is still limited. Therefore, revealing the structure and assembly of the bacterial communities in soil, rhizosphere and soybean root related to soybean cyst nematode could provide a better understanding of the promising beneficial bacteria and factors affecting bacterial assembly.

Objectives:

1) Map, fine-map, and characterize the QTL using bi-parental populations derived from resistant sources identified in previous study
2) Determine composition of microbiome on the roots of soybean genotypes with resistance and susceptibility to SCN; and investigate the impact of root microbiome on host plant resistance to SCN

References

Ahmad, I., Akhtar, M.J., Asghar, H.N., Ghafoor, U., and Shahid, M. (2016). Differential effects of plant growth-promoting rhizobacteria on maize growth and cadmium uptake. *J. Plant Growth Regul.* 35(2)**,** 303-315. doi: 10.1007/s00344-015-9534-5.

Al-ani, R.A., Adhab, M.A., Mahdi, M.H., and Abood, H. (2018). *Rhizobium japonicum* as a biocontrol agent of soybean root rot disease caused by *Fusarium solani* and *Macrophomina phaseolina*. *Plant Prot. Sci.* 48**,** 149-155.

Ali, N. (2010). "Soybean processing and utilization," in *The soybean: botany, production and uses,* ed. G. Singh. (Oxfordshire, UK: CAB International), 345-375.

Arelli, P.R., Concibido, V.C., and Young, L.D. (2010). QTLs associated with resistance in soybean PI567516 C to synthetic nematode population infecting cv. 'Hartwig'. *J Crop Sci Biotechnol* 13(3)**,** 163-167. doi: 10.1007/s12892-010-0060-z.

Azarias Guimarães, A., Duque Jaramillo, P.M., Simão Abrahão Nóbrega, R., Florentino, L.A., Barroso Silva, K., and de Souza Moreira, F.M. (2012). Genetic and symbiotic diversity of nitrogen-fixing bacteria isolated from agricultural soils in the Western Amazon by using cowpea as the trap plant. *Appl. Environ* 78(18)**,** 6726-6733. doi: 10.1128/aem.01303-12.

Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S., and Vivanco, J.M. (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* 57**,** 233-266. doi: 10.1146/annurev.arplant.57.032905.105159.

Bayless, A.M., Smith, J.M., Song, J., McMinn, P.H., Teillet, A., August, B.K., et al. (2016). Disease resistance through impairment of alpha-SNAP-NSF interaction and vesicular trafficking by soybean *Rhg1*. *Proc Natl Acad Sci.* 113(47)**,** E7375-E7382. doi: 10.1073/pnas.1610150113.

Bayless, A.M., Zapotocny, R.W., Grunwald, D.J., Amundson, K.K., Diers, B.W., and Bent, A.F. (2018). An atypical N-ethylmaleimide sensitive factor enables the viability of nematode-resistant *Rhg1* soybeans. *Proc Natl Acad Sci.* 115(19)**,** E4512-E4521. doi: 10.1073/pnas.1717070115.

Beneduzi, A., Ambrosini, A., and Passaglia, L.M.P. (2012). Plant growth-promoting rhizobacteria (PGPR): Their potential as antagonists and biocontrol agents. *Genet. Mol. Biol.* 35(4 (suppl))**,** 1044-1051. doi: 10.1590/s1415-47572012000600020.

Berendsen, R.L., Vismans, G., Yu, K., Song, Y., de Jonge, R., Burgman, W.P., et al. (2018). Disease-induced assemblage of a plant-beneficial bacterial consortium. *The ISME J.*12(6)**,** 1496-1507. doi: 10.1038/s41396-018-0093-1.

Bergelson, J., Mittelstrass, J., and Horton, M.W. (2019). Characterizing both bacteria and fungi improves understanding of the *Arabidopsis* root microbiome. *Sci. Rep.* 9(1)**,** 24. doi: 10.1038/s41598-018-37208-z.

Bernard, R.L. (1972). Two genes affecting stem termination in soybeans1. *Crop Sci.* 12(2)**,** doi: https://doi.org/10.2135/cropsci1972.0011183X001200020028x**.**

Bokulich, N.A., Thorngate, J.H., Richardson, P.M., and Mills, D.A. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proc. Natl. Acad. Sci.* 111(1)**,** E139-E148. doi: 10.1073/pnas.1317377110.

Brim, C.A., and Ross, J.P. (1966). Registration of 'Pickett' soybeans. *Crop Sci.* 6**,** 305.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13(7)**,** 581-583. doi: 10.1038/nmeth.3869.

Carlson, J.B., and Lersten, N.R. (2004). "Reproductive Morphology," in *Soybeans: Improvement, Production, and Uses*., 59-95.

Carter Jr., T.E., Nelson, R.L., Sneller, C.H., and Cui, Z. (2004). "Genetic diversity in soybean," in *Soybeans: Improvement, Production, and Uses*., 303-416.

Chang, S., Thurber, C.S., Brown, P.J., Hartman, G.L., Lambert, K.N., and Domier, L.L. (2014). Comparative mapping of the wild perennial *Glycine latifoli*a and soybean (*G. max*) reveals extensive chromosome rearrangements in the genus *Glycine*. *PLOS ONE* 9(6)**,** e99427. doi: 10.1371/journal.pone.0099427.

Chaparro, J.M., Badri, D.V., and Vivanco, J.M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *The ISME J.* 8(4)**,** 790-803. doi: 10.1038/ismej.2013.196.

Chin-A-Woeng, T.F.C., Bloemberg, G.V., and Lugtenberg, B.J.J. (2003). Phenazines and their role in biocontrol by *Pseudomonas* bacteria. *New Phytol.* 157(3)**,** 503-523. doi: 10.1046/j.1469-8137.2003.00686.x**.**

Colgrove, A.L., and Niblack, T.L. (2005). The effect of resistant soybean on male and female development and adult sex ratios of *Heterodera glycines*. *J. Nematol.* 37(2), 161-167.

Concibido, V.C., Denny, R.L., Lange, D.A., Orf, J.H., and Young, N.D. (1996). RFLP mapping and marker-assisted selection of soybean cyst nematode resistance in PI 209332. *Crop Sci.* 36(6). doi: cropsci1996.0011183X003600060038x**.**

Concibido, V.C., Lange, D.A., Denny, R.L., Orf, J.H., and Young, N.D. (1997). Genome mapping of soybean cyst nematode resistance genes in 'Peking', PI 90763, and PI 88788 using DNA markers. *Crop Sci.* 37(1). doi: cropsci1997.0011183X003700010046x**.**

Cook, D.E., Bayless, A.M., Wang, K., Guo, X., Song, Q., Jiang, J., et al. (2014a). Distinct copy number, coding sequence, and locus methylation patterns underlie *Rhg1*-mediated

soybean resistance to soybean cyst nematode. *Plant Physiol* 165(2)**,** 630-647. doi: 10.1104/pp.114.235952.

Cook, D.E., Lee, T.G., Guo, X., Melito, S., Wang, K., Bayless, A.M., et al. (2012). Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338(6111)**,** 1206-1209. doi: 10.1126/science.1228746.

Corbin, K.R., Bolt, B., and Rodríguez López, C.M. (2020). Breeding for beneficial microbial communities using epigenomics. *Front. Microbiol.* 11(937). doi: 10.3389/fmicb.2020.00937.

Davis, E.L., and Tylka, G.L. (2000). Soybean cyst nematode disease. *Plant Health Instr.* doi: 10.1094/phi-i-2000-0725-01.

Dong, J., and Hudson, M.E. (2021). WI12*Rhg1* interacts with DELLAs and mediates soybean cyst nematode resistance through hormone pathways. *Plant Biotechnol. J.* . doi: https://doi.org/10.1111/pbi.13709**.**

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci.* 112(8)**,** E911-E920. doi: 10.1073/pnas.1414592112.

Egamberdieva, D., Wirth, S.J., Alqarawi, A.A., Abd_Allah, E.F., and Hashem, A. (2017). Phytohormones and beneficial microbes: Essential components for plants to balance stress and fitness. *Front. Microbiol.* 8(2104). doi: 10.3389/fmicb.2017.02104.

Epps, J.M., and Hartwig, E.E. (1967). 'Dyer', a new nematode-resistant soybean variety. *University of Tennessee Agricultural Experiment Station*. doi: https://trace.tennessee.edu/utk_agbulletin/170/.

Glover, K.D., Wang, D., Arelli, P.R., Carlson, S.R., Cianzio, S.R., and Diers, B.W. (2004). Near isogenic lines confirm a soybean cyst nematode resistance gene from PI 88788 on linkage group J. *Crop Sci.* 44(3)**,** 936-941. doi: https://doi.org/10.2135/cropsci2004.9360**.**

Golden, A., and Et, A. (1970). Terminology and identity of infraspecific forms of the soybean cyst nematode (*Heterodera glycine*s). *Plant Dis.* 54**,** 544-546.

Guo, W., Zhang, F., Bao, A., You, Q., Li, Z., Chen, J., et al. (2019). The soybean *Rhg1* amino acid transporter gene alters glutamate homeostasis and jasmonic acid-induced resistance to soybean cyst nematode. *Mol Plant Pathol* 20(2)**,** 270-286. doi: 10.1111/mpp.12753.

Han, Y., Zhao, X., Cao, G., Wang, Y., Li, Y., Liu, D., 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 Genet.* 16**,** 598. doi: 10.1186/s12864-015-1800-1.

Hart, C. (2017). "The Economic evolution of the soybean industry," in *The Soybean Genome,* eds. H.T. Nguyen & M.K. Bhattacharyya. (Cham: Springer International Publishing), 1-9.

Horton, M.W., Bodenhausen, N., Beilsmith, K., Meng, D., Muegge, B.D., Subramanian, S., et al. (2014). Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nat. Commun.* 5(1)**,** 5320. doi: 10.1038/ncomms6320.

Huang, R., Li, Z., Mao, C., Zhang, H., Sun, Z., Li, H., et al. (2020). Natural variation at OsCERK1 regulates arbuscular mycorrhizal symbiosis in rice. *New Phytol.* 225(4)**,** 1762-1776. doi: https://doi.org/10.1111/nph.16158**.**

Hungria, M., Campo, R.J., Souza, E.M., and Pedrosa, F.O. (2010). Inoculation with selected strains of *Azospirillum brasilense* and *A. lipoferum* improves yields of maize and wheat in Brazil. *Plant Soil* . 331(1)**,** 413-425. doi: 10.1007/s11104-009-0262-0.

Hymowitz, T. (1987). Introduction of the soybean to Illinois. *Econ. Bot.* 41(1)**,** 28-32. doi: 10.1007/BF02859342.

Hymowitz, T., and Harlan, J.R. (1983). Introduction of soybean to North America by Samuel Bowen in 1765. *Econ. Bot.* 37(4)**,** 371-379. doi: 10.1007/BF02904196.

Igiehon, N.O., and Babalola, O.O. (2018). Rhizosphere microbiome modulators: Contributions of nitrogen fixing bacteria towards sustainable agriculture. *Int J Environ Res Public Health* 15(4). doi: 10.3390/ijerph15040574.

Jiao, Y., Vuong, T.D., Liu, Y., Li, Z., Noe, J., Robbins, R.T., et al. (2015). Identification of quantitative trait loci underlying resistance to southern root-knot and reniform nematodes in soybean accession PI 567516 C. *Mol Breed* 35(6)**,** 131. doi: 10.1007/s11032-015-0330-5.

Kato, S., Sayama, T., Taguchi-Shiobara, F., Kikuchi, A., Ishimoto, M., and Cober, E. (2019). Effect of change from a determinate to a semi-determinate growth habit on the yield and lodging resistance of soybeans in the northeast region of Japan. *Breed. Sci* 69(1)**,** 151-159. doi: 10.1270/jsbbs.18112.

Kazi, S., Shultz, J., Afzal, J., Hashmi, R., Jasim, M., Bond, J., 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(3)**,** 633-644. doi: 10.1007/s00122-009-1181-4.

Kecskés, M.L., Choudhury, A.T.M.A., Casteriano, A.V., Deaker, R., Roughley, R.J., Lewin, L., et al. (2016). Effects of bacterial inoculant biofertilizers on growth, yield and nutrition of rice in Australia. *J. Plant Nutr.* 39(3)**,** 377-388. doi: 10.1080/01904167.2015.1016172.

Kim, Y.H., Kim, K.S., and Riggs, R.D. (2012). Initial subcellular responses of susceptible and resistant soybeans infected with the soybean cyst nematode. *Plant Pathol. J.* 28(4)**,** 401-408. doi: 10.5423/ppj.Oa.04.2012.0054.

Köberl, M., Müller, H., Ramadan, E.M., and Berg, G. (2011). Desert farming benefits from microbial potential in arid soils and promotes diversity and plant health. *PLOS ONE* 6(9)**,** e24452. doi: 10.1371/journal.pone.0024452.

Lakhssassi, N., Liu, S., Bekal, S., Zhou, Z., Colantonio, V., Lambert, K., et al. (2017). Characterization of the Soluble NSF Attachment Protein gene family identifies two members involved in additive resistance to a plant pathogen. *Sci Rep* 7**,** 45226. doi: 10.1038/srep45226.

Lee, T.G., Kumar, I., Diers, B.W., and Hudson, M.E. (2015). Evolution and selection of *Rhg*1, a copy-number variant nematode-resistance locus. *Mol Ecol* 24(8)**,** 1774-1791. doi: 10.1111/mec.13138.

Liu, S., Kandoth, P.K., Lakhssassi, N., Kang, J., Colantonio, V., Heinz, R., et al. (2017). The soybean *GmSNAP18* gene underlies two types of resistance to soybean cyst nematode. *Nat. Commun.* 8(1)**,** 14822. doi: 10.1038/ncomms14822.

Liu, S., Kandoth, P.K., Warren, S.D., Yeckel, G., Heinz, R., Alden, J., et al. (2012). A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492(7428)**,** 256-260. doi: 10.1038/nature11651.

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., et al. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488(7409)**,** 86-90. doi: 10.1038/nature11237.

MacGuidwin, A.E. (2012). Distribution of SCN HG types in Wisconsin. 2011. *Wiscosin Crop Production Association Distinguished Service Awards* 107.

Matson, A.L., and Williams, L.F. (1965). Evidence of a fourth gene for resistance to the soybean cyst nematode. *Crop Sci.* 5(5)**.** doi: 10.2135/cropsci1965.0011183X000500050032xa.

McCarville, M.T., Marett, C.C., Mullaney, M.P., Gebhart, G.D., and Tylka, G.L. (2017). Increase in soybean cyst nematode virulence and reproduction on resistant soybean varieties in Iowa from 2001 to 2015 and the effects on soybean yields. *Plant Health Prog.* 18(3)**,** 146-155. doi: 10.1094/php-rs-16-0062.

Meksem, K., Pantazopoulos, P., Njiti, V.N., Hyten, D., Arelli, P.R., and Lightfoot, D. (2001). 'Forrest' resistance to the soybean cyst nematode is bigenic: Saturation mapping of the *Rhg1* and *Rhg4* loci. *Theor. Appl. Genet.* 103**,** 710-717. doi: 10.1007/s001220100597.

Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J.H.M., et al. (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* 332(6033)**,** 1097-1100. doi: 10.1126/science.1203980.

Meyer, J.B., Lutz, M.P., Frapolli, M., Péchy-Tarr, M., Rochat, L., Keel, C., et al. (2010). Interplay between wheat cultivars, biocontrol *Pseudomonads,* and soil. *Appl. Environ* 76(18)**,** 6196-6204. doi: 10.1128/aem.00752-10.

Micallef, S.A., Shiaris, M.P., and Colón-Carmona, A. (2009). Influence of *Arabidopsis thaliana* accessions on rhizobacterial communities and natural variation in root exudates. *J Exp Bot* 60(6)**,** 1729-1742. doi: 10.1093/jxb/erp053.

Mitchum, M.G. (2016). Soybean resistance to the soybean cyst nematode *Heterodera glycines*: An Update. *Phytopathology* 106(12)**,** 1444-1450. doi: 10.1094/phyto-06-16-0227-rvw.

Mitchum, M.G., Wrather, J.A., Heinz, R.D., Shannon, J.G., and Danekas, G. (2007). Variability in distribution and virulence phenotypes of *Heterodera glycines* in Missouri during 2005. *Plant Dis* 91(11)**,** 1473-1476. doi: 10.1094/pdis-91-11-1473.

Naz, I., Bano, A., and Ul-Hassan, T. (2008). Isolation of phytohormones producing plant growth promoting rhizobacteria from weeds growing in Khewra salt range, to Pakistan and their implication in providing salt tolerance *Glycine max* L. *Afr. J. Biotechnol.* 8**,** 5762-5766.

Niblack, T.L., Arelli, P.R., Noel, G.R., Opperman, C.H., Orf, J.H., Schmitt, D.P., et al. (2002). A revised classification scheme for genetically diverse populations of *Heterodera glycines*. *J. Nematol.* 34(4)**,** 279-288.

Niblack, T.L., Colgrove, A.L., Colgrove, K., and Bond, J.P. (2008). Shift in virulence of soybean cyst nematode is associated with use of resistance from PI 88788. *Plant Health Prog.* 9(1)**,** 29. doi: 10.1094/PHP-2008-0118-01-RS.

Niblack, T.L., Lambert, K.N., and Tylka, G.L. (2006). A model plant pathogen from the kingdom Animalia: *Heterodera glycines*, the soybean cyst nematode. *Annu. Rev. Phytopathol.* 44(1)**,** 283-303. doi: 10.1146/annurev.phyto.43.040204.140218.

Noel, G.R. (1986). "The Soybean Cyst Nematode," in *Cyst Nematodes,* eds. F. Lamberti & C.E. Taylor. (Boston, MA: Springer US), 257-268.

Patel, K., and N.Amaresan (2014). Antimicrobials compounds from extreme environment rhizosphere organisms for plant growth. *Int.J.Curr.Microbiol.App.Sci* 3(7)**,** 651-664.

Patil, G.B., Lakhssassi, N., Wan, J., Song, L., Zhou, Z., Klepadlo, M., et al. (2019). Whole-genome re-sequencing reveals the impact of the interaction of copy number variants of

the *rhg1* and *Rhg4* genes on broad-based resistance to soybean cyst nematode. *Plant Biotechnol. J.* 17(8)**,** 1595-1611. doi: 10.1111/pbi.13086.

Perry, R.N. (1996). Chemoreception in plant parasitic nematodes. *Annu. Rev. Phytopathol.* 34(1)**,** 181-199. doi: 10.1146/annurev.phyto.34.1.181.

Pimentel, M.F., Arnão, E., Warner, A.J., Subedi, A., Rocha, L.F., Srour, A., et al. (2020). *Trichoderma* isolates inhibit *Fusarium virguliforme* growth, reduce root rot, and induce defense-related genes on soybean seedlings. *Plant Dis.* 104(7)**,** 1949-1959. doi: 10.1094/pdis-08-19-1676-re.

Qiao, Q., Wang, F., Zhang, J., Chen, Y., Zhang, C., Liu, G., et al. (2017). The variation in the rhizosphere microbiome of cotton with soil type, genotype and developmental stage. *Sci. Rep.* 7(1)**,** 3940-3940. doi: 10.1038/s41598-017-04213-7.

Rascovan, N., Carbonetto, B., Perrig, D., Díaz, M., Canciani, W., Abalo, M., et al. (2016). Integrated analysis of root microbiomes of soybean and wheat from agricultural fields. *Sci. Rep.* 6(1)**,** 28084. doi: 10.1038/srep28084.

Riggs, R.D. (1977). Worldwide distribution of soybean-cyst nematode and its economic importance. *J. Nematol.* 9(1)**,** 34-39.

Riggs, R.D., and Schmitt, D.P. (1988). Complete characterization of the race scheme for *Heterodera glycines*. *J. Nematol.* 20(3)**,** 392-395.

Rincker, K., Cary, T., and Diers, B.W. (2017). Impact of soybean cyst nematode resistance on soybean yield. *Crop Sci.* 57(3)**,** 1373-1382. doi: 10.2135/cropsci2016.07.0628.

Rizzo, G., and Baroni, L. (2018). Soy, soy foods and their role in vegetarian diets. *Nutrients* 10(1)**,** 43. doi: 10.3390/nu10010043.

Roman-Reyna, V., Pinili, D., Borja, F.N., Quibod, I.L., Groen, S.C., Alexandrov, N., et al. (2020). Characterization of the leaf microbiome from whole-genome sequencing data of the 3000 Rice Genomes Project. *Rice* 13(1)**,** 72. doi: 10.1186/s12284-020-00432-1.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463(7278)**,** 178-183. doi: 10.1038/nature08670.

Schulz-Bohm, K., Martín-Sánchez, L., and Garbeva, P. (2017). Microbial Volatiles: Small molecules with an important role in intra- and inter-kingdom interactions. *Front. Microbiol.* 8(2484). doi: 10.3389/fmicb.2017.02484.

Schuster, I., Abdelnoor, R.V., Marin, S.R.R., Carvalho, V.P., Kiihl, R.A.S., Silva, J.F.V., et al. (2001). Identification of a new major QTL associated with resistance to soybean cyst nematode (*Heterodera glycines*). *Theor. Appl. Genet.* 102(1)**,** 91-96. doi: 10.1007/s001220051622.

Shi, Z., Liu, S., Noe, J., Arelli, P., Meksem, K., and Li, Z. (2015). SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. *BMC Genet.* 16(1)**,** 314-314. doi: 10.1186/s12864-015-1531-3.

Singh, R.J. (2017). "Botany and Cytogenetics of Soybean," in *The Soybean Genome*, eds. H.T. Nguyen & M.K. Bhattacharyya. (Cham: Springer International Publishing), 11-40.

Smith, G.J., Wiebold, W.J., Niblack, T.L., Scharf, P.C., and Blevins, D.G. (2001). Macronutrient concentrations of soybean infected with soybean cyst nematode. *Plant Soil* . 235(1)**,** 21-26. doi: 10.1023/A:1011872818069.

Suzuki, C., Taguchi-Shiobara, F., Ikeda, C., Iwahashi, M., Matsui, T., Yamashita, Y., et al. (2020). Mapping soybean *rhg2* locus, which confers resistance to soybean cyst nematode

race 1 in combination with *rhg1* and *Rhg4* derived from PI 84751. *Breed Sci* 70(4)**,** 474-480. doi: 10.1270/jsbbs.20035.

Tabor, G.M., Tylka, G.L., Behm, J.E., and Bronson, C.R. (2003). *Heterodera glycines* infection increases incidence and severity of brown stem rot in both resistant and susceptible soybean. *Plant Dis.* 87(6)**,** 655-661. doi: 10.1094/pdis.2003.87.6.655.

Tian, Y., Liu, B., Shi, X., Reif, J.C., Guan, R., Li, Y.-h., et al. (2019). Deep genotyping of the gene *GmSNAP* facilitates pyramiding resistance to cyst nematode in soybean. *Crop J.* 7(5)**,** 677-684. doi: 10.1016/j.cj.2019.04.003.

Tran, D.T., Steketee, C.J., Boehm, J.D., Noe, J., and Li, Z. (2019). Genome-wide association analysis pinpoints additional major genomic regions conferring resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe). *Front. Plant Sci.* 10(401). doi: 10.3389/fpls.2019.00401.

Tsavkelova, E.A., Cherdyntseva, T.A., and Netrusov, A.I. (2005). Auxin production by bacteria associated with orchid roots. *Microbiology* 74(1)**,** 46-53. doi: 10.1007/s11021-005-0027-6.

Turner, T.R., James, E.K., and Poole, P.S. (2013). The plant microbiome. *Genome Biology* 14(6)**,** 209. doi: 10.1186/gb-2013-14-6-209.

Tylka, G.L. (2020). SCN-resistant soybean varieties for Iowa - By the Numbers. *ICM News* 2660. doi: https://lib.dr.iastate.edu/cropnews/2660**.**

Tylka, G.L., Gebhart, G.D., Marett, C.C., and Mullane, M.P. (2020). Evaluation of soybean varieties resistant to soybean cyst nematode in Iowa—2020. *IPM 52 Iowa State University*.

Tylka, G.L., and Marett, C.C. (2021). Known distribution of the soybean cyst nematode, *Heterodera glycines*, in the United States and Canada in 2020. *Plant Health Prog.* 22(1)**,** 72-74. doi: 10.1094/php-10-20-0094-br.

Usovsky, M., Ye, H., Vuong, T.D., Patil, G.B., Wan, J., Zhou, L., et al. (2021). Fine-mapping and characterization of qSCN18, a novel QTL controlling soybean cyst nematode resistance in PI 567516C. *Theor Appl Genet* 134(2)**,** 621-631. doi: 10.1007/s00122-020-03718-6.

Valliyodan, B., Cannon, S.B., Bayer, P.E., Shu, S., Brown, A.V., Ren, L., et al. (2019). Construction and comparison of three reference-quality genome assemblies for soybean. *Plant J.* 100(5)**,** 1066-1082. doi: https://doi.org/10.1111/tpj.14500**.**

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

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

Wallace, J.G., Kremling, K.A., Kovar, L.L., and Buckler, E.S. (2018). Quantitative genetics of the maize leaf microbiome. *Phytobiomes J.* 2(4)**,** 208-224. doi: 10.1094/PBIOMES-02-18-0008-R.

Wang, D., Diers, B.W., Arelli, P.R., and Shoemaker, R.C. (2001). Loci underlying resistance to Race 3 of soybean cyst nematode in *Glycine soja* plant introduction 468916. *Theor. Appl. Genet.* 103(4)**,** 561-566. doi: 10.1007/PL00002910.

Wang, J., Niblack, T.L., Tremain, J.A., Wiebold, W.J., Tylka, G.L., Marett, C.C., et al. (2003). Soybean cyst nematode reduces soybean yield without causing obvious aboveground symptoms. *Plant Dis.* 87(6)**,** 623-628. doi: 10.1094/pdis.2003.87.6.623.

Webb, D.M., Baltazar, B.M., Rao-Arelli, A.P., Schupp, J., Clayton, K., Keim, P., et al. (1995). Genetic mapping of soybean cyst nematode race-3 resistance loci in the soybean PI 437.654. *Theor. Appl. Genet.* 91(4)**,** 574-581. doi: 10.1007/BF00223282.

Weinert, N., Piceno, Y., Ding, G.C., Meincke, R., Heuer, H., Berg, G., et al. (2011). PhyloChip hybridization uncovered an enormous bacterial diversity in the rhizosphere of different potato cultivars: many common and few cultivar-dependent taxa. *FEMS Microbiol Ecol* 75(3)**,** 497-506. doi: 10.1111/j.1574-6941.2010.01025.x.

Westcott, S.L., and Schloss, P.D. (2015). De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. *PeerJ* 3**,** e1487. doi: 10.7717/peerj.1487.

Wilson, R.F. (2004). "Seed Composition," in *Soybeans: Improvement, Production, and Uses*.), 621-677.

Winsor, S. (2020). Soybean cyst nematode management in the Western Corn Belt. *Crops & Soils* 53(5)**,** 4-12. doi: https://doi.org/10.1002/crso.20063**.**

Wu, X., Blake, S., Sleper, D.A., Shannon, J.G., Cregan, P., and Nguyen, H.T. (2009). QTL, additive and epistatic effects for SCN resistance in PI 437654. *Theor Appl Genet* 118(6)**,** 1093-1105. doi: 10.1007/s00122-009-0965-x.

Xing, L., and Westphal, A. (2006). Interaction of *Fusarium solani f. sp. glycines* and *Heterodera glycines* in Sudden Death Syndrome of Soybean. *Phytopathology* 96(7)**,** 763-770. doi: 10.1094/phyto-96-0763.

XongHong, L., Jian-qiang, L., and Dong-sheng, Z. (1997). History and status of soybean cyst nematode in China. *International J. Nematol.* 7**,** 18-25.

Xu, Y., Wang, G., Jin, J., Liu, J., Zhang, Q., and Liu, X. (2009). Bacterial communities in soybean rhizosphere in response to soil type, soybean genotype, and their growth stage. *Soil Biol. Biochem.*41(5)**,** 919-925. doi: 10.1016/j.soilbio.2008.10.027**.**

Yu, N., and Diers, B.W. (2017). Fine mapping of the SCN resistance QTL cqSCN-006 and cqSCN-007 from *Glycine soja* PI 468916. *Euphytica* 213(2)**,** 54. doi: 10.1007/s10681-016-1791-2.

Yu, N., Lee, T.G., Rosa, D.P., Hudson, M., and Diers, B.W. (2016). Impact of *Rhg1* copy number, type, and interaction with *Rhg4* on resistance to *Heterodera glycines* in soybean. *Theor. Appl. Genet.* 129(12)**,** 2403-2412. doi: 10.1007/s00122-016-2779-y.

Yue, P., Sleper, D.A., and Arelli, P.R. (2001). Mapping resistance to multiple races of *Heterodera glycines* in soybean PI 89772. *Crop Sci.* 41(5)**,** 1589-1595. doi: https://doi.org/10.2135/cropsci2001.4151589x**.**

Zhang, L.X., Kyei-Boahen, S., Zhang, J., Zhang, M.H., Freeland, T.B., Watson, C.E., et al. (2007). Modifications of optimum adaptation zones for soybean maturitygroups in the USA. *Crop Management* 6(1)**,** 1-11. doi: 10.1094/cm-2007-0927-01-rs.

Zhou, L., Song, L., Lian, Y., Ye, H., Usovsky, M., Wan, J., et al. (2021). Genetic characterization of qSCN10 from an exotic soybean accession PI 567516 C reveals a novel source conferring broad-spectrum resistance to soybean cyst nematode. *Theor Appl Genet* 134(3)**,** 859-874. doi: 10.1007/s00122-020-03736-4.

Zhu, H., Choi, H.-K., Cook, D.R., and Shoemaker, R.C. (2005). Bridging model and crop legumes through comparative genomics. *Plant Physiol.* 137(4)**,** 1189-1196. doi: 10.1104/pp.104.058891.

Tables and Figures

Table 1.1. Race classification system for soybean cyst nematode (SCN) populations. Race determination is based on the female index of four indicator lines using 'Lee' as a susceptible check.

Race	PI 548988 ('Pickett')	PI 548402 ('Peking')	PI 88788	PI 90763
1	- [a]	-	+	-
2	+	+	+	-
3	-	-	-	-
4	+	+	+	+
5	+	-	+	-
6	+	-	-	-
7	-	-	+	+
8	-	-	-	+
9	+	+	-	-
10	+	-	-	+
11	-	+	+	-
12	-	+	-	+
13	-	+	-	-
14	+	+	-	+
15	+	-	+	+
16	-	+	+	+

[a] + denoted that the female index of an indicator line is >10% and "- indicated that the female index is <10%.

Table 1.2. HG type test using seven indicator lines to describe diverse SCN populations. HG type is determined by calculating the female index using 'Lee 74' as a standard susceptible check and a 10% rule with "+" and "-" is similar to the race classification system.
HG type is named as a list of number listed in the table. Ex: HG type 1.2.- indicates a population that has a FI ≥ 10% on PI 548402 (No.1); PI 88788 (No.2); PI 209332 (No.5) and PI 548316 (No.7)

Number	Indicator line
1	PI 548402 ('Peking')
2	PI 88788
3	PI 90763
4	PI 437654
5	PI 209332
6	PI 89772
7	PI 548316 ('Cloud')

Table 1.3 List of designated QTLs controlling for soybean cyst nematode resistance by Soybean Genetics Committee.

QTL name	Chromosome	Position (Mb) [a]	Resistance to SCN [a]	Resistant parents [a]
cqSCN-001 (Rhg1)	18	16.2-19.7	Race 1, 2, 3, 5, 14 and LY1	PI 89772, 'Forrest', PI 90763. PI 437654, PI 404198 A, PI 438489 B, 'Hartwig', 'Peking', PI 88788
cqSCN-002 (Rhg4)	8	6.8- 7.8	Race 1, 2, 3	'Forrest', 'Peking', PI 90763, PI 437654, PI 404198 A, PI 438489 B, 'Hartwig'
cqSCN-003 (Rhg5)	16	34.0- 36.2	Race 2, 3, 14	PI 90763, 'Hartwig', PI 437655, PI 88788
cqSCN-005 (Rhg3)	17	20.1-27.3	Race 1, 14	'Hartwig', PI 89772, L10
cqSCN-006	15	7.0-39.6	Race 1, 3	PI 468916
cqSCN-007	18	50.7-53.8	Race 1, 3	PI 468916
Rhg2	7	16.5-18.3	Race 3	PI 437654

[a] The position is based on 'William 82' version 2. Information is from Soybase.org.

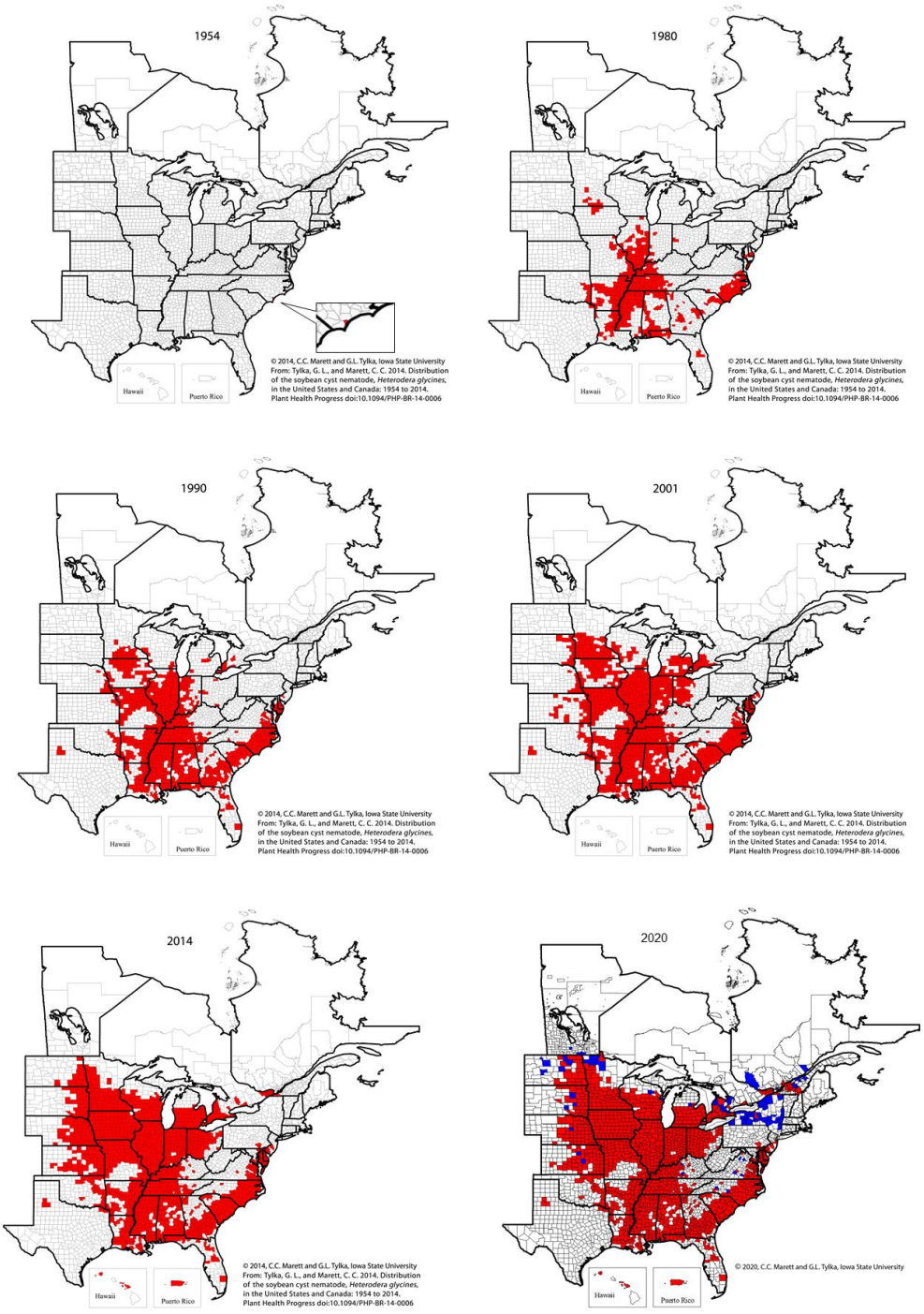


Fig 1.1 Distribution of soybean cyst nematode (SCN) in the U.S. counties in 1954 and 2020. Infested areas before 2017 are indicated in red and infested areas after 2017 are indicated in blue. Source: Tylka and Marett (2021)

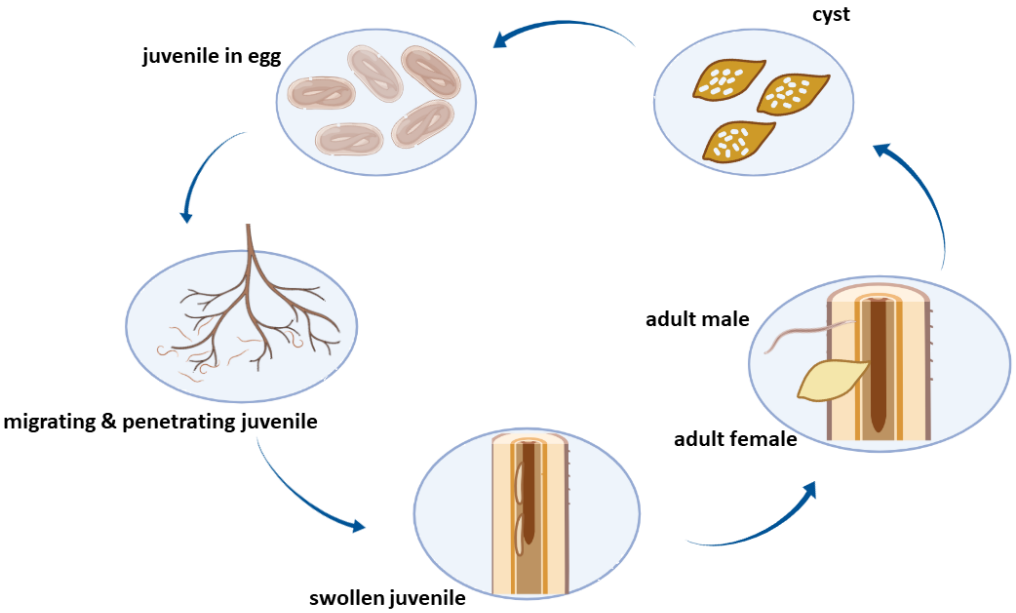


Fig 1.2 Life cycle of the soybean cyst nematode (SCN). The 1st stage juvenile is in the egg and the 2nd stage juvenile hatches from eggs and penetrates plant roots. The 3rd and 4th stage juveniles feed off the plants. Adult males migrate out of the roots in search of females. Adult females rupture the root and after death, their bodies become the cysts containing eggs. The graph was created by BioRender.com

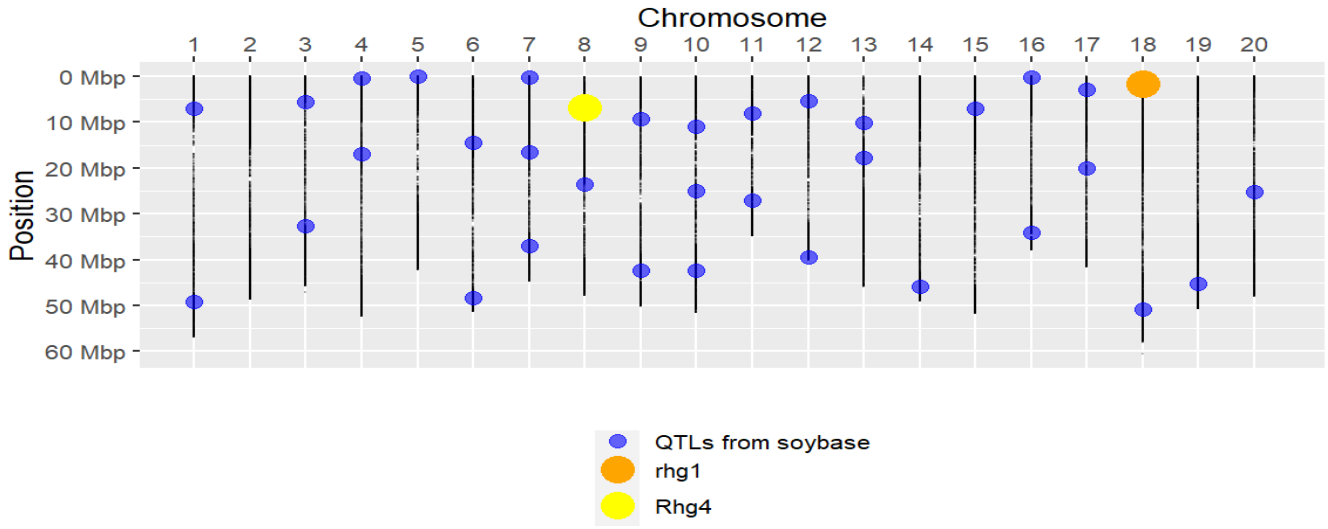


Fig 1.3 QTLs from mapping studies were reported on the SoyBase database for SCN resistance. *Rhg1* and *Rhg4* on Chrs 18 and 8, respectively are two major QTLs for SCN resistance.

CHAPTER 2

IDENTIFICATION AND CHARACTERIZATION OF QUANTITATIVE TRAIT LOCI UNDERLYING RESISTANCE TO SOYBEAN CYST NEMATODE (*HETERODERA GLYCINES*) IN SOYBEAN PI 567295[1]

[1] D. T. Tran, M. Mitchum, and Z. Li. To be submitted to *BMC Genomics*

Abstract

Soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is the most damaging pest 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 sources has led some SCN populations to overcome the resistance provided by *the rhg1* and *Rhg4* alleles. It is critical to identify new sources of SCN resistance beyond the defense mechanisms found at the *Rhg1* and *Rhg4* loci. Our greenhouse screening studies revealed that PI 567295 was resistant to two SCN populations but did not carry *rhg1* and *Rhg4* resistance alleles. Thus, the objective of this study was to identify quantitative trait loci (QTLs) underlying HG type 0 (SCN race 3) and HG type 1.2.- (SCN race 2) in PI 567295. Using an $F_{2:3}$ population derived from 'Bossier' × PI 567295 that was phenotyped with HG type 0 (race 3) in greenhouse, the bulked segregant analysis (BSA) using the SoySNP50K Infinium BeadChip identified six putative genomic regions on Chromosomes (Chrs) 2, 5, 10, 11, 14, and 19 for SCN resistance. A recombinant inbred line (RIL) population derived from the same parents was then phenotyped with HG type 0 and HG type 1.2.- and genotyped with SoySNP6K BeadChip for QTL mapping. Based on SoySNP50K BeadChip data, 70 KASP markers were also designed in the BSA regions for genotyping. QTL analysis identified a major QTL on Chr 10 for both HG types and two minor QTLs on Chrs 6 and 8 for HG type 1.2.-. The major QTL on Chr 10 explained 39.3 and 53.7% of phenotypic variation for HG type 1.2.- and 0, respectively; and two minor QTLs on Chrs 6 and 8 accounted for 13 and 6.2% of phenotypic variation for HG type 1.2.-, respectively. The result provides an additional resistant source and SNP marker information for developing SCN resistant cultivars.

Keywords: Soybean cyst nematode (SCN); *Heterodera glycines* (HG) type; Resistance to *Heterodera glycines* (*Rhg*); Quantitative trait locus (QTL); Whole genome resequencing (WGRS)

Introduction

Soybean [*Glycine max* (L.) Merr.] is an important source of protein for humans and livestock. The U.S. is the second largest producer, accounting for 29% of the world's soybean production with 96.7 million metric tons annually (soystats.com, 2021). However, soybean production in the U.S is economically challenged by the soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) with an estimated loss of 76.14 Million Metric Tons in 2020 (Bradley et al., 2021), accounting for more than $1.5 billion loss annually (Bandara et al., 2020). More than 30% of yield loss can occur in soybean cyst nematode infested fields without noticeable aboveground symptoms (Wang et al., 2003).

SCN was first reported in North Carolina in the U.S. in 1954 and to date, it has spread to most soybean-producing states (Tylka and Marett, 2021). SCN is an obligate plant parasite that requires infection of the host plants to complete its life cycle. After hatching, the second stage juveniles (J2) will travel through soil and penetrate soybean roots using their stylets. Moving toward the vascular cylinder, SCN selects a single cell near the vascular tissue to establish a permanent feeding site called the syncytium that absorbs essential nutrients. The feeding sites are formed by progressive cell wall dissolution and merging of hundreds of cells (Mitchum, 2016).

Different levels of virulence were firstly classified using a race designation system, consisting of four indicator lines: Plant Introduction (PI) 548988 ('Pickett'), PI 548402 ('Peking'), PI 88788, and PI 90763 and a susceptible check, PI 548656 ('Lee') (Golden and Et, 1970b). However, this classification system was not suitable to capture proper variability of SCN

due to its genetic diversity. Therefore, a new classification system, HG type, was developed to characterize SCN virulence. The reproductive capability of a SCN population is based on seven indicator lines that are marked numerically from number 1 to 7 based on the order of reported source of resistance. They are: 1) PI 548402 ('Peking'), 2) PI 88788, 3) PI 90763, 4) PI 437654, 5) PI 209332, 6) PI 89772, and 7) PI 548316 ('Cloud') (Niblack et al., 2002).

To better understand the molecular mechanism of resistance and increase efficiency in development of new resistant varieties, quantitative trait locus (QTL) mapping for SCN resistance has been conducted since 1992 with the discovery of the first *Rhg1* locus (resistant to *Heterodera glycines*) on Chromosome (Chr) 18, which is associated with resistance to multiple SCN populations including HG type 0 (race 3) (Concibido et al., 2004). This QTL was detected in the majority of resistant accessions including all seven indicator lines in HG type classification, accounting for up to 50% of the phenotypic variation observed for SCN race 3 resistance (Mudge et al., 1997; Concibido et al., 2004). This *Rhg1* locus comprises four genes but only three of these four genes conferred SCN resistance including *Glyma.18g022400* (putative amino acid transporter), *Glyma.18g022500* (α-Soluble NSF Attachment Protein. α - SNAP), and *Glyma.18g022700* (putative wound-inducible protein, WIP) (Cook et al., 2012), . Silencing any of three genes reduced SCN resistance in 'Fayette', a resistant variety that carries *rhg1* resistance allele (Cook et al., 2012). SCN resistance is also mediated by copy number variation of these genes (Cook et al., 2012; Cook et al., 2014). A SCN susceptible cultivar, 'William 82' carries only a single copy, while resistant genotypes carry more than two copies (Cook et al., 2014). Based on number of copies, SCN resistant genotypes have been classified as *rhg1 HC* (high copy) including PI 88788 carrying more than six copies; and *rhg1 LC* (low copy) including 'Peking' carrying from 2 to 4 copies (Cook et al., 2012; Cook et al., 2014).

Multiple studies of haplotype analysis showed two unique resistant haplotypes that also correlates with the copy number of *rhg1: LC* and *HC* groups (Liu et al., 2017; Patil et al., 2019; Tran et al., 2019). Two resistance groups also exhibited two structural classes based on the unique C terminal of the *α-SNAP* gene (*Glyma.18g022500)*. The 'Peking' resistance group exhibited a copy of retrotransposon called "RAC" within the first intron of *α-SNAP* resistance gene in each repeat (Bayless et al., 2018).

The second major QTL, *Rhg4* locus, a *serine hydroxymethyltransferase* (*GmSHMT08*) gene was mapped on Chr 8 (Concibido et al., 2004). This QTL explained about 28% of phenotypic variation (Weisemann et al., 1992). With presence of the *Rhg4* susceptible allele, the resistance was lost in 'Peking'-derived SCN resistance (Brucker et al., 2005; Yu et al., 2016). The copy number variation pattern was also detected at the *Rhg4* locus. Susceptible genotypes and PI 88788-type resistance only carried one copy, while the 'Peking'-type resistance carried up to 4.3 copies (Patil et al., 2019). The causal gene *GmSHMT8* was a predicted cytosolic enzyme that differs by two amino acids that reside in the folate and pyridoxal-5-phosphate binding sites of enzyme (Liu et al., 2012).

Based on cloning information of these genes at *Rhg1* and *Rhg4* loci 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 selection of SCN resistance at *Rhg1* and *Rhg4* loci using a Kompetitive Allele Specific PCR (KASP) assay. Two of these KASP markers (GSM381 and GSM383) were designed for selection of the r*hg1* resistance alleles, of which GSM383 was used to distinguish the 'Peking' and PI 88788-type resistance. GSM191 was designed for selection of the resistance allele at the *Rhg4* locus.

To date, both types of resistance have been extensively utilized by soybean breeders to develop SCN resistant cultivars, especially PI 88788-type resistance. Of 849 resistant cultivars surveyed, 95% had SCN resistance derived from PI 88788 (Tylka, 2020). Due to the fact that soybean breeders have primarily used 'Peking' and PI 88788-type resistance to develop soybean cultivars for SCN resistance, some of SCN populations have shifted, rendering the PI 88788 SCN resistance not effective (Gardner et al., 2017). In fact, some SCN populations have been reported in Missouri, Illinois, and North Dakota as virulent to some indicator lines including PI 88788 and 'Peking' (Howland et al., 2018; Chowdhury et al., 2021). Therefore, it is critical to find new sources of SCN resistance beyond *Rhg1* and *Rhg4* loci to support SCN resistance breeding efforts.

In addition to *Rhg1* and *Rhg4* loci, more than 200 genomic regions for SCN resistance have been reported but only six QTLs were confirmed and designated as cqSCN by the Soybean Genetic Committee. Of these six, *Rhg1* was designated as *cqSCN-001* and *Rhg4* as *cqSCN-002*. The *cqSCN-003* located between two SSR markers: Satt547 and Satt431 on Chr 16 conferring HG type 7 (race 3) and HG type 1.3.5.6.7 (race 14) in PI 88788 (Glover et al., 2004). This QTL was also mapped in PI 90763 (Concibido et al., 1997) and PI 209332 (Concibido et al., 1996) for HG type 0 resistance. The *cqSCN-005* on Chr 17 was detected in 'Hartwig', a SCN resistant cultivar derived from PI 437654, explaining 11% of phenotypic variation for HG type 1.2.5- (SCN race 2) (Kazi et al., 2010).

Two QTLs: *cqSCN-006* and c*qSCN-007* on Chrs 15 and 18, respectively, conferring resistance for SCN race 3 were detected in a wild soybean accession PI 468916 (Kim and Diers, 2013). The *cqSCN-006* was fine-mapped to a 212.1 kb region, explaining 23% of phenotypic variation for HG type 2.5.7 resistance, while *cqSCN-007* was fine mapped to a 103.2 kb region

on Chr 18, accounting for 27% of phenotypic variation. PI 567516 C was reported as a new SCN resistance source, carrying two QTLs: *qSCN10* and *qSCN18* for HG type 0 resistance on Chrs 10 and 18, respectively (Vuong et al., 2010). The QTLs were fine mapped to 166 kb and 379 kb regions on Chrs 18 and 10, respectively using an NIL population derived from PI 567516 C (Usovsky et al., 2021; Zhou et al., 2021). Although there have been many QTL mapping studies for SCN resistance, most of them have not been confirmed in different genetic backgrounds or had minor effects on SCN resistance. Moreover, due to the variation of SCN populations, not every reported QTL was related to all SCN populations.

Herein, in an effort to identify novel QTLs for SCN resistance, nearly 700 soybean accessions with various origins were screened for resistance to HG type 0 (SCN race 3) using an inoculation assay conducted at the University of Georgia greenhouse. Of these accessions, PI 567295 from China was identified as being moderately resistant to both HG type 0 (SCN race 3) and HG type 1.2.-1.2.- (SCN race 2). Genotyping with three functional markers: GSM381; GSM383 and GSM191 (Shi et al., 2015) indicated that PI 567295 did not carry the same resistance alleles as 'Peking' and PI 88788 at the *Rhg1* or *Rhg4* locus (Tran et al., 2019).

Thus, the objective of this work was to discover the underlying QTLs associated with HG type 0 and 1.2.- in PI 567925 using $F_{2:3}$ and RIL populations derived from a cross of 'Bossier' (susceptible) × PI 567295 and develop breeder-friendly SNP markers for marker-assisted selection of SCN resistance.

Materials and methods

Population development

The $F_{2:3}$ and F_5-drived RIL populations were developed by crossing 'Bossier' (PI 567789) with PI 567295. 'Bossier' is a maturity group (MG) VIII cultivar and highly susceptible to HG

type 0 (race 3) and HG type 1.2.- (race 2). PI 567295 is a plant introduction in MG VIII which was identified as moderate resistance to HG type 0 (FI =12%) and HG type 1.2.- (FI=23.8%). The cross of 'Bossier' × PI 567295 was made in the summer of 2017. The F_1 seeds from this cross were grown in the winter nursery in Puerto Rico and then the F_2 plants were grown in the summer of 2018 in the University of Georgia Iron Horse Farm (IHF), Watkinsville, GA and harvested individually to create $F_{2:3}$ families. In the meantime, F_3 and F_4 generations were advanced through a single seed descent method in winter nursery in Puerto Rico during the winter of 2018-2019. The F_5 generation was planted at the IHF in the summer of 2019. The 259 F_5 plants were individually harvested as RILs. Of these 259 F_5-derived RILs, 144 were randomly chosen for this study.

Phenotyping of the $F_{2:3}$ and RIL populations in greenhouse

Inbred soybean cyst nematode populations HG type 0 and HG type 1.2.- were originally from Missouri then maintained on 'Williams 82' and PI 88788, respectively, in a greenhouse at the University of Georgia. The "-" in the name of HG type 1.2.- indicates that the HG classification was based on the modified HG type test using first four indicator lines by excluding PI 209332, PI 89772, and 'Cloud'. PI 209332 and 'Cloud' carry the same *rhg1-b* resistance allele as PI 88788. It is expected that both PI 209332 and 'Cloud' would have a similar response to the SCN population as PI 88788. Based on the both SCN classification systems, the HG types 0 and 1.2.- are equivalent to races 3 and 2 respectively. HG type 1.2.- is more aggressive than HG type 0 because both 'Peking' and PI 88788 were susceptible (Table 2.1). SCN phenotyping assays were conducted in the University of Georgia Plant Pathology Greenhouse in Athens, GA.

The 120 $F_{2:3}$ families derived from 'Bossier' × PI 567295 were evaluated for HG type 0 resistance with two replications for a bulked segregant analysis. The 198 F_5 -derived RILs from the same parents were evaluated with both HG type 0 and HG type 1.2.- with six replications per line for the mapping study. Soybean seeds were first germinated in germination paper pouches (CYG Seed Germination Pouch, Newport, MN, U.S.) for 3 days. Each seedling was then transplanted into a plastic cone with a length of 20.6 cm that was filled with fumigated sandy loam soil. Cones were arranged into a randomized complete block design (RCBD). Due to a large number of plants, six replicates of the RIL population were divided into three sub-tests with two replicates per test, so each RIL was phenotyped for a total of six plants. Each cone representing a single replicate was inoculated with 1,500 eggs of HG type 0 and HG type 1.2.-, respectively using an automatic pipetting machine. Approximately 28 days after inoculation, roots were washed individually free of soil, and then flushed with water to collect cysts through a 20-60 µm nested aperture sieve. Cysts on all roots from each individual plant were counted under a stereo microscope. The level of reaction was classified based on Female Index (FI) [(average number of cysts on a given line/ average number of cysts on 'Lee 74') × 100%]. The rating scores by Schmitt and Shannon (1992) were followed with FI < 10% (Resistant), 10 < FI < 30% (Moderately Resistant), 30 < FI < 60 (Moderately Susceptible), and FI > 60% (Susceptible).

Bulked segregant analysis

Based on the phenotyping score of the $F_{2:3}$ population in a greenhouse test with two replicates, two bulks: resistant and susceptible were created for each replicate. An equal amount of leaf tissue from each of 20 highly resistant families with FI <10% was combined as a resistant bulk. Similarly, an equal amount of leaf tissue of 20 highly susceptible families with FI > 60% were combined as a susceptible bulk. Genomic DNA was isolated using a modified CTAB

protocol and DNA was diluted in water to obtain a final concentration of at least 100 ng/µL. Four resistant and susceptible bulks for two replicates along with two parents were genotyped with the SoySNP50K iSelect BeadChip in Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD. SNP genotypes were called using GenomeStudio 2.0 (Illumina, San Diego, U.S.). Putative genomic regions from BSA were determined when the homozygous genotype alleles of the susceptible bulk matched with those of the susceptible parent and SNP alleles of the resistant bulk matched (both heterozygous and homozygous) with those from the resistant parent.

KASP SNP assay design and genotyping

Bulked 12 young leaves of each RIL and parents were used for genomic DNA extraction using the protocol mentioned previously. The RIL population along with two parents were then genotyped with the SoySNP6K iSelect BeadChip in Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD. Monomorphic markers between susceptible and resistant parents were removed, leaving 2,470 polymorphic SNP markers to be used for construction of a linkage map.

Based on the BSA result, putative genomic regions associated with HG type 0 resistance were selected for genotyping. Three KASP SNP markers were first designed for each genomic region. To further saturate the genomic regions, additional KASP SNP markers were added. KASP markers were then used to genotype both $F_{2:3}$ families and F_{5}–derived RILs. The genotyping was conducted by following Pham et al. (2013) for the master mix preparation and thermocycling condition. The endpoint reading was carried out using Roche Light Cycler 480 II (Roche Diagnostics Corporation Indianapolis, IN).

QTL Analysis

Polymorphic SNP markers from SoySNP6K iSelect BeadChip and designed KASP SNP markers from informative BSA regions were used to construct the genetic map. The linkage map of the RIL population was created using Kosambi's regression model function in JoinMap 4.1 (Ooijen et al., 2006). JoinMap calculated an estimate of pairwise recombination frequencies and LOD scores, and logarithm of odds criterion of greater than 5.0 and a maximum distance < 50 cM were used to establish linkage groups. Regression mapping with the Kosambi function was used to convert recombination frequencies to centiMorgans (cM). SNP quality control for segregation distortion *(P<0.005*) and minor allele frequency (MAF < 0.15) were performed, leaving a total of 1,804 polymorphic SNP markers to be used in QTL mapping.

R/qtl (Broman et al., 2003) was used to map the QTLs in the RIL population derived from 'Bossier' × PI 567295. The covariates of three markers and a window size of 5 cM were selected to perform composite interval mapping (CIM). The significance of the LOD threshold was determined by 1000 permutations with a significance level of $\alpha = 0.05$. The QTL confidence interval was estimated with the function 'bayesint' using a Bayesian credible interval.

Genotype response to nematode development in roots

To understand SCN invasion and development in two parents, seedlings of both parents were inoculated with 3,000 HG type 0 eggs. Roots were carefully removed from the soil at 2, 4, and 8 days after inoculation (DAI), respectively. Two root samples of each parent were then harvested at each time point. Roots were stained by following the protocol of Byrd et al. (1983). In brief, roots were washed with tap water to clear soil and soaked in 5.25% sodium hypochlorite for 7 min. Then, roots were rinsed with water for a few min and boiled in the acid fuchsin staining solution for one min followed by cooling down at room temperature for 30 mins and

rinsing the roots with water. The roots were then de-stained with warm glycerin and then incubated at 65^{o}C overnight and nematodes were examined under a dissecting microscope.

Whole genome re-sequencing

To identify the candidate genes for SNP marker development, one highly resistant RIL G19-SPT1748, one highly susceptible RIL G19-SPT1719, and two parents were selected and sequenced (Table 2.2). Library preparation and sequencing were performed using Illumina Novaseq/Hiseq PE150 (Illumina, San Diego, CA, U.S) in CD Genomics (Shirley, NY, U.S). The Illumina reads were trimmed with Trimmomatic version 0.39 (Bolger et al., 2014), then the trimmed reads were mapped on to the reference sequence, 'Williams 82' version 2 (Phytozome.com) using Burrows-Wheeler Alinger (BWA) version 0.7.17 (Li and Durbin, 2009). Reads were mapped using the default settings of the mem function. Alignment was further processed using the program Picard v 2.21.6 (Broad Institute, accessed Aug 2021) to mark PCR duplicates. Sequence alignments were processed for variant discovery using the GATK software package v4.1.8.1 (McKenna et al., 2010). After the variant filtration step, the dataset containing 3,696,678 SNPs and 831,752 Indels was used for further analysis. Variants where G19-SPT1748 and PI 567925 were homozygous and different from 'Bossier' and RIL G19-SPT1719 were evaluated. Variants that change amino acids in candidate genes were selected to design KASP markers for testing association between phenotype and genotype of the RIL population derived from 'Bossier' × PI 567295.

Statistical analysis

To minimize the environmental effect, Best Linear Unbiased Prediction (BLUP) values were used as phenotyping scores for QTL analysis. BLUP values were calculated across all three tests using SAS Studio Version 9.4 (SAS Institute Inc., Cary, NC, U.S). The model was built by

treating genotype and genotype by test interaction as random variables. Use of BLUP values for each genotype across tests helped to account for variation caused by environmental factors.

Results

Resistance response of PI 567295 to SCN

PI 567295, the resistant parent had 12% in FI for HG type 0 and 23.8% for HG type 1.2.-, while 'Bossier' had FIs of 100.8 and 78.5% for HG type 0 and HG type 1.2.-, respectively. To investigate the difference of two parents in response to SCN, the SCN growth and development were observed at three time points: 2, 4, and 8 DAI. As indicated in Fig. 2.1, no difference was observed between two parents at the J2 stage inside the roots at 2 DAI. J2s were able to penetrate both parents, suggesting that PI 567295 did not affect the penetration of SCN in soybean roots. However, the difference in nematode development was observed at 4 DAI in the roots between two parents. A delayed nematode development that was at the J2 stage was observed in the roots of PI 567295, while it was at the J3 stage in the 'Bossier' roots. A degenerating syncytium was observed around the head of J2 SCN in PI 567295's roots. At 8 DAI, normal development of SCN in 'Bossier' roots and clear necrosis in PI 567295 roots were observed (Fig. 2.1).

To prevent SCN damage, resistant plants cause necrosis around syncytia which is the only nutrient source for SCN development (Mitchum, 2016). Based on our observation of nematode development, PI 567295 showed slower degeneration of the syncytium at 4 DAI. It means that a resistant plant allowed the initiation of a syncytium, but its expansion was restricted by formation of a layer of necrotic cells around the feeding site. Besides, male SCN was observed in the PI 567295 roots at 8 DAI. It indicated that at a later stage, the layer of necrosis was formed to disconnect the syncytium from a vascular cylinder, therefore the female development was arrested due to starvation.

Phenotypes of the $F_{2:3}$ and RIL populations

Two parents and 120 $F_{2:3}$ families were evaluated with HG type 0 with two replicates. The FI data showed a variation of 0 to 185% for the first replicate and 0 to 172.4% for the second replicate. Based on the phenotype scores of each replicate, two bulks: resistant and susceptible bulks were created for each replicate due to genetic segregation in the $F_{2:3}$ population. The resistant bulks included 20 families with FI < 10% per rep and the susceptible bulks consisted of 20 families with FI > 60% per rep. Four bulks and two parents were used to find putative genomic regions controlling for HG type 0 resistance.

$F_{5:6}$ RIL showed a large variation at both SCN populations with a range of 0.4 to 144.8% for HG type 0 and 0.5 to 96.1% for HG type 1.2.-. Of 144 RILs evaluated for HG type 0, 15 RILs were rated as highly resistant (FI < 10%); 24 rated as moderately resistant (10% < FI < 30%); 28 rated as moderately susceptible (30% < FI < 60%); and 131 rated as highly susceptible (FI > 60%). Compared to the HG type 0, HG type 1.2.- is a more aggressive population. Five RILs were rated as highly resistant, 30 rated as moderately resistant, 34 rated as moderately susceptible, and 75 rated as highly susceptible. Genotype, test set, and interaction of genotype × test set were statistically significant ($P < 0.05$) for both SCN populations. Combining both phenotyping scores with two SCN populations, three $F_{5:6}$ RILs showed highly resistant to both SCN populations with FI < 10%. The highest resistant RIL, G19-SPT1748 was selected to develop a backcrossing population to introgress the target alleles to an elite recurrent parent, 'Bossier' and to confirm the QTLs.

Bulked segregant analysis

BSA was conducted to preliminarily determine the putative genomic regions controlling SCN resistance. BSA was performed in the $F_{2:3}$ 'Bossier' × PI 567295 population by selecting

homozygous SNPs that were polymorphic between 'Bossier' and PI 567295. The genomic regions consisting of at least 10 informative SNP markers were considered to be associated with a putative resistance locus. Based on these criteria, six genomic regions consisting of a total of 159 informative SNPs on Chrs 2, 5, 10, 11, 14, and 19 were identified as putative QTL regions for further mapping work. Of these 159 SNPs, 58 SNPs were identified from a genomic region on Chr 10 in the interval of SNP markers ss715606718 (39,310,253 bp) and ss715607341 (44,146,772 bp) (Wm82.a2 reference genome). Number of informative SNPs on each chromosome was presented in Table 2.3. All SNP marker coordinates, and physical positions were defined based on the Wm82.a2 genome sequence. To confirm and map the QTLs, 70 KASP markers on six chromosomes were designed for QTL mapping.

QTLs for SCN resistance

A total of 1,734 polymorphic SNPs from SoySNP6K BeadChip and 70 KASP markers designed from the BSA regions were used to construct a genetic linkage map for QTL analysis (Table 2.4). The linkage map spanned approximately 2,165 cM across 20 chromosomes with an average distance of 1.2 cM between SNPs. The largest number of SNPs (138) was placed on Chr 2, while Chr 4 only had 44 markers. The SNP positions on the linkage map were generally in agreement with the physical positions on the Wm82.a2 reference genome with some SNPs showing the differences between linkage map and physical orders.

Based on a genome-wide permutation test, a LOD threshold of 4.6 and 4.4 (permutation of 1,000 with $\alpha = 0.05$) were used to identify QTLs that are significant for HG type 0 and HG type 1.2.- resistance in the RIL population, respectively (Table 2.5 and Figs. 2.2 and 2.3). CIM analysis identified one major QTL on Chr 10 for HG type 0 resistance, which explained 53.7% of the phenotypic variation. Seven significant SNPs detected in the interval on Chr 10 had LOD

scores ranging from 15.0 to 21.8. Of these SNPs, GSM825 is the most significant SNP marker for the QTL resistant to HG type 0 with a LOD score of 21.8 (Fig. 2.2 A). The favorable allele (G) of this SNP is inherited from the resistant parent, PI 567295 with an average BLUP value of -27.5 (Fig. 2.2 B). Based on the confidence interval defined by the Bayesian credible interval method in R/QTL, the QTL position was located between GSM825 (ss715607225; 42,959,979 bp) and Gm10_42845506 (ss 715607271; 43,406,182 bp) on Chr 10 with an interval of 0.6 cM (Fig. 2.2 A; Table 2.5).

Based on a LOD threshold of 4.4, a significant genomic region consisting of seven significant SNPs on Chr 10 was also identified for resistance to HG type 1.2.-. Same to the resistance to HG type 0, the SNP marker GSM825 had the highest LOD score of 18.3, which explained 39.3% of phenotypic variation for HG type 1.2.- resistance (Fig. 2.2 A). Using a Bayesian credible interval method in R/QTL, the confidence interval of this QTL was same to the one for HG type 0 resistance on Chr 10 (Fig. 2.2 A; Table 2.5). The favorable allele of GSM825 is also from PI 567295. The average BLUP values of the lines carrying the favorable allele is -13.9, while the average BLUP values of lines carrying the unfavorable allele is 11.4 (Fig. 2.2 C).

In addition to the major QTL on Chr 10, two minor QTLs were also identified for resistance to HG type 1.2.- on Chrs 6 and 8, respectively. The Chr 6 QTL was located in the interval of Gm06_7323345 to Gm06_6797764 and explained 13% of phenotypic variation. The most significant marker Gm06_6914278 in the interval has a LOD score of 9.2 (Fig. 2.3 A). Lines carrying favorable alleles from PI 567295 had a BLUP value of -7.5, which is significantly different from those carrying unfavorable alleles ($P = 0.001$) (Fig. 2.3 B)

The Chr 8 QTL for resistance to HG type 1.2.- was identified in the interval from Gm08_17399376 to Gm08_18225805, spanning 1.9 cM and the most significant marker was Gm08_18225805 that explained 6.2% of phenotypic variation (Fig. 2.3 C). This QTL is located 8.7 Mb away from *Rhg4* locus. The difference of the BLUP values between two genotype groups at Gm08_18225805 was illustrated in Fig. 2.3 D. All QTLs had all favorable alleles from the resistant parent PI 567295. This negative effect indicated that the RILs carrying the alleles from PI 567295 with low or negative BLUP values provided better resistance than those carrying QTL alleles from 'Bossier' with higher BLUP values.

QTL mapping showed that no genomic region overlapped with *Rhg1* and *Rhg4* loci on Chrs 18 and 8, suggesting that PI 567295 has a different genetic background of resistance from two most common sources: 'Peking' and PI 88788. To understand the interaction of these three QTLs for resistance to HG type 1.2.-., the effect of these QTLs was tested by classifying them into eight groups: 1) QTLs on Chrs 6, 8, and 10; 2) QTLs on Chrs 6 and 10; 3) QTLs on Chrs 8 and 10; 4) QTLs on Chrs 6 and 8; 5) QTL on Chr 6 only; 6) QTL on Chr 8 only; 7) QTL on Chr 10 only; 8) absence of QTLs on these three chromosomes (Fig. 2.4). The average BLUP value of RILs carrying resistance alleles at the QTL on Chr 10 was 4.5. In comparison, the RILs possessing two QTLs on Chr 10 and 6 or 8 had average BLUP values of -30.2 and -27.7 respectively. Combining these three QTLs had further reduced the average BLUP value to -33.3 while lines with absence of these three QTLs had an average BLUP value of 16.5. The QTL interaction was tested by using an ANOVA analysis and Tukey HSD comparison. Combining the major QTL on Chr 10 with other minor QTLs had significantly reduced BLUP values when comparing with single QTL alone. With absence of the QTL on Chr 10, the combination of QTLs from Chrs 6 and 8 did not provide adequate resistance (average BLUP value of 5.03).

Therefore, the QTL on Chr 10 was required in PI 567295 resistance. In addition, the lines carrying only two minor QTLs without the major QTL on Chr 10 did not show a significant difference in BLUP value from the lines with absence of these three QTLs (Table 2.6). The results indicated that top five RILs that were highly resistant to HG type 1.2.- carried all three QTLs on Chrs 6, 8, and 10 with a BLUP value of less than -40.5. and three of these five RILs were also highly resistant to HG type 0 with a BLUP value of less than -53.0.

Candidate gene analysis and development of KASP markers on Chr 10 for HG type 0 and 1.2.- resistance

To identify candidate genes for SCN resistance, whole genome resequencing was performed with four genotypes including two parents, the most resistant RIL G19-SPT1748 and the most susceptible RIL G19-SPT1719. Based on QTL mapping results, the RIL G19-SPT1748 carries resistance alleles while RIL G19-SPT1719 has susceptible alleles at three QTLs on Chrs 6, 8, and 10. Sequence analysis identified 3,696,678 SNPs and 831,752 Indels across 20 chromosomes among four genotypes.

In this study, a major QTL on Chr 10 was mapped for both SCN populations: HG type 0 and HG type 1.2.- using 144 RILs. The genomic region associated with HG type 0 resistance was in the interval of GSM825 (ss715607225; 42,959,979 bp) to Gm10_42845506 (ss 715607271; 43,406,182 bp) for both HG types. Using the flanking region of 50 kb from both sides, candidate genes from 42,909,979 to 43,456,182 bp, spanning 546 kb were investigated. Using WGRS data, 18 gene models where SNPs and Indels were identified to be inside exons were detected based on the SoyBase database (Table 2.7). A genomic DNA sequence comparison between susceptible (‘Bossier’ and G19-SPT1719, a susceptible RIL) and resistant (PI 567295 and G19-SPT1748, a resistant RIL) groups identified 30 SNPs and three Indels within the exons of 18

gene models, resulting in changes in amino acids. These models had gene ontology terms including protein, nucleic acid, and iron ion binding as well as regulation of signal transduction, intracellular protein transport, biosynthetic process, protein kinase activity, and carbohydrate metabolism. Of 18 gene models, nine had root specific expression in SoyBase including *Glyma.10g198000, Glyma.10g198300, Glyma.10g198600, Glyma.10g199300, Glyma.10g200000, Glyma.10g200200, Glyma.10g200600, Glyma.10g200700, and Glyma.10g201900*.

Based on previous studies related to SCN resistance, the *GmSNAP18* at *Rhg1* locus encoded α-SNAP protein that interacts with the protein NSF to disrupt vesicle trafficking (Bayless et al., 2016). Therefore, gene models related to intracellular transport might be considered as candidate genes for SCN resistance. Of these 18 gene models, two gene models: *Glyma.10g199300* (43,021,060 to 43,026,276 bp) and *Glyma.10g201900* (43,302114 to 43,309,900 bp) were involved in vesicular transport. Based on the SoyBase expression database, both gene models *Glyma.10g199300* and *Glyma.10g201900* were related to intracellular transportation and had root specific expression. Seven variants were located within two exons of *Glyma.10g199300,* encoding α subunit Coat Protein I (COPI), which was involved in many cellular processes including intracellular protein transport, retrograde vesicle-mediated transport (soybase.org). Two of these seven variants were selected to design KASP markers GSM1021 and GSM1022 due to the change in amino acids. The GSM1021 is located within the 2nd exon with the favorable allele from resistant parent (C allele) and amino acid change from Threonine (T) to Alanine (A). Another KASP marker, GSM1022 resided in the 3rd exon of *Glyma.10g199300* with the amino acid change from Valine (V) to Isoleucine (I). The gene model, *Glyma.10g201900,* encoding Vps51/ Vps67 protein, a component of vesicular transport

carried three SNPs within their two exons that caused the change in amino acids including $C^{283}T$ ($A^{45}V$) in the 1st exon; and $C^{922}T$ ($V^{258}A$); $A^{2287}G$ ($K^{668}R$) in 2nd exon. Two KASP markers (GSM1018 and GSM1019), corresponding to two variants in the 2nd exon were designed; no marker was designed for the variant in the 1st exon due to non-specificity of the flanking region.

Three gene models: *Glyma.10g201800, Glyma.10g198500* and *Glyma.10g198600,* residing in the vicinity of the most significant marker GSM825 were also investigated. The gene model *Glyma.10g201800,* has two domains: the RING/U-box and TRAF. The RING/U box was involved in DNA replication (Jia et al., 2011), while the TRAF domain was involved in signal transduction (Fu et al., 2018). The gene model had a biological function of multicellular development and ubiquitin-dependent protein catabolic process (Soybase.org). Based on a comparison between DNA sequences of susceptible and resistant groups, a SNP in the 7th exon ($C^{1714}T$) changed amino acid ($M^{516}E$). The KASP marker GSM1020 was designed to validate this variant. The *Glyma.10g198500* encoded histone superfamily protein and was involved in nucleosome assembly (Soybase.org). Two SNPs including ss715607224 ($C^{136}T$ and $A^{388}G$) located in the 1st exon caused amino acid change in this gene model. Two KASP markers: GSM1024 and GSM1025, corresponding to $C^{136}T$ and $A^{388}G$, respectively, were designed. The 3rd gene model, *Glyma.10g198600* that encoded the RING/ U-box protein family, related to signal transduction. This protein family has been reported in responding to pathogen attack including U-box spotted leaf 11 to rice blast disease in rice (Zeng et al., 2004) and RING finger protein RMP1-interacting protein 2 and 3 to in *Arabidopsis* (Kawasaki et al., 2005). One variant in the first exon is a nonsynonymous SNP with an amino acid change from Leucine (L) to Valine (V). The KASP marker, GSM1023, corresponding to SNP ID ss715607231 in SoySNP50K Infinium Chip was designed for this variant.

All eight KASP markers from five candidate genes on Chr 10 were used to genotype the RIL population and analyze for the association between phenotypes and genotypes using a single marker analysis method. The results indicated that all these markers were significantly associated with resistance to both HG types with *P <0.001* (Table S2.2). GSM1021 was the most significant marker with *P <1.4e-14* and *2.2e-16* for HG types 1.2.- and 0, respectively. These eight KASP markers were also included with the marker set from SoySNP6K Infinium BeadChip for further QTL analysis. After quality control, 1,745 markers were used for linkage and QTL analysis (Fig. S2.2). Using a CIM method with a permutation of 1,000 and $\alpha = 0.001$, the marker GSM1021, located in the 2nd exon of *Glyma.10g199300* was most significant with LOD scores of 23.6 and 20.3 for HG type 0 and 1.2.-, respectively (Fig. 2.5 A). Favorable allele for resistance to both HG types at this marker locus was from the resistant parent, PI 567295. Comparing the BLUP values of two genotype groups at this locus revealed a significant difference with *P* values of *3.59e-26* and *4.95e-25*, respectively (Fig. 2.5 B & C).

Candidate gene analysis and development of KASP markers on Chrs 6 and 8 for HG type 1.2.-

Two minor QTLs on Chrs 6 and 8 were identified for HG type 1.2.- resistance. The genomic region on Chr 6 was flanked by Gm06_6797764 (ss715595458, 6,803,119 bp) and Gm06_7323345 (ss71559550, 7,329,187 bp). Candidate genes were reported in the region from 6,753,119 bp to 7,379,187, spanning 576 kb including 50 kb flanking of both sides (Table 2.8). Genomic comparison between resistant and susceptible groups identified four genes (*Glyma.06g090300, Glyma.06g091100, Glyma.06g092600, and Glyma.06g093000)* that carry five SNPs within exons, causing the changes in five amino acids. Based on the SoyBase database, these genes had biological functions including cellular lipid metabolic process, transmembrane transport, nucleic acid phosphodiester bond hydrolysis, and glycolytic process.

Interestingly, the flanking marker, Gm06_7323345 (ss71559550, 7,329,187 bp) with LOD score of 6.0 associated with resistance to HG type 1.2.- on Chr 6 was located within the 1st exon of *Glyma.06g093000,* causing an amino acid change. Therefore, the gene model *Glyma.06g093000* involve in gluconeogenesis could be responsible for resistance to HG type 1.2.-. Additionally, the *Glyma.06g091100* encoded a MATE efflux family protein that was known to play a role in transmembrane transport. Based on a genome wide analysis of MATE transporters in soybean (Liu et al., 2016), the *Glyma.06g09110* was grouped with three known MATE proteins including *A. thalian activated disease susceptibility 1 (ATADS1),* which negatively regulated the accumulation of salicylic acid, a plant immune system activator responding to pathogen attack (Sun et al., 2011). Four SNP markers located within exons of three candidate genes (*Glyma.06g090300, Glyma.06g091100, Glyma.06g092600)* on Chr 6 were selected to design KASP markers. Single marker analysis indicated that two SNPs GSM1028 and GSM1029 were significantly associated with resistance to HG type 1.2.- (*P = 0.02*) (Table 2.9). These SNPs were located within the 1st and 10th exon of *Glyma.06g90300*. Other two markers within coding regions of *Glyma.06g091100* and *Glyma.06g092600* did not show a significant contribution to SCN resistance (*P = 0.04* and *0.06*, respectively). Based on a haplotype analysis, GSM1028, GSM1029, and GSM1030 were all located within the same haplotype block with the most significant marker, *Gm06_6914278* (Fig. S2.3).

The QTL was identified in the genomic region between Gm08_17399376 (ss715599932; 17,333,566 bp) to Gm08_18225805 (ss715600030; 18,160,728 bp) on Chr 8. Whole genome resequencing analysis identified 11 SNPs and Indels within exons of six gene models in the region of 827 kb (Table 2.9). Of six gene models, three SNPs and one INDELs were identified within two exons of *Glyma.08g207600,* encoding UDP-Glycosyltransferase superfamily protein,

which performed glycosylation involving in biosynthetic pathway for plant defensive compounds such as phenolics, silicates, and salicylates (Rehman et al., 2018). Seven KASP SNP markers were designed to track six candidate genes. Of these seven markers, GSM1039 was the most significant marker associated with resistance to HG type 1.2.- with a *P* value of 0.0003 based on single marker analysis (Table 2.10). This SNP was located within the 6th exon of *Glyma.08g210000*, encoding protein binding. The GSM1040 was the second significant marker with a *P* value of 0.004 and this SNP was located within the 1st exon of *Glyma.08g207600*, encoding glycosyltransferase. Based on haplotype analysis, four KASP SNP markers: GSM1039, GSM1040, GSM1041, and GSM1042 all were located in same haplotype block with the significant marker Gm08_17303109 of the QTL on Chr 8 with a LOD score of 6 after adding new KASP markers (Fig. S2.4). While GSM1039 was located in the 6th exon of *Glyma.08210000*, three SNPs: GSM1040, GSM1041, and GSM1042 are within a coding region of *Glyma.08207600*. This result showed two genes might contribute together to SCN resistance on Chr 8.

Discussion

The objective of this project was to find genomic regions controlling SCN resistance in PI 567295 to provide additional sources of resistance while two main sources of resistance: 'Peking' and PI 88788 have been extensively exploited. Based on our previous study, PI 567295 did not carry the same resistance allele as 'Peking' and PI 88788 at the *Rhg1* and *Rhg4* loci and it had FIs of 12 and 23.8% for HG type 0 and HG type 1.2.-, respectively.

In this study, bulked segregant analysis (BSA) technique with SoySNP50K Infinium BeadChip was used to survey putative QTLs related to HG type 0 resistance in an $F_{2:3}$ population derived from a cross of 'Bossier' × PI 567295. The major advantage of this method is a simple

approach to quickly identify the putative genomic regions with two bulked DNA samples that were formed from the progenies with contrasting phenotypes: resistant and susceptible responses; and genotyped with a high-density SNP array. Based on BSA, six putative genomic regions on Chrs 2, 5, 10, 11, 14, and 19 were identified for HG type 0 resistance. Of these six regions, 58 informative SNPs were located within a region on Chr 10. QTL mapping with $F_{5:6}$ RIL population derived from the same parents for HG type 0 resistance was used to confirm these genomic regions using SoySNP6K Infinium BeadChip, which is a subset of SoySNP50K Infinium BeadChip. QTL mapping could only confirm the QTL on Chr 10, overlapping with the BSA result using a CIM approach. The QTL on Chr 10 was considered a major QTL with an LOD score of 21.8 for HG Type 0 resistance and explained 53.7% of phenotypic variation. It demonstrated that BSA was an effective method to target a QTL with a relatively large effect. Besides, the limitation of the BSA technique depends strongly on selection of highly contrasting phenotypes to create bulks so that the $F_{2:3}$ population with large segregation could affect the results. In this study, multiple resistant and susceptible bulks were created from each replication of the $F_{2:3}$ population and 20 informative SNPs on Chr 10 were found in each of the replications. Therefore, multiple bulks are recommended for constructing BSA when using an $F_{2:3}$ population in order to verify genomic regions. Additionally, since BSA was used in a subset of individuals, it could cause false hits or not capture the minor QTLs. Therefore, BSA should be used as a preliminary survey to identify putative genomic regions and the QTL mapping should follow up to confirm.

QTLs for resistance to HG type 0 or HG type 1.2.-

HG type 0 (SCN race 3) is the most common SCN population, and HG type 1.2.- (SCN race 2) is more aggressive than HG type 0 because both ‘Peking’ and PI 88788, the most

common sources of resistance, have shown susceptible responses to HG type 1.2.- (Table 2.1). Both *Rhg1 and Rhg4* loci have been repeatedly mapped as major QTLs in different genetic backgrounds for resistance to HG type 0 (race 3). They have also been mapped for resistance to HG type 1.2.- and other QTLs are required with them to provide full resistance to HG type 1.2.-. For example, *Rhg1* and *Rhg4* were detected in PI 438489 B for resistance to HG type 1.2.-, but they only explained 12.8 and 7.4% of phenotypic variation (Yue et al., 2001a). Other QTLs on Chrs 4, 11, and 15 were mapped with *Rhg1* and Rhg4 loci for resistance to HG type 1.2.- in PI 438489 B (Yue et al., 2001a).

QTL mapping using the $F_{5:6}$ RIL population identified a genomic region on Chr 10 for both SCN populations using the CIM method. The GSM 825 is the most significant marker for both SCN populations but it does not locate within any genes, so it is considered as a flanking marker for marker assisted selection. To be more accurate, candidate genes should be identified and polymorphic SNPs within the gene could be developed as functional markers for selection. Therefore, eight KASP SNP markers that are located within the exons of five candidate genes on Chr 10 were designed to evaluate candidate genes and were included in QTL analysis. GSM1021, located within *Glyma.10g199300*, was the most significant marker for both HG types with LOD score of 23.6 and 20.3 for HG type 0 and 1.2.-, respectively.

Previous studies indicated multiple QTLs were required for full resistance to HG type 1.2.- (Yue et al., 2001a; Yue et al., 2001b; Jiao et al., 2015b). In this study, two minor QTLs on Chrs 6 and 8 were identified for resistance to HG type 1.2.- in PI 567295. On Chr 8, a 827 kb region was located between Gm08_17399376 (17,333,566 bp) and Gm08_18225805 (18,160,728 bp) with an LOD score of 5.02, which was 8.7 Mb away from *Rhg4* locus. According to SoyBase (soybase.org), our mapped QTL overlapped with QTLs SCN 37-4 and

SCN 50-2. The SCN 37-4 was identified in PI 567516 C for resistance to HG type 1.2.5.7 and located between Satt233 (17,232,368 bp) and Sat_040 (39,682,787 bp) with an LOD value of 4.3 (Vuong et al., 2010). The QTL located on Chr 6 and identified in this study for resistance to HG type 1.2.- does not overlap any reported QTLs on Chr 6 according to SoyBase (Soybase.org).

Candidate genes and KASP markers for SCN resistance on Chr 10

After attacking soybean roots, survival of cyst nematode depends on the feeding site called syncytium which is formed through cell wall dissolution and neighboring cells because syncytium is the only source of nutrients (Mitchum, 2016). In resistant plants, the cyst nematode is unable to establish or maintain this feeding site. In soybeans, the genes responsible for cyst nematode are identified and well characterized for only major genes at *Rhg1* and *Rhg4* loci with two main molecular mechanisms of SCN resistance: 1) disruption of SNARE complexes and vesicle transportation by accumulation of resistant α-SNAP protein at the syncytium (Bayless et al., 2016); and 2) alteration of folate homeostasis which may lead hypersensitive response at feeding site (Liu et al., 2012). Within the fine mapped region on Chr 10, no *SNAP* or *SHMT* family genes were found, suggesting PI 567295 might have a different mechanism of resistance.

WGRS of two parents, and a highly resistant and susceptible RILs revealed 18 gene models that carried polymorphic SNPs or deletions in coding regions. Of the 18 genes, five gene models including *Glyma.10g199300, Glyma.10g198500; Glyma.10g198600, Glyma.201800* and *Glyma.201900* were considered as candidate genes based on SoyBase database and Arabidopsis Information Resource (TAIR) (arabidopsis.org). The gene model, *Glyma.10g201900* encoding Vps 51/ Vps 67 family, is a component of vesicular transport. Vesicular trafficking is a complicated process including SNAREs, an integral membrane protein. The interaction of *NSF* gene on Chr 7 and *α-SNAP* gene at Chr 18 affects SNARE complexes, then disrupting vesicle

fusion and related to SCN resistance (Bayless et al., 2016). In yeast, VFT (Vps fifty-three) complexes including Vps51 might affect SNARE activity by Vps51's ability to bind via a short N terminal peptide, which competes for the same binding site with SNARE (Siniossoglou and Pelham, 2002). The KASP SNP marker, GSM1019, located in the 2nd exon of this gene, *Glyma.10g201900* was also significantly associated with resistance to HG type 1.2.- (LOD = 7.6) and HG type 0 (LOD = 5.0 This SNP is non-synonymous, leading an amino acid change from Lysine (K) to Arginine (R). This marker was located within the 2nd exon of *Glyma.10g201900.* The average of BLUP values of RILs with favorable allele, (C from PI 567295) were -11.9 and -23.2 for HG type 1.2.- and 0, respectively.

Another molecular mechanism of SCN resistance could be hormone signaling including salicylic acid (SA) and jasmonic acid (JA) which could trigger a resistance response (Kandoth et al., 2011; Matthews et al., 2014). QTL mapping identified a region of ~330 kb on Chr 10 including 42 gene models. The *Glyma.10g199300* encoding the α subunit coatomer that is a member of Coat Protein I complex related to protein transport between the endoplasmic reticulum and Golgi compartments. In *Arabidopsis,* α subunit coatomer was identified as SINAT1 (Seven in Absentia *of Arabidopsis thaliana*) interactors that is one of E3 ubiquitin ligases, functioning in hormone signaling (Xia et al., 2020). Ubiquitination related components have been identified that were involved in plant pathogen or plant insect interactions. For example, a RING E3 ligase was reported for plant response to powdery mildew invasion, probably in the initial signaling process to pathogen infection (Tör et al., 2002). The SNP marker GSM1021 contributed most to the resistance to HG type 0 and 1.2.- (LOD score > 20). This SNP locates in the 2nd exon of *Glyma.10g199300* and caused an amino acid change from Threonine

(T) to Alanine (A), Based on these results, these markers: GSM1019 and GSM1021 can be used in effective marker-assisted selection of the major QTL on Chr 10 for SCN resistance.

Interestingly, of five candidate genes, *Glyma.10g201800* encoded a protein that is a member of the Seven in Absentia homolog (SIAH) family (soybase.org). The protein is also an E3 ligase and involved in ubiquitination. In *Arabidopsis*, the α coatomer protein interacted with SINAT1 in regulation of abscisic acid; (ABA signaling) (Xia et al., 2020), therefore, the SIAH and α coatomer protein encoded by *Glyma.10g199300* might interact with hormone signaling including ABA, influencing plant immunity. Some studies detected ABA enhanced rice's susceptibility to *Magnaporthe grisea*, rice blast fungus (Jiang et al., 2010). Nahar et al. (2012) showed that ABA negatively affected *Hirschmanniella oryzae,* the migratory nematode, increasing the resistance response of rice. The KASP marker, GSM1020 that was located within the 7th exon of *Glyma.10g201800* was significantly associated with the resistance to both HG type 0 and 1.2.-, respectively, with $P < 1.5e\text{-}15$ and *4.2e-11* (ranked 3rd after GSM1019 and GSM1021) based on single marker analysis (Table S2.2).

Lastly, the *Glyma.10g198500*, identified near GSM825 is in the Histone H2B protein family. Histones play a role in chromatin structure and DNA replication. In plants, histone modification affects DNA methylation (Zhao et al., 2019). Each histone family often encodes related but functionally distinct proteins, which are referred to as "histone variants" during evolution (Talbert and Henikoff, 2017). Most studies in *Arabidopsis* related to H2B chemical modification focused on mono-ubiquitination that is involved in the regulation of various developmental processes (Fleury et al., 2007) in plants' defenses against fungi, (Dhawan et al., 2009) and in controlling the expression of disease resistance genes (Zou et al., 2014). Single marker analysis indicated two KASP markers: GSM1024 and GSM1025 that are located within

the 1st exon of *Glyma.10g198500* were significantly associated with resistance to both HG types (Table S2.2)

Comparison of QTLs on Chr 10

Based on the information obtained from SoyBase, four QTLs on Chr 10 associated with SCN resistance have been reported previously. However, our QTL on Chr 10 in this study does not overlap with these four QTLs. The first QTL was flanked by *Satt466* (29,480,265 bp) and *Satt173 (*36,987,008 bp). Using an $F_{4:5}$ RIL population derived from PI 464925B (*G. soja*), Winter et al. (2007) reported this QTL with a LOD score of 2.0, explaining 5.1% of phenotypic variation. This QTL was responsible for resistance to the Ruthven isolate of *H. glycines*, which was collected in Canada, characterized as HG type 1.2.- (race 2). The second QTL (*Qscn14-2*) on Chr 10 was identified in L-10 for HG type 0 resistance. Using an $F_{5:7}$ RIL population derived from Heinong 37 × L-10, the *Qscn14-2* was mapped between *Satt550* (25,055,615 bp) and *Sat_282* (38,033,347 bp) with an LOD score of 2.3 that overlapped with the first reported QTL (Chang et al., 2011). The third QTL (*Qscn14-1*) on Chr 10 was from the same study, located in the interval of *Sat_145* (11,017,809 bp) to *Satt345* (12,355,600 bp) with an LOD score of 2.5 (Chang et al., 2011). Both QTLs (*Qscn14-1* and *Qscn14-2*) were associated with HG type 3.- (race 14) resistance, explaining 5.0-8.8% of the phenotypic variation. The last reported QTL, *qSCN10,* flanked by Satt592 (43,509,732 bp) and Sat_038 (46,052,341 bp) on Chr 10 was associated with multiple SCN populations (HG type 0; 2.5.7; 2.7; 1.4.5.6.7 and LY1) and derived from PI 567516 C (Vuong et al., 2010). Later, the *qSCN10* was fine mapped to a 379 kb (42,430,173 – 42,809,800 bp), (Zhou et al., 2021), which was located 150 kb away from our QTL position on Chr 10. Of four reported QTLs, only *qSCN10* was closely located to our mapped QTL on Chr 10 while the other three QTLs were located at least 5 Mb away. The

difference between *qSCN10* and our mapped QTL could be due to the different genetic backgrounds or variation of the SCN populations used in the studies.

Implication of Breeding for SCN resistance

SCN has been the most destructive pest for soybean production and has the ability of adaptation to current resistant sources. PI 88788-type resistance has become less effective due to overexploitation in resistant cultivars (Mitchum, 2016). It is critical to breed resistant cultivars with diverse resistant sources or new resistance mechanisms. Screening soybean accessions for SCN resistance provided additional sources for SCN resistance. With our designed KASP markers for genotyping resistance alleles at *Rhg1* and *Rhg4* loci, novel SCN resistance sources that do not carry the same resistance alleles as 'Peking' or PI 88788 have been discovered. PI 567295 is a MG VIII accession and has resistance to HG type 0 and 1.2.-. PI 567295, originally from China has a brown seed coat color, and moderate resistance to lodging and shattering.

A QTL with a large effect on resistance to both HG type 0 and HG type 1.2.- was identified on Chr 10 in this study. All favorable alleles were from PI 567295, accounting for 39.3 to 53.7% of phenotypic variation for HG type 0 and 1.2.-, respectively. These favorable alleles could be introgressed into elite lines to improve HG type 1.2.- resistance. The KASP SNP marker, GSM1021 was shown to have significant association with resistance to both SCN populations based on both single marker analysis and QTL analysis. This SNP could serve as a functional marker for selection of this QTL. Pyramiding favorable alleles of QTLs have been shown to enhance SCN resistance. Soybean accessions carrying *GmSNAP11, rhg1-a* and *Rhg4* alleles had an FI of 1.9%, and the accessions carrying resistance alleles at *GmSNAP11* and *rhg1-a* had a FI of 12.4% (Tian et al., 2019). The *GmSNAP11* gene on Chr 11 belongs to the *SNAP* gene family with *GmSNAP18* gene at *Rhg1* locus. Previous studies indicated that besides

GmSNAP18, the *GmSNAP11* gene had a minor effect on SCN resistance. In this study, single QTL even the major QTL on Chr 10 could not provide full resistance. It indicates the SCN resistance is complicated quantitative traits that required multiple QTLs for resistance. Combining this major QTL with a minor QTL on Chr 6 or 8 showed enhanced resistance response, significantly difference with lines carrying a single QTL. The best combination was all favorable alleles from all QTLs, reducing the BLUP value down to -33.3. Therefore, pyramiding resistance alleles at the QTL on Chr 10 with other QTLs in this study could provide broad resistance to SCN populations.

Conclusion

Soybean cyst nematode is the most devastating plant parasite affecting soybean production in the U.S. SCN is difficult to eradicate due to the structure of the cysts that protect a large number of eggs in the soil for a long time (Niblack et al., 2006). Using resistant cultivars and crop rotation are the most effective ways to SCN management. QTL mapping using a RIL population derived from 'Bossier' x PI 567295 in this study pinpointed the major QTL on Chr 10 with an interval of ~446 kb for resistance to two SCN populations HG type 0 and 1.2.-. Two minor QTLs with intervals of ~476 kb on Chr 6 and ~827 kb on Chr 8, respectively were significantly associated with HG type 1.2.- resistance. The favorable alleles came from the resistant parent, PI 567295. Of 144 $F_{5:6}$ RIL progenies, the G19-SPT1748 was rated as highly resistant to HG type 0 and 1.2.- with FIs of 0.4 and 0.5%, respectively. Whole genome resequencing and candidate gene analysis identified two gene models *Glyma.10g199300* and *Glyma.10g201900* that could contribute to SCN resistance. Soybean breeders also need combinations of resistance genes from different sources of resistance to provide broad resistance

to multiple SCN populations. QTLs associated with resistance to HG type 0 and 1.2.- could be pyramided with PI 88788-type resistance to enhance resistance.

References

Bandara, A.Y., Weerasooriya, D.K., Bradley, C.A., Allen, T.W., and Esker, P.D. (2020). Dissecting the economic impact of soybean diseases in the United States over two decades. *PLOS ONE* 15(4), e0231141. doi: 10.1371/journal.pone.0231141.

Bayless, A.M., Smith, J.M., Song, J., McMinn, P.H., Teillet, A., August, B.K., et al. (2016). Disease resistance through impairment of α-SNAP–NSF interaction and vesicular trafficking by soybean *Rhg1*. *Proc Natl Acad Sci.* 113(47), E7375. doi: 10.1073/pnas.1610150113.

Bayless, A.M., Zapotocny, R.W., Grunwald, D.J., Amundson, K.K., Diers, B.W., and Bent, A.F. (2018). An atypical N-ethylmaleimide sensitive factor enables the viability of nematode-resistant *Rhg1* soybeans. *Proc Natl Acad Sci.* 115(19), E4512. doi: 10.1073/pnas.1717070115.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30(15), 2114-2120. doi: 10.1093/bioinformatics/btu170.

Bradley, C., Allen, T., Tenuta, A., Mehl, K., and Sisson, A. (2021). *Soybean disease loss estimates from the United States and Ontario, Canada-2020* [Online]. Available: https://cropprotectionnetwork.org/resources/publications/soybean-disease-loss-estimates-from-the-united-states-and-ontario-canada-2020 [Accessed].

Broman, K.W., and Speed, T.P. (1999). A Review of Methods for Identifying QTLs in Experimental Crosses. *Lect. Notes Ser.* 33, 114-142. doi: 10.1214/lnms/1215455550

Broman, K. W., H. Wu, S. Sen, and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics*. 19, 889-890. doi: 10.1093/bioinformatics/btg112.

Brucker, E., Carlson, S., Wright, E., Niblack, T., and Diers, B. (2005). *Rhg1* alleles from soybean PI 437654 and PI 88788 respond differentially to isolates of *Heterodera glycines* in the greenhouse. *Theor Appl Genet.* 111(1), 44-49. doi: 10.1007/s00122-005-1970-3.

Chang, W., Dong, L., Wang, Z., Hu, H., Han, Y., Teng, W., et al. (2011). QTL underlying resistance to two HG types of *Heterodera glycines* found in soybean cultivar 'L-10'. *BMC Genet.* 12(1), 233. doi: 10.1186/1471-2164-12-233.

Chowdhury, I.A., Yan, G., Plaisance, A., and Markell, S. (2021). Characterization of virulence phenotypes of soybean cyst nematode (*Heterodera glycines*) Populations in North Dakota. *Phytopathology* 111(11), 2100-2109. doi: 10.1094/PHYTO-01-21-0031-R.

Concibido, V.C., Denny, R.L., Lange, D.A., Orf, J.H., and Young, N.D. (1996). RFLP Mapping and marker-assisted selection of soybean cyst nematode resistance in PI 209332. *Crop Sci.* 36(6). doi: 10.2135/cropsci1996.0011183X003600060038x.

Concibido, V.C., Lange, D.A., Denny, R.L., Orf, J.H., and Young, N.D. (1997). Genome mapping of soybean cyst nematode resistance genes in 'Peking', PI 90763, and PI 88788 using DNA markers. *Crop Sci.* 37(1). doi: 10.2135/cropsci1997.0011183X003700010046x.

Concibido, V.C., Diers, B.W., and Arelli, P.R. (2004). A decade of QTL mapping for cyst nematode resistance in soybean. *Crop Sci.* 44(4), 1121-1131. doi: 10.2135/cropsci2004.1121.

Cook, D.E., Bayless, A.M., Wang, K., Guo, X., Song, Q., Jiang, J., et al. (2014). Distinct copy number, coding sequence, and locus methylation patterns underlie *Rhg1*-mediated soybean resistance to soybean cyst nematode. *Plant Physiol.* 165(2), 630-647. doi: 10.1104/pp.114.235952.

Cook, D.E., Lee, T.G., Guo, X., Melito, S., Wang, K., Bayless, A.M., et al. (2012). Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338(6111), 1206. doi: 10.1126/science.1228746.

Dhawan, R., Luo, H., Foerster, A.M., Abuqamar, S., Du, H.N., Briggs, S.D., et al. (2009). Histone Monoubiquination1 interacts with a subunit of the mediator complex and regulates defense against necrotrophic fungal pathogens in *Arabidopsis*. *Plant Cell* 21(3), 1000-1019. doi: 10.1105/tpc.108.062364.

Fleury, D., Himanen, K., Cnops, G., Nelissen, H., Boccardi, T.M., Maere, S., et al. (2007). The *Arabidopsis thaliana* homolog of yeast BRE1 has a function in cell cycle regulation during early leaf and root growth. *Plant Cell* 19(2), 417-432. doi: 10.1105/tpc.106.041319.

Fu, T.-M., Shen, C., Li, Q., Zhang, P., and Wu, H. (2018). Mechanism of ubiquitin transfer promoted by TRAF6. *Proc Natl Acad Sci.* 115(8), 1783. doi: 10.1073/pnas.1721788115.

Gardner, M., Heinz, R., Wang, J., and Mitchum, M.G. (2017). Genetics and adaptation of soybean cyst nematode to broad spectrum soybean resistance. *G3-GENES GENOM GENET* 7(3), 835-841. doi: 10.1534/g3.116.035964.

Golden, A.M., and Et, A.L. (1970). Terminology and identity of infraspecific forms of the soybean cyst nematode (*Heterodera glycines)*. *Plant Dis.* 54, 544-546.

Glover, K.D., Wang, D., Arelli, P.R., Carlson, S.R., Cianzio, S.R., and Diers, B.W. (2004). Near isogenic lines confirm a soybean cyst nematode resistance gene from PI 88788 on linkage group J. *Crop Sci.* 44(3), 936-941. doi: 10.2135/cropsci2004.9360.

Holland, J.B., Nyquist, W.E., and Cervantes-Martínez, C.T. (2002). Estimating and interpreting heritability for plant breeding: An Update, in *Plant Breed. Rev*., 9-112. doi: 10.1002/9780470650202.ch2.

Howland, A., Monnig, N., Mathesius, J., Nathan, M., and Mitchum, M.G. (2018). Survey of *Heterodera glycines* population densities and virulence phenotypes during 2015–2016 in Missouri. *Plant Dis.* 102(12), 2407-2410. doi: 10.1094/PDIS-04-18-0650-SR.

Jia, L., Bickel, J.S., Wu, J., Morgan, M.A., Li, H., Yang, J., et al. (2011). RBX1 (RING box protein 1) E3 ubiquitin ligase is required for genomic integrity by modulating DNA replication licensing proteins. *The J. Biol. Chem.* 286(5), 3379-3386. doi: 10.1074/jbc.M110.188425.

Jiang, C.-J., Shimono, M., Sugano, S., Kojima, M., Yazawa, K., Yoshida, R., et al. (2010). Abscisic acid interacts antagonistically with salicylic acid signaling pathway in rice–magnaporthe grisea interaction. *Mol. Plant Microbe Interact.* 23(6), 791-798. doi: 10.1094/mpmi-23-6-0791.

Jiao, Y., Vuong, T.D., Liu, Y., Meinhardt, C., Liu, Y., Joshi, T., et al. (2015). Identification and evaluation of quantitative trait loci underlying resistance to multiple HG types of soybean cyst nematode in soybean PI 437655. *Theor Appl Genet.* 128(1), 15-23. doi: 10.1007/s00122-014-2409-5.

Kandoth, P.K., Ithal, N., Recknor, J., Maier, T., Nettleton, D., Baum, T.J., et al. (2011). The soybean *Rhg1* locus for resistance to the soybean cyst nematode *Heterodera glycines* regulates the expression of a large number of stress- and defense-related genes in degenerating feeding cells. *Plant Physiol.* 155(4), 1960-1975. doi: 10.1104/pp.110.167536.

Kazi, S., Shultz, J., Afzal, J., Hashmi, R., Jasim, M., Bond, J., 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(3), 633-644. doi: 10.1007/s00122-009-1181-4.

Kawasaki, T., Nam, J., Boyes, D.C., Holt III, B.F., Hubert, D.A., Wiig, A., et al. (2005). A duplicated pair of *Arabidopsis* RING-finger E3 ligases contribute to the RPM1- and RPS2-mediated hypersensitive response. *Plant J.* 44(2), 258-270. doi: 10.1111/j.1365-313X.2005.02525.x.

Kim, M., and Diers, B.W. (2013). Fine Mapping of the SCN Resistance QTL cqSCN-006 and cqSCN-007 from Glycine soja PI 468916. *Crop Sci.* 53(3), 775-785. doi: 10.2135/cropsci2012.07.0425.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14), 1754-1760. doi: 10.1093/bioinformatics/btp324.

Liu, J., Li, Y., Wang, W., Gai, J., and Li, Y. (2016). Genome-wide analysis of MATE transporters and expression patterns of a subgroup of MATE genes in response to aluminum toxicity in soybean. *BMC Genet.* 17(1), 223. doi: 10.1186/s12864-016-2559-8.

Liu, S., Kandoth, P.K., Lakhssassi, N., Kang, J., Colantonio, V., Heinz, R., et al. (2017). The soybean *GmSNAP18* gene underlies two types of resistance to soybean cyst nematode. *Nat. Commun.* 8(1), 14822. doi: 10.1038/ncomms14822.

Liu, S., Kandoth, P.K., Warren, S.D., Yeckel, G., Heinz, R., Alden, J., et al. (2012). A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492 (7428), 256-260. doi: 10.1038/nature11651.

Matthews, B.F., Beard, H., Brewer, E., Kabir, S., MacDonald, M.H., and Youssef, R.M. (2014). Arabidopsis genes, AtNPR1, AtTGA2 and AtPR-5, confer partial resistance to soybean cyst nematode (*Heterodera glycines*) when overexpressed in transgenic soybean roots. *BMC Plant Biol.* 14(1), 96. doi: 10.1186/1471-2229-14-96.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9), 1297-1303. doi: 10.1101/gr.107524.110.

Mitchum, M.G. (2016). Soybean Resistance to the Soybean Cyst Nematode *Heterodera glycines*: An Update. *Phytopathology* 106(12), 1444-1450. doi: 10.1094/phyto-06-16-0227-rvw.

Mudge, J., Cregan, P.B., Kenworthy, J.P., Kenworthy, W.J., Orf, J.H., and Young, N.D. (1997). Two Microsatellite Markers That Flank the Major Soybean Cyst Nematode Resistance Locus. *Crop Sci.* 37(5). doi: cropsci1997.0011183X003700050034x.

Nahar, K., Kyndt, T., Nzogela, Y.B., and Gheysen, G. (2012). Abscisic acid interacts antagonistically with classical defense pathways in rice–migratory nematode interaction. *New Phytol.* 196(3), 901-913. doi: 0.1111/j.1469-8137.2012.04310.x.

Niblack, T.L., Arelli, P.R., Noel, G.R., Opperman, C.H., Orf, J.H., Schmitt, D.P., et al. (2002). A Revised Classification Scheme for Genetically Diverse Populations of *Heterodera glycines. J. Nematol.* 34(4), 279-288.

Niblack, T.L., Lambert, K.N., and Tylka, G.L. (2006). A model plant pathogen from the kingdom animalia: *Heterodera glycines*, the soybean cyst nematode. *Annu. Rev. Phytopathol.* 44(1), 283-303. doi: 10.1146/annurev.phyto.43.040204.140218.

Niblack, T.L., Colgrove, A.L., Colgrove, K., and Bond, J.P. (2008). Shift in virulence of soybean cyst nematode is associated with use of resistance from PI 88788. *Plant Health Prog.* 9(1), 29. doi: 10.1094/PHP-2008-0118-01-RS.

Ooijen, J., Verlaat, J.v.t., Tol, J., Dalh n, J., Buren, J., Meer, J.M., et al. (2006). "JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations". Kyazma B.V., Wageningen.

Patil, G.B., Lakhssassi, N., Wan, J., Song, L., Zhou, Z., Klepadlo, M., et al. (2019). Whole-genome re-sequencing reveals the impact of the interaction of copy number variants of the *rhg1* and *Rhg4* genes on broad-based resistance to soybean cyst nematode. *Plant Biotechnol. J.* 17(8), 1595-1611. doi: 10.1111/pbi.13086.

Pham, A.T., McNally, K., Abdel-Haleem, H., Roger Boerma, H., and Li, Z. (2013). Fine mapping and identification of candidate genes controlling the resistance to southern root-knot nematode in PI 96354. *Theor Appl Genet.* 126(7), 1825-1838. doi: 10.1007/s00122-013-2095-8.

Rehman, H.M., Nawaz, M.A., Shah, Z.H., Ludwig-Müller, J., Chung, G., Ahmad, M.Q., et al. (2018). Comparative genomic and transcriptomic analyses of Family-1 UDP glycosyltransferase in three *Brassica* species and *Arabidopsis* indicates stress-responsive regulation. *Sci. Rep.* 8(1), 1875. doi: 10.1038/s41598-018-19535-3.

Schmitt, D.P., and Shannon, G. (1992). Differentiating soybean responses to *Heterodera glycines* races. *Crop Sci.* 32(1). doi: 10.2135/cropsci1992.0011183X003200010056x.

Shi, Z., Liu, S., Noe, J., Arelli, P., Meksem, K., and Li, Z. (2015). SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. *BMC Genet.* 16(1), 314-314. doi: 10.1186/s12864-015-1531-3.

Siniossoglou, S., and Pelham, H.R.B. (2002). Vps51p Links the VFT Complex to the SNARE Tlg1p*. *J. Biol. Chem.* 277(50), 48318-48324. doi: 10.1074/jbc.M209428200.

Sun, X., Gilroy, E.M., Chini, A., Nurmberg, P.L., Hein, I., Lacomme, C., et al. (2011). ADS1 encodes a MATE-transporter that negatively regulates plant disease resistance. *New Phytol.* 192(2), 471-482. doi: 10.1111/j.1469-8137.2011.03820.x.

Talbert, P.B., and Henikoff, S. (2017). Histone variants on the move: substrates for chromatin dynamics. *Nat. Rev* 18(2), 115-126. doi: 10.1038/nrm.2016.148.

Tian, Y., Liu, B., Shi, X., Reif, J.C., Guan, R., Li, Y.-h., et al. (2019). Deep genotyping of the gene *GmSNAP* facilitates pyramiding resistance to cyst nematode in soybean. *Crop J.* 7(5), 677-684. doi: 10.1016/j.cj.2019.04.003.

Tran, D.T., Steketee, C.J., Boehm, J.D., Noe, J., and Li, Z. (2019). Genome-wide association analysis pinpoints additional major genomic regions conferring resistance to soybean cyst nematode (*Heterodera glycine*s Ichinohe). *Front. Plant Sci.* 10(401). doi: 10.3389/fpls.2019.00401.

Tör, M., Gordon, P., Cuzick, A., Eulgem, T., Sinapidou, E., Mert-Türk, F., et al. (2002). *Arabidopsis* SGT1b is required for defense signaling conferred by several downy mildew resistance genes. *The Plant Cell* 14(5), 993-1003. doi: 10.1105/tpc.001123.

Tylka, G.L. (2020). SCN-resistant Soybean Varieties for Iowa - By the Numbers. *ICM News* 2660. doi: https://lib.dr.iastate.edu/cropnews/2660.

Tylka, G.L., and Marett, C.C. (2021). Known distribution of the soybean cyst nematode, *Heterodera glycines*, in the United States and Canada in 2020. *Plant Health Prog.* 22(1), 72-74. doi: 10.1094/php-10-20-0094-br.

Usovsky, M., Ye, H., Vuong, T., Patil, G., Wan, J., Zhou, L., et al. (2021). Fine-mapping and characterization of qSCN18, a novel QTL controlling soybean cyst nematode resistance in PI 567516C. *Theor Appl Genet.* 134, 1-11. doi: 10.1007/s00122-020-03718-6.

Voorrips, R.E. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93(1), 77-78. doi: 10.1093/jhered/93.1.77.

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

Xia, F.-N., Zeng, B., Liu, H.-S., Qi, H., Xie, L.-J., Yu, L.-J., et al. (2020). SINAT E3 Ubiquitin Ligases Mediate FREE1 and VPS23A Degradation to Modulate Abscisic Acid Signaling. *The Plant Cell* 32(10), 3290-3310. doi: 10.1105/tpc.20.00267.

Wang, J., Niblack, T.L., Tremain, J.A., Wiebold, W.J., Tylka, G.L., Marett, C.C., et al. (2003). Soybean cyst nematode reduces soybean yield without causing obvious aboveground symptoms. *Plant Dis* 87(6), 623-628. doi: 10.1094/pdis.2003.87.6.623.

Weisemann, J.M., Matthews, B.F., and Devine, T.E. (1992). Molecular markers located proximal to the soybean cyst nematode resistance gene, *Rhg4*. *Theor Appl Genet.* 85(2-3), 136-138. doi: 10.1007/bf00222850.

Winter, S.M., Shelp, B.J., Anderson, T.R., Welacky, T.W., and Rajcan, I. (2007). QTL associated with horizontal resistance to soybean cyst nematode in *Glycine soja* PI464925B. *Theor Appl Genet.* 114(3), 461-472. doi: 10.1007/s00122-006-0446-4.

Yu, N., Lee, T.G., Rosa, D.P., Hudson, M., and Diers, B.W. (2016). Impact of Rhg1 copy number, type, and interaction with *Rhg4* on resistance to *Heterodera glycines* in soybean. *Theor Appl Genet.* 129(12), 2403-2412. doi: 10.1007/s00122-016-2779-y.

Yue, P., Arelli, P.R., and Sleper, D.A. (2001a). Molecular characterization of resistance to *Heterodera glycines* in soybean PI 438489B. *Theor Appl Genet.* 102(6), 921-928. doi: 10.1007/s001220000453.

Yue, P., Sleper, D.A., and Arelli, P.R. (2001b). Mapping Resistance to Multiple Races of *Heterodera glycines* in Soybean PI 89772. *Crop Sci.* 41(5), 1589-1595. doi: 10.2135/cropsci2001.4151589x.

Zeng, L.R., Qu, S., Bordeos, A., Yang, C., Baraoidan, M., Yan, H., et al. (2004). Spotted leaf11, a negative regulator of plant cell death and defense, encodes a U-box/armadillo repeat protein endowed with E3 ubiquitin ligase activity. *Plant Cell* 16(10), 2795-2808. doi: 10.1105/tpc.104.025171.

Zhao, T., Zhan, Z., and Jiang, D. (2019). Histone modifications and their regulatory roles in plant development and environmental memory. *J Genet Genomics* 46(10), 467-476. doi: 10.1016/j.jgg.2019.09.005.

Zheng, J., and Chen, S. (2013). Estimation of virulence type and level of soybean cyst nematode field populations in response to resistant cultivars. *Int. J. Nematology Entomol*. 1111–116

Zhou, L., Song, L., Lian, Y., Ye, H., Usovsky, M., Wan, J., et al. (2021). Genetic characterization of qSCN10 from an exotic soybean accession PI 567516 C reveals a novel source conferring broad-spectrum resistance to soybean cyst nematode. *Theor Appl Genet.* 134(3), 859-874. doi: 10.1007/s00122-020-03736-4.

Zou, B., Yang, D.-L., Shi, Z., Dong, H., and Hua, J. (2014). Monoubiquitination of Histone 2B at the disease resistance gene locus regulates its expression and impacts immune responses in *Arabidopsis*. *Plant Physiol.* 165(1), 309-318. doi: 10.1104/pp.113.227801.

Tables and Figures

Table 2.1 HG type designation tests with soybean cyst nematode HG type 0 and 1.2.- populations in the University of Georgia Nematology Greenhouse.

Indicator number	Genotypes	HG type 0			HG type 1.2.-[a]		
		Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
1	'Peking'	1 [b]	2	5	16	21	12
2	PI 88788	1	1	3	95	82	95
3	PI 90763	0	3	3	1.3	0.3	0.4
4	PI 437654	0	0	1	0	0	0
5	PI 209332	1	2	6	NA	NA	NA
6	PI 89772	0	0	3	NA	NA	NA
7	'Cloud'	4	6	9	NA	NA	NA
	'Pickett'[c]	5	4	10	65	68	74

[a] Modified designation test with five indicator lines was used for HG type 1.2.-.
[b] The female index in percentage using the susceptible check, 'Lee 74'.
[c] 'Pickett' was used to classify SCN populations based on an old designation test. HG type 0 is equivalent to SCN race 3, while HG type 1.2.- = race 2.
[d] The number of cysts on 'Lee 74' was 155, 125.8, 156.3 in HG type 0 test 1, 2 and 3. The number of cysts on 'Lee 74' was 243, 225, 220 in HG type 1.2- test 1, 2 and 3

Table 2.2 Female index of two recombinant inbred lines (RILs) along with their parents included in whole genome resequencing

Genotype	Type	HG type 0 (Female index, %)	HG type 1.2.- (Female index, %)
G19-SPT1748	RIL	0.5	0.5
G19-SPT1719	RIL	144.9	63.3
PI 567295	Parent	12	23.8
'Bossier'	Parent	107.5	71.4

Table 2.3 Number of informative SNPs and their physical positions identified with a bulked segregant analysis in a $F_{2:3}$ population derived from 'Bossier' × PI 567295 for resistance to HG type 0

Chromosome	Number of informative SNPs	Physical position [a]
2	12	13.9-39.7
5	23	1.5-2.8
10	58	39.3-44.1
11	31	0.4-3.6
14	22	43.4-49.5
19	13	12.0-46.8

[a]Physical position in Mb was based on Wm82 version 2 reference genome.

Table 2.4 Linkage groups constructed using SoySNP6K BeadChip in an $F_{5:6}$ RIL population derived from a cross of 'Bossier' × PI 567295

Chromosome	Number of markers	Length
1	90	114.8
2	138	132.7
3	83	127.8
4	44	106.9
5	82	101.5
6	66	83.9
7	99	124.6
8	136	154.1
9	89	104.9
10	124	97.3
11	87	102.7
12	81	92.5
13	97	103.0
14	83	119.3
15	75	139.3
16	74	97.4
17	117	131.3
18	58	65.4
19	110	71.1
20	71	95.3
Total	1,804	2,165.7

Table 2.5 QTLs identified on chromosomes 6, 8, and 10 that were associated with resistance to HG type 0 and HG type 1.2.-.

HG type	Chr [a]	Significant marker	Confidence interval	LOD score	R2	Additive effect	Source of favorable Allele
1.2.-	6	Gm06_6914278	Gm06_6797764-Gm06_7323345	9.2	13.0	-9.6	PI 567295
	8	Gm08_18225805	Gm08_17399376-Gm08_18225805	5.02	6.2	-5.9	PI 567295
	10	GSM 825	GSM825-Gm10_42845506	18.3	39.3	-13.9	PI 567295
0	10	GSM 825	GSM825-Gm10_42845506	21.8	53.7	-23.4	PI 567295

[a] Chr is Chromosome.

Table 2.6 6 QTL interaction and effect of QTLs on chromosome 6, 8, and 10 that were associated with resistance to HG type 1.2.-

QTL	BLUP value	No of lines	Tukey HSD test (p=0.01)
Absence of QTLs	16.6	25	A
Chr 8 only	10.3	17	AB
Chr 6 only	6.6	15	AB
Chr 10 only	4.52	23	B
Chr 6+ 8	5.03	16	B
Chr 8+ 10	-27.7	11	C
Chr 6+ 10	-30.2	4	C
Chr 6 + 8 + 10	-33.3	13	C

Table 2.7 List of genes carrying non-synonymous SNPs from the confidence interval (556 kb) of the QTL on Chr 10 for resistance to HG types 0 and 1.2.-

Gene model	Position (bp) [a]	Annotation	Gene ontology
Glyma.10g198000	42923799-42930121	Galactose oxidase/kelch repeat superfamily protein	protein binding
Glyma.10g198300	42942649-42951239	binding to TOMV RNA 1L (long form)	nucleic acid binding
Glyma.10g198500	42958437-42959619	Histone superfamily protein	nucleosome
Glyma.10g198600	42979388-42985630	RING/U-box superfamily protein	regulation of signal transduction
Glyma.10g198800	42995962-42997094	Haemoglobin 2	iron ion binding
Glyma.10g198900	43000237-43003014	Haemoglobin 2	transport
Glyma.10g199000	43004558-43005947	Haemoglobin 2	transport
Glyma.10g199100	43009251-43011119	Haemoglobin 2	transport
Glyma.10g199300	43021060-43026276	Coatomer, alpha subunit	Intracellular protein transport
Glyma.10g199800	43069752-43071805	F-box and associated interaction domains-containing protein	protein binding
Glyma.10g200000	43077358-43086081	UDP-Glycosyltransferase superfamily protein	biosynthetic process
Glyma.10g200200	43089459-43103008	phosphate transporter traffic facilitator1	protein binding
Glyma.10g200400	43118470-43121988	F-box/RNI-like superfamily protein	protein binding
Glyma.10g200600	43131862-43139512	3-phosphoinositide-dependent protein kinase	protein kinase activity
Glyma.10g200700	43142100-43147237	glycosyltransferase family protein 47	carbohydrate metabolic process
Glyma.10g201300	43244789-43259165	Myotubularin-like phosphatases II superfamily	phosphatase activity
Glyma.10g201800	43296070-43300577	Protein with RING/U-box and TRAF-like domains	ubiquitin-dependent protein catabolic process
Glyma.10g201900	43302114-43309900	Vps51/Vps67 family (components of vesicular transport) protein	intra-Golgi vesicle-mediated transport

[a] The annotation and gene ontology are from Soybase database (Soybase.org). The position is based on the reference genome, Williams 82 version 2.

Table 2.8 List of candidate genes from the confidence interval (576kb) of the QTL on Chr 6 for resistance to HG type 1.2.-

Gene model	Position (bp) [a]	Annotation	Gene ontology
Glyma.06g090300	7005790-7019054	Putative serine esterase family protein	cellular lipid metabolic process
Glyma.06g090400	7022045-7025738	xylulose kinase-2	carbohydrate metabolic process
Glyma.06g091100	7112836-7115139	MATE efflux family protein	transmembrane transport
Glyma.06g092600	7304843-7306654	Pentatricopeptide repeat (PPR-like) superfamily protein	nucleic acid phosphodiester bond hydrolysis
Glyma.06g093000	7329072-7330965	Pentatricopeptide repeat (PPR) superfamily protein	glycolytic process

[a] The annotation and gene ontology are from Soybase database (Soybase.org). The position is based on the reference genome, Williams 82 version 2.

Table 2.9 List of candidate genes from the confidence interval (927 kb) of the QTL on Chr 8 for resistance to HG type 1.2.-

Gene model	Position (bp) [a]	Annotation	Gene ontology
Glyma.08G207400	16756340-16761352	CASC3/Barentsz eIF4AIII binding	metal ion transport
Glyma.08G207600	16763273-16767452	UDP-Glycosyltransferase superfamily protein	biosynthetic process
Glyma.08g210000	16998098-17005096	WD40 repeat protein	protein binding
Glyma.08g214200	17296769-17297666	Actin-like ATPase superfamily protein	phosphotransferase activity, alcohol group as acceptor
Glyma.08g220200	17889781-17898867	hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase	transferase activity, transferring acyl groups other than amino-acyl groups
Glyma.08g221600	17990599-17991857	B-box 32	intracellular

[a] The annotation and gene ontology are from Soybase database (Soybase.org). The position is based on the reference genome, Williams 82 version 2.

Table 2.10 Single factor analysis of KASP SNP markers that were located within the exons of candidate genes on chromosomes 6 and 8 for resistance to HG type 1.2.-.

Chr	Candidate gene	KASP SNP marker	Position (bp) [a]	P-value
6	Glyma.06g090300	GSM1028	7,006,724	0.02
	Glyma.06g090300	GSM1029	7,013,960	0.02
	Glyma.06g091100	GSM1030	7,114,426	0.04
	Glyma.06g092600	GSM1031	7,305,514	0.06
8	Glyma.08G207400	GSM1033	16,757,883	0.07
	Glyma.08g221600	GSM1034	17,991,318	0.01
	Glyma.08g220200	GSM1037	17,897,946	0.01
	Glyma.08g214200	GSM1038	17,297,291	0.12
	Glyma.08g210000	GSM1039	17,001,799	0.0003
	Glyma.08G207600	GSM1040	16,763,950	0.004
	Glyma.08G207600	GSM1041	16,765,577	0.08
	Glyma.08G207600	GSM1042	16,765,596	0.04

[a] The position is based on the reference genome, Williams 82 version 2.

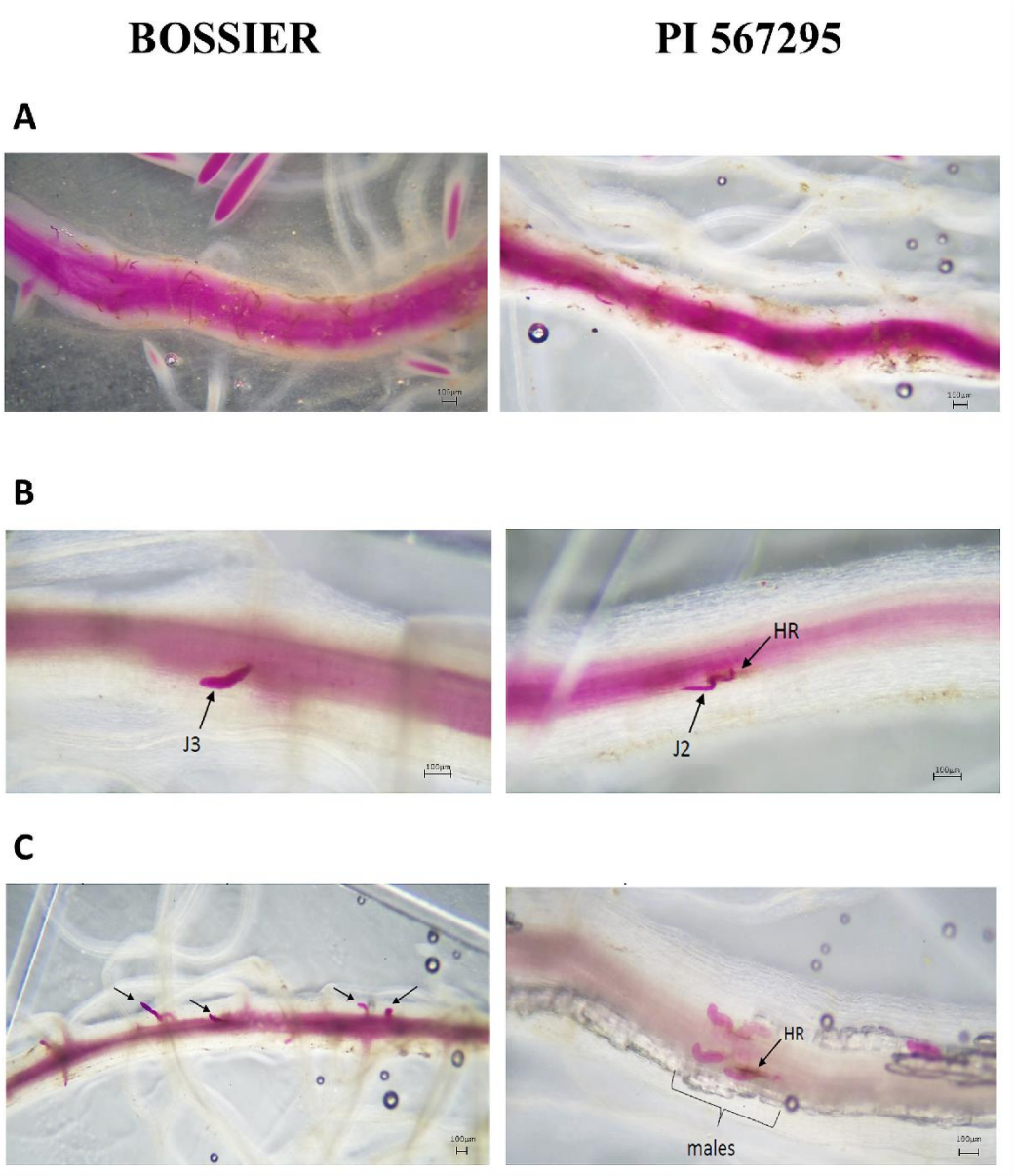


Fig 2.1. Development of soybean cyst nematode in the roots of 'Bossier' and PI 567295. **(A)** 2 days after inoculation (DAI). J2, second stage juvenile penetrated and PI 567295, the resistant parent did not affect the penetration; **(B)** 4 DAI, SCN were at J3, third stage juvenile in 'Bossier' but still J2 stage in PI 567295; and (**C**) 8 DAI, the HR was also identified in PI 567295 and juvenile shifted to males which require less nutrient than female.

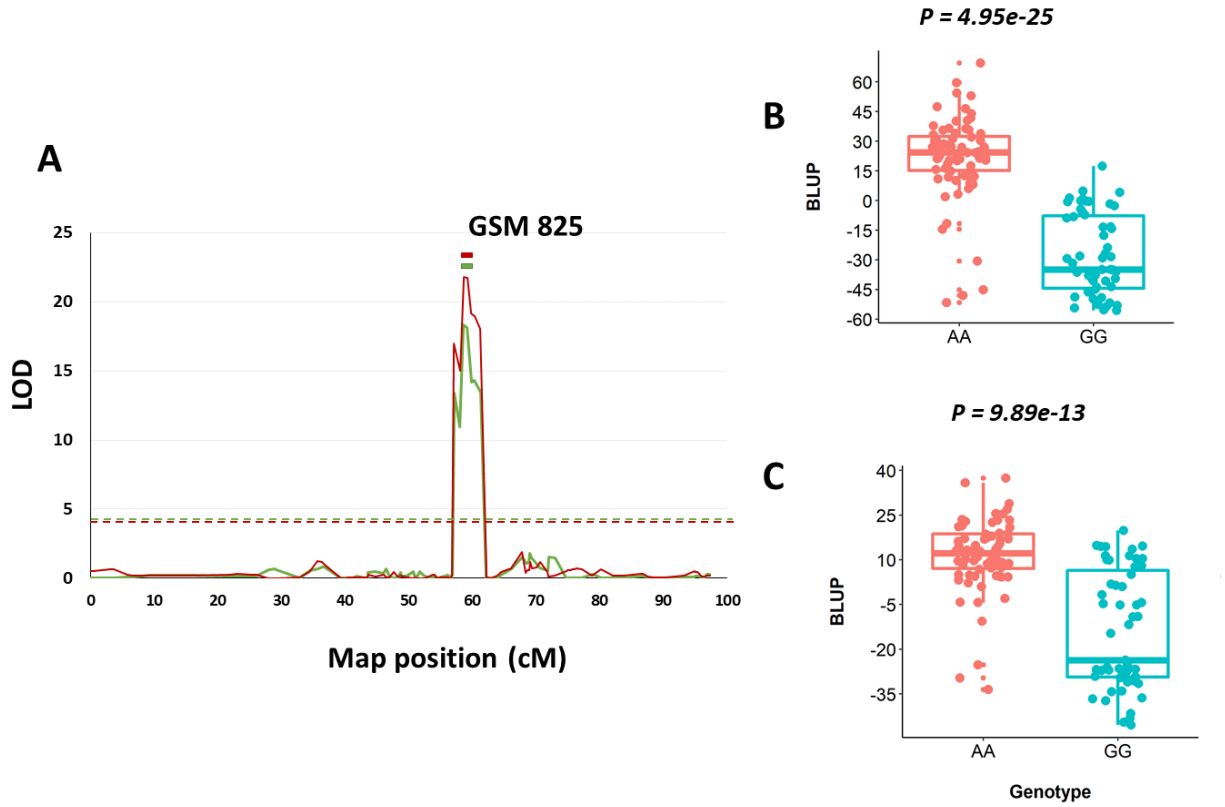


Fig 2.2 QTL on chromosome 10 identified in an RIL population derived from 'Bossier' × PI 567295 and allele effects of the linked SNP GSM825
(A) QTL plot for HG type 0 (Red) and HG type 1.2.- (Green). The significant markers: GSM825 with LOD scores of 21.8 and 18.3 for resistance to HG type 0 and 1.2.-, respectively. LOD thresholds were 4.6 and 4.4for HG type 0 and 1.2.-, respectively. Red and green bars indicated the confidence interval (LOD-1).
(B-C) Box plot showing the genotype-phenotype correlation with two different genotypes at GSM825 using a t-test comparison**.** The solid horizontal lines are the median of BLUP values were **(B**) AA: -34.9; GG: 24.4. **(C)** AA: 12.1; GG: -23.7.

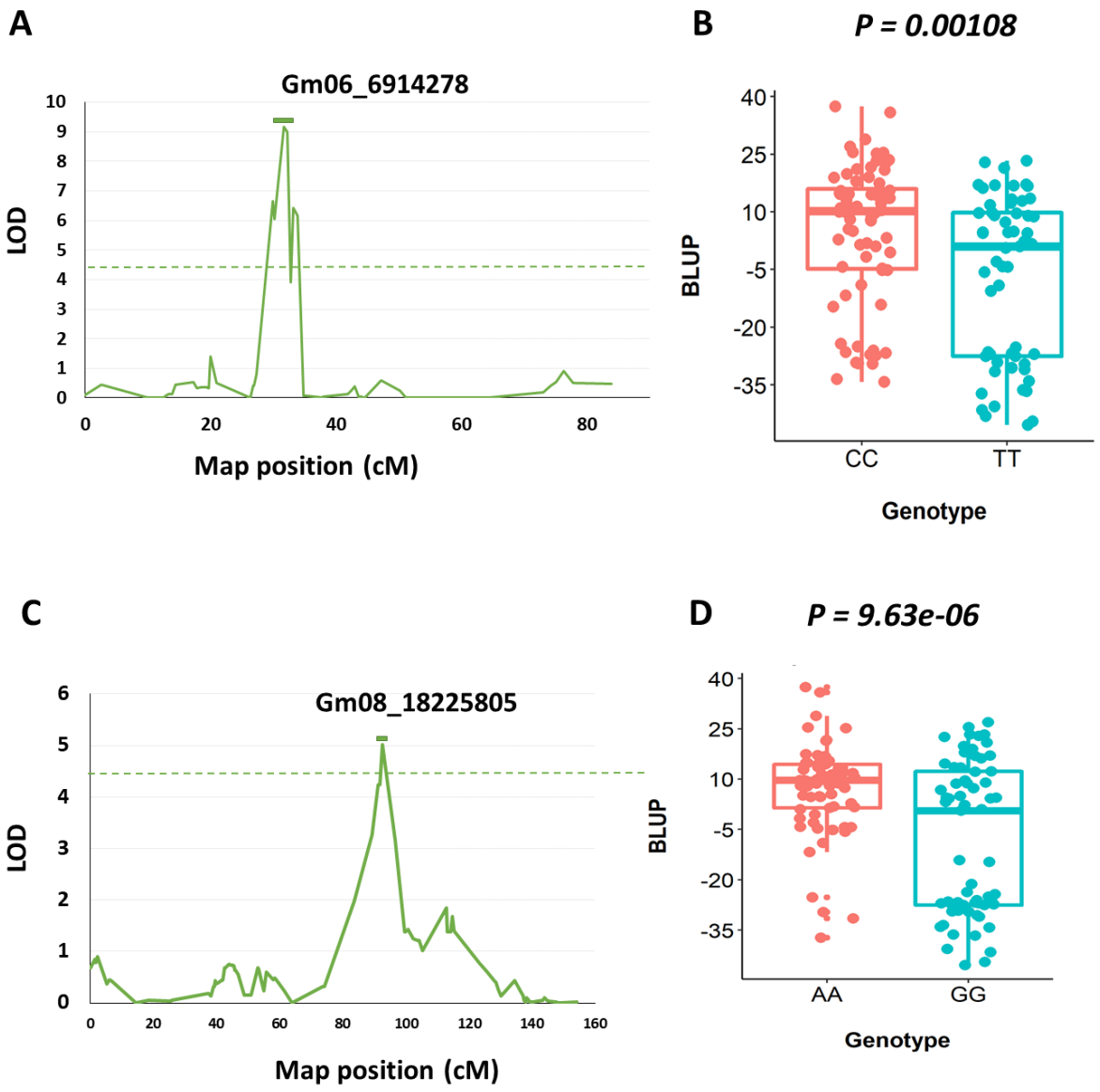


Fig 2.3. QTLs on chromosomes 6 and 8 identified for resistance to HG type 1.2.- and allele effects of linked SNPs, Gm06_6914278 and Gm08_18255805
(A-C) QTL plots showing peaks of Gm06_6914278 and Gm08_18255805 with LOD scores of 9.2 and 5.0, respectively. The LOD threshold was 4.4. The green bar indicates the confidence interval (LOD-1).
(B-D) Box plot showing the genotype-phenotype correlation at two significant markers on **(B)** chromosome 6 and **(D**) chromosome 8 using a t- test comparison**.** The solid horizontal lines are the median of BLUP values were **(B**) CC: 10.2; TT: 0.80. **(D)**; AA: 9.7; GG: 0.61.

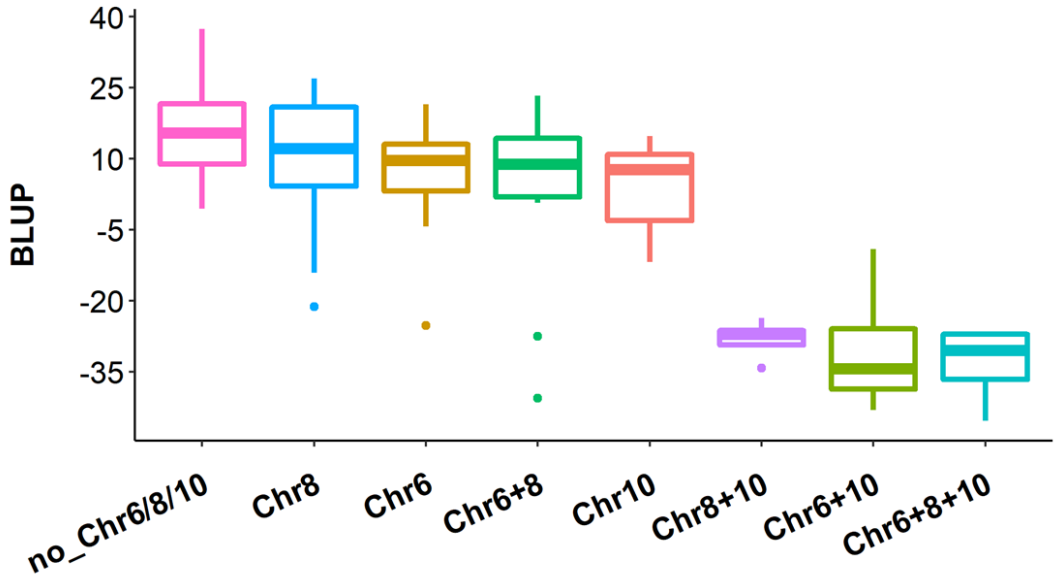


Fig 2.4 BLUP values of the $F_{5:6}$ RILs phenotyped with different combinations of QTLs on chromosomes 6, 8, and 10 that were phenotyped using HG type 1.2.-. The solid horizontal lines are the median of the female index.

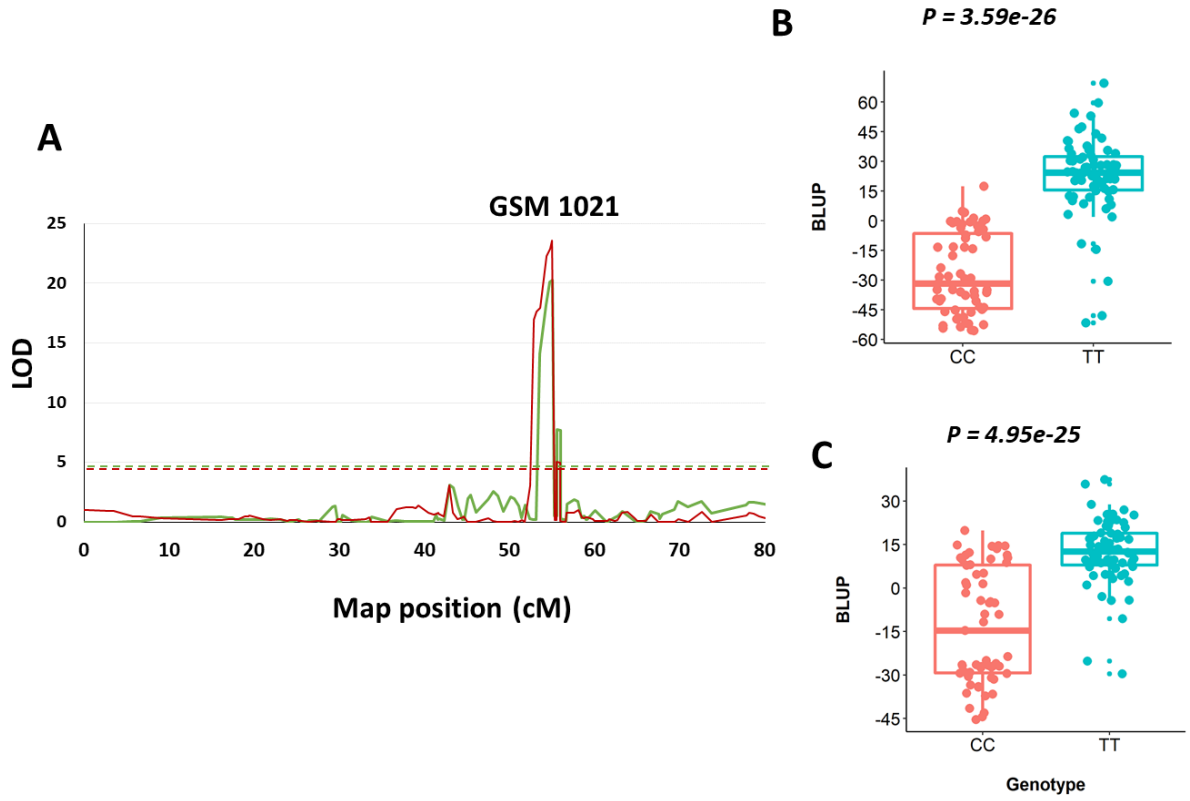


Fig 2.5. Effects of the SNP marker GSM1021, located within the 2nd exon of *Glyma.10g199300* on HG type 0 (Red) and 1.2.- (Green).
(A) QTL plots showing peaks of GSM1021 with LOD scores of 23.6 and 20.26. LOD threshold was ~ 4.6 for both HG types.
(B-C) Box plot showing the genotype-phenotype correlation at GSM1021 using a t-test comparison on **(B)** HG type 1.2.- (Race 2) and **(C)** HG type 0 (Race 3)**.** The solid horizontal lines are the median of BLUP values.

Supplementary tables and figures

Tables S2.1 Primer sequence of significant SNP KASP markers on Chr 10 designed based on bulked segregant and whole resequencing analyses

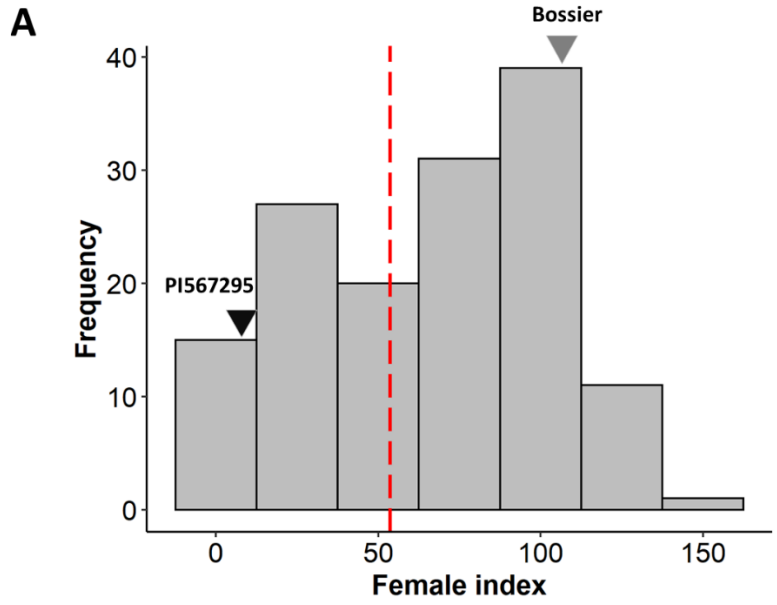
No	Marker name	Primer design	Sequence (5' – 3')	Favorable allele
1	GSM 825	FAM	GAAGGTGACCAAGTTCATGCTTGTGTTGTTT TGGTTCTACCTTAGTACA	A
		VIC	GAAGGTCGGAGTCAACGGATTTGTGTTGTT TTGGTTCTACCTTAGTACG	
		Reverse	ACCAAAATTAGCTCCACACCTCA	
2	GSM 1021	FAM	GAAGGTGACCAAGTTCATGCTAGTCAAAAC GGAGATGCAAGC	C
		VIC	GAAGGTCGGAGTCAACGGATTCAGTCAAAA CGGAGATGCAAGT	
		Reverse	CAAACCATTCGCATCTGGAAC	

[a] FAM and VIC header sequences are in red fonts.

Table S2.2 Single factor analysis of KASP SNP markers that were located within the exons of candidate genes on chromosomes 10 for resistance to HG types 0 and 1.2.-.

Candidate gene	**KASP SNP marker**	**Position (bp) [a]**	**P-value HG type 0**	**P-value HG type 1.2.-**
Glyma.10g201900	GSM1018	43,303,263	4.9e-12	4.1e-10
Glyma.10g201900	GSM1019	43,304,628	2.2e-16	2.5e-13
Glyma.10g201800	GSM1020	43,296,257	1.5e-15	4.2e-11
Glyma.10g199300	GSM1021	43,025,683	2.2e-16	1.4e-14
Glyma.10g199300	GSM1022	43,023,774	5.0e-08	2.4e-07
Glyma.10g198600	GSM1023	42,981,020	1.5e-12	3.5e-06
Glyma.10g198500	GSM1024	42,392,851	1.5e-11	3.3e-09
Glyma.10g198500	GSM1025	42,958,572	3.4e-09	1.07e-05
Glyma.10g196100	GSM1026	42,761,455	3.5e-07	8.8e-05

[a] The position is based on the reference genome, Williams 82 version 2.



.

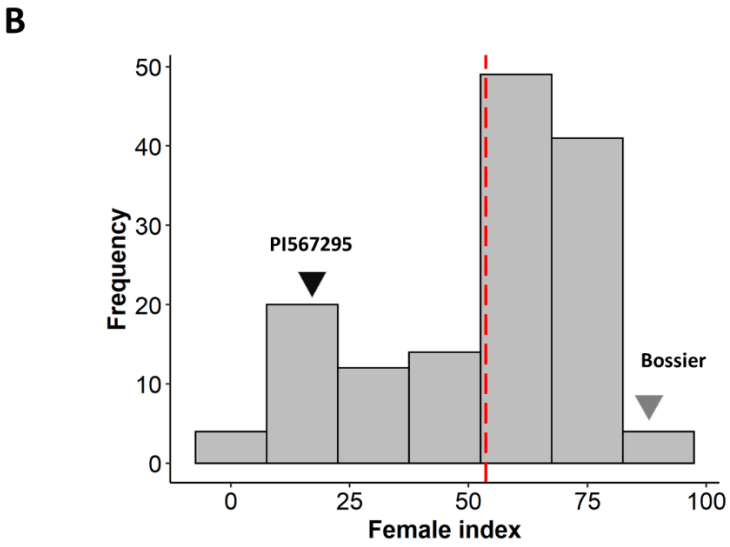
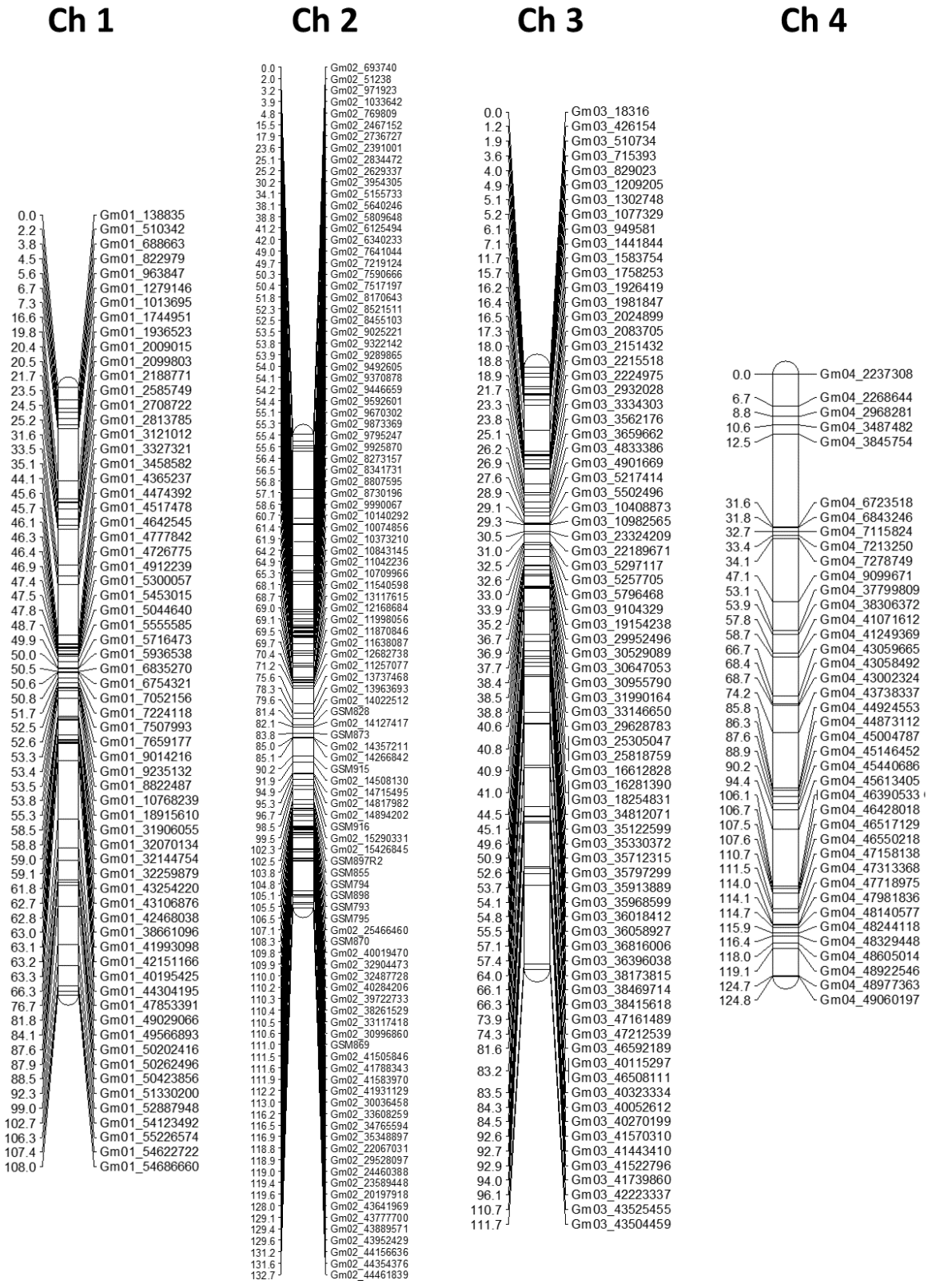
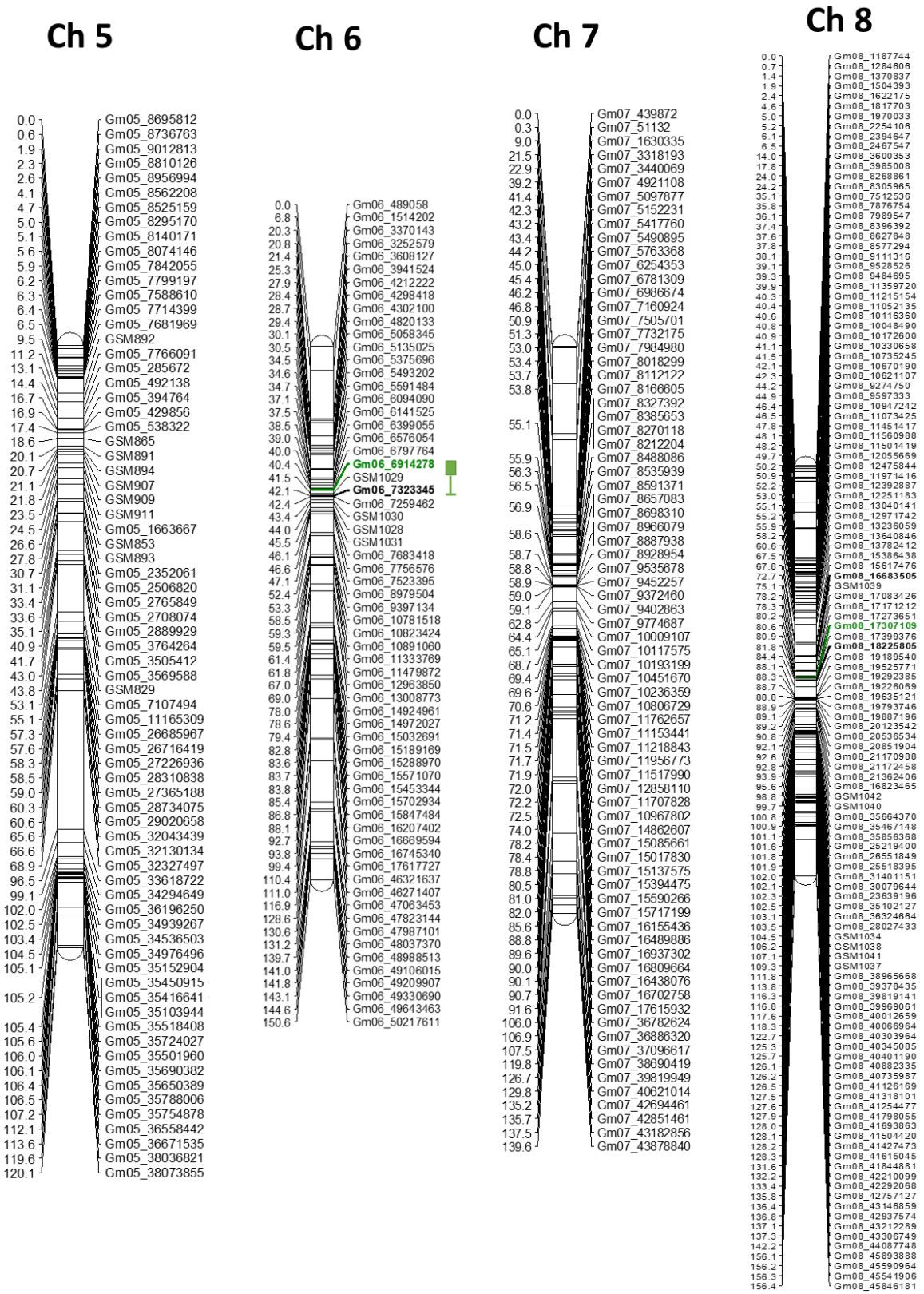
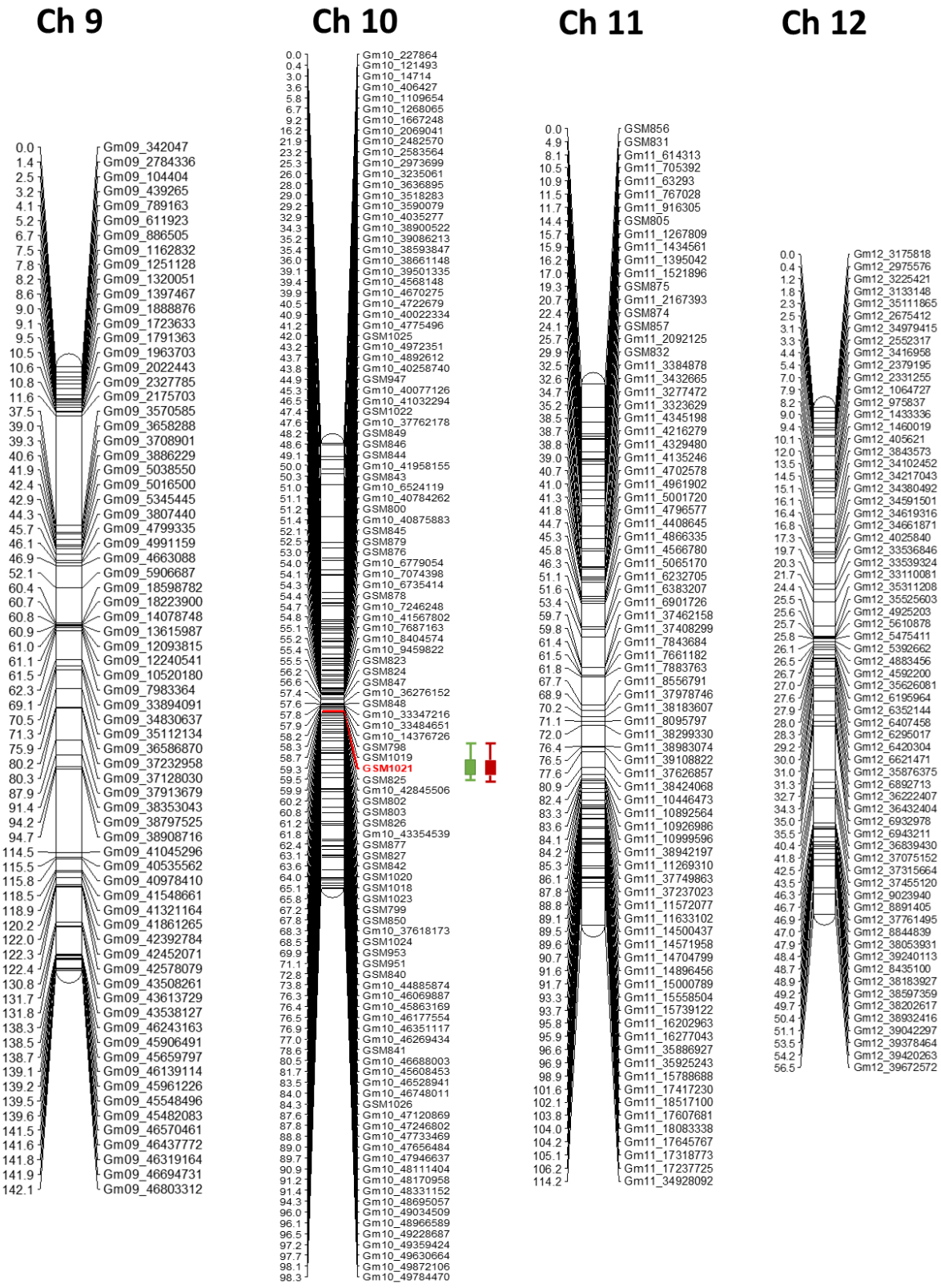
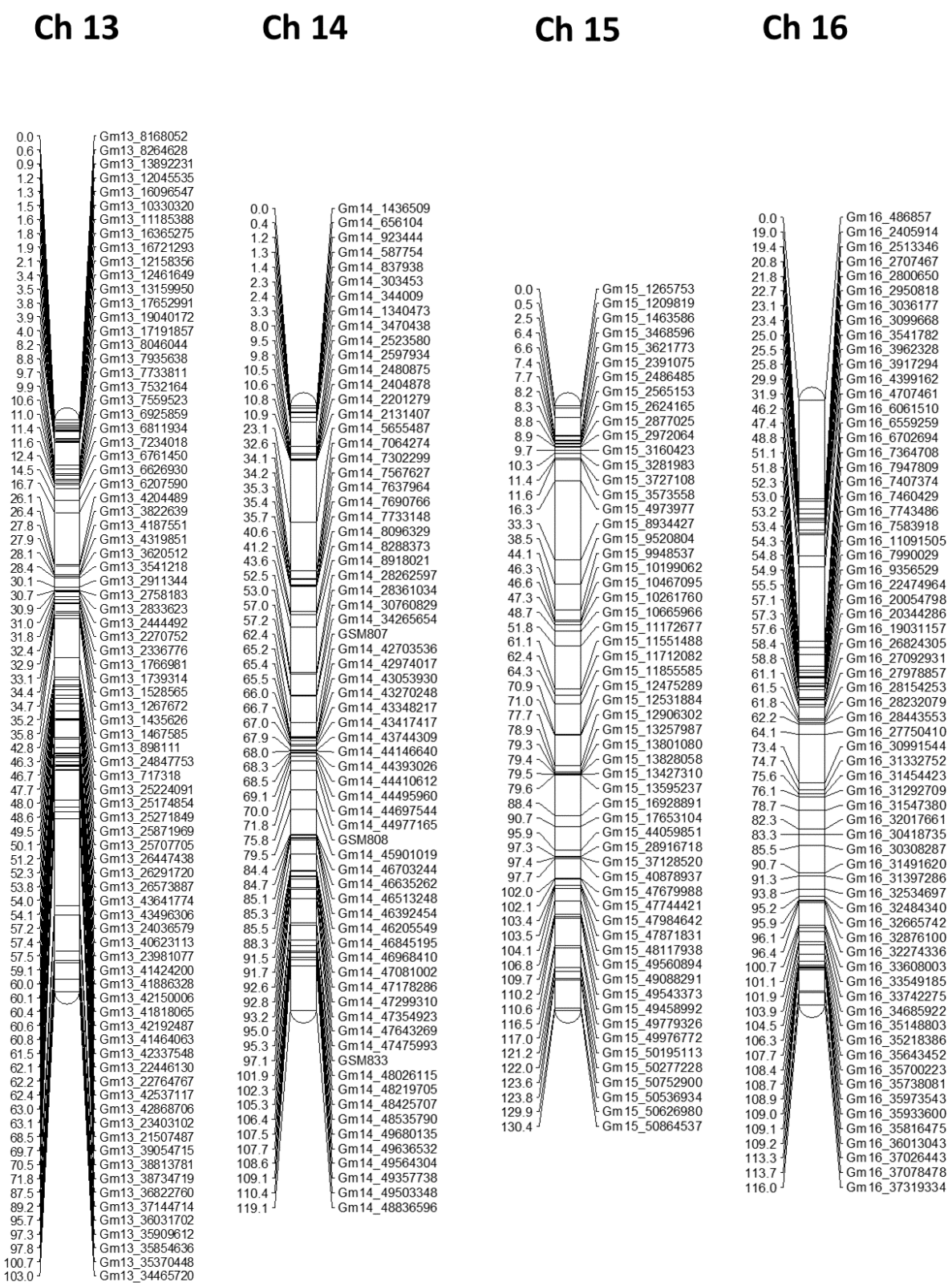


Fig S2.1 Distribution of female indices among 144 $F_{5:6}$ RILs evaluated in greenhouse assays for resistance to HG type 0 and HG type 1.2.-.
(A) HG type 0; (B) HG type 1.2.-. The red dash line indicates a mean value. The black arrow indicates the female index of resistance parent, PI 567295 while the gray arrow is for 'Bossier', the susceptible parent.


Ch 1
Ch 2
Ch 3
Ch 4
0.0 Gm01_138835
2.2 Gm01_510342
3.8 Gm01_688663
4.5 Gm01_822979
5.6 Gm01_963847
6.7 Gm01_1279146
7.3 Gm01_1013695
16.6 Gm01_1744951
19.8 Gm01_1936523
20.4 Gm01_2009015
20.5 Gm01_2099803
21.7 Gm01_2188771
23.5 Gm01_2585749
24.5 Gm01_2708722
25.2 Gm01_2813785
31.6 Gm01_3121012
33.5 Gm01_3327321
35.1 Gm01_3458582
44.1 Gm01_4365237
45.6 Gm01_4474392
45.7 Gm01_4517478
46.1 Gm01_4642545
46.3 Gm01_4777842
46.4 Gm01_4726775
46.9 Gm01_4912239
47.4 Gm01_5300057
47.5 Gm01_5453015
47.8 Gm01_5044640
48.7 Gm01_5555585
49.9 Gm01_5716473
50.0 Gm01_5936538
50.5 Gm01_6835270
50.6 Gm01_6754321
50.8 Gm01_7052156
51.7 Gm01_7224118
52.5 Gm01_7507993
52.6 Gm01_7659177
53.3 Gm01_9014216
53.4 Gm01_9235132
53.5 Gm01_8822487
53.8 Gm01_10768239
55.3 Gm01_18915610
58.5 Gm01_31906055
58.8 Gm01_32070134
59.0 Gm01_32144754
59.1 Gm01_32259879
61.8 Gm01_43254220
62.7 Gm01_43106876
62.8 Gm01_42468038
63.0 Gm01_38661096
63.1 Gm01_41993098
63.2 Gm01_42151166
63.3 Gm01_40195425
66.3 Gm01_44304195
76.7 Gm01_47853391
81.8 Gm01_49029066
84.1 Gm01_49566893
87.6 Gm01_50202416
87.9 Gm01_50262496
88.5 Gm01_50423856
92.3 Gm01_51330200
99.0 Gm01_52887948
102.7 Gm01_54123492
106.3 Gm01_55226574
107.4 Gm01_54622722
108.0 Gm01_54686660
0.0 Gm02_693740
2.0 Gm02_51238
3.2 Gm02_971923
3.9 Gm02_1033642
4.8 Gm02_769809
15.5 Gm02_2467152
17.9 Gm02_2736727
23.6 Gm02_2391001
25.1 Gm02_2834472
25.2 Gm02_2629337
30.2 Gm02_3954305
34.1 Gm02_5155733
38.1 Gm02_5640246
38.8 Gm02_5809648
41.2 Gm02_6125494
42.0 Gm02_6340233
49.0 Gm02_7641044
49.7 Gm02_7219124
50.3 Gm02_7590666
50.4 Gm02_7517197
51.8 Gm02_8170643
52.3 Gm02_8521511
52.5 Gm02_8455103
53.5 Gm02_9025221
53.8 Gm02_9322142
53.9 Gm02_9289865
54.0 Gm02_9492605
54.1 Gm02_9370878
54.2 Gm02_9446659
54.4 Gm02_9592601
55.1 Gm02_9670302
55.3 Gm02_9873369
55.4 Gm02_9795247
55.6 Gm02_9925870
56.4 Gm02_8273157
56.5 Gm02_8341731
56.8 Gm02_8807595
57.1 Gm02_8730196
58.6 Gm02_9990067
60.7 Gm02_10140292
61.4 Gm02_10074856
61.9 Gm02_10373210
64.2 Gm02_10843145
64.9 Gm02_11042236
65.3 Gm02_10709966
68.1 Gm02_11540598
68.7 Gm02_13117615
69.0 Gm02_12168684
69.1 Gm02_11998056
69.5 Gm02_11870846
69.7 Gm02_11638087
70.4 Gm02_12682738
71.2 Gm02_11257077
75.6 Gm02_13737468
78.3 Gm02_13963693
79.6 Gm02_14022512
81.4 GSM828
82.1 Gm02_14127417
83.8 GSM873
85.0 Gm02_14357211
85.1 Gm02_14266842
90.2 GSM915
91.9 Gm02_14508130
94.9 Gm02_14715495
95.3 Gm02_14817982
96.7 Gm02_14894202
98.5 GSM916
99.5 Gm02_15290331
102.3 Gm02_15426845
102.5 GSM897R2
103.8 GSM855
104.8 GSM794
105.1 GSM898
105.5 GSM793
106.5 GSM795
107.1 Gm02_25466460
108.3 GSM870
109.8 Gm02_40019470
109.9 Gm02_32904473
110.0 Gm02_32487728
110.2 Gm02_40284206
110.3 Gm02_39722733
110.4 Gm02_38261529
110.5 Gm02_33117418
110.6 Gm02_30996860
111.0 GSM869
111.5 Gm02_41505846
111.6 Gm02_41788343
111.9 Gm02_41583970
112.2 Gm02_41931129
113.0 Gm02_30036458
116.2 Gm02_33608259
116.5 Gm02_34765594
116.9 Gm02_35348897
118.8 Gm02_22067031
118.9 Gm02_29528097
119.0 Gm02_24460388
119.4 Gm02_23589448
119.6 Gm02_20197918
128.0 Gm02_43641969
129.1 Gm02_43777700
129.4 Gm02_43889571
129.6 Gm02_43952429
131.2 Gm02_44156636
131.6 Gm02_44354376
132.7 Gm02_44461839
0.0 Gm03_18316
1.2 Gm03_426154
1.9 Gm03_510734
3.6 Gm03_715393
4.0 Gm03_829023
4.9 Gm03_1209205
5.1 Gm03_1302748
5.2 Gm03_1077329
6.1 Gm03_949581
7.1 Gm03_1441844
11.7 Gm03_1583754
15.7 Gm03_1758253
16.2 Gm03_1926419
16.4 Gm03_1981847
16.5 Gm03_2024899
17.3 Gm03_2083705
18.0 Gm03_2151432
18.8 Gm03_2215518
18.9 Gm03_2224975
21.7 Gm03_2932028
23.3 Gm03_3334303
23.8 Gm03_3562176
25.1 Gm03_3659662
26.2 Gm03_4833386
26.9 Gm03_4901669
27.6 Gm03_5217414
28.9 Gm03_5502496
29.1 Gm03_10408873
29.3 Gm03_10982565
30.5 Gm03_23324209
31.0 Gm03_22189671
32.5 Gm03_5297117
32.6 Gm03_5257705
33.0 Gm03_5796468
33.9 Gm03_9104329
35.2 Gm03_19154238
36.7 Gm03_29952496
36.9 Gm03_30529089
37.7 Gm03_30647053
38.4 Gm03_30955790
38.5 Gm03_31990164
38.8 Gm03_33146650
40.6 Gm03_29628783
40.8 Gm03_25305047
Gm03_25818759
40.9 Gm03_16612828
41.0 Gm03_16281390
Gm03_18254831
44.5 Gm03_34812071
45.1 Gm03_35122599
49.6 Gm03_35330372
50.9 Gm03_35712315
52.6 Gm03_35797299
53.7 Gm03_35913889
54.1 Gm03_35968599
54.8 Gm03_36018412
55.5 Gm03_36058927
57.1 Gm03_36816006
57.4 Gm03_36396038
64.0 Gm03_38173815
66.1 Gm03_38469714
66.3 Gm03_38415618
73.9 Gm03_47161489
74.3 Gm03_47212539
81.6 Gm03_46592189
83.2 Gm03_40115297
Gm03_46508111
83.5 Gm03_40323334
84.3 Gm03_40052612
84.5 Gm03_40270199
92.6 Gm03_41570310
92.7 Gm03_41443410
92.9 Gm03_41522796
94.0 Gm03_41739860
96.1 Gm03_42223337
110.7 Gm03_43525455
111.7 Gm03_43504459
0.0 Gm04_2237308
6.7 Gm04_2268644
8.8 Gm04_2968281
10.6 Gm04_3487482
12.5 Gm04_3845754
31.6 Gm04_6723518
31.8 Gm04_6843246
32.7 Gm04_7115824
33.4 Gm04_7213250
34.1 Gm04_7278749
47.1 Gm04_9099671
53.1 Gm04_37799809
53.9 Gm04_38306372
57.8 Gm04_41071612
58.7 Gm04_41249369
66.7 Gm04_43059665
68.4 Gm04_43058492
68.7 Gm04_43002324
74.2 Gm04_43738337
85.8 Gm04_44924553
86.3 Gm04_44873112
87.6 Gm04_45004787
88.9 Gm04_45146452
90.2 Gm04_45440686
94.4 Gm04_45613405
106.1 Gm04_46390533
106.7 Gm04_46428018
107.5 Gm04_46517129
107.6 Gm04_46550218
110.7 Gm04_47158138
111.5 Gm04_47313368
114.0 Gm04_47718975
114.1 Gm04_47981836
114.7 Gm04_48140577
115.9 Gm04_48244118
116.4 Gm04_48329448
118.0 Gm04_48605014
119.1 Gm04_48922546
124.7 Gm04_48977363
124.8 Gm04_49060197


Ch 5
0.0 Gm05_8695812
0.6 Gm05_8736763
1.9 Gm05_9012813
2.3 Gm05_8810126
2.6 Gm05_8956994
4.1 Gm05_8562208
4.7 Gm05_8525159
5.0 Gm05_8295170
5.1 Gm05_8140171
5.6 Gm05_8074146
5.9 Gm05_7842055
6.2 Gm05_7799197
6.3 Gm05_7588610
6.4 Gm05_7714399
6.5 Gm05_7681969
9.5 GSM892
11.2 Gm05_7766091
13.1 Gm05_285672
14.4 Gm05_492138
16.7 Gm05_394764
16.9 Gm05_429856
17.4 Gm05_538322
18.6 GSM865
20.1 GSM891
20.7 GSM894
21.1 GSM907
21.8 GSM909
23.5 GSM911
24.5 Gm05_1663667
26.6 GSM853
27.8 GSM893
30.7 Gm05_2352061
31.1 Gm05_2506820
33.4 Gm05_2765849
33.6 Gm05_2708074
35.1 Gm05_2889929
40.9 Gm05_3764264
41.7 Gm05_3505412
43.0 Gm05_3569588
43.8 GSM829
53.1 Gm05_7107494
55.1 Gm05_11165309
57.3 Gm05_26685967
57.6 Gm05_26716419
58.3 Gm05_27226936
58.5 Gm05_28310838
59.0 Gm05_27365188
60.3 Gm05_28734075
60.6 Gm05_29020658
65.6 Gm05_32043439
66.6 Gm05_32130134
68.9 Gm05_32327497
96.5 Gm05_33618722
99.1 Gm05_34294649
102.0 Gm05_36196250
102.5 Gm05_34939267
103.4 Gm05_34536503
104.5 Gm05_34976496
105.1 Gm05_35152904
Gm05_35450915
105.2 Gm05_35416641
Gm05_35103944
105.4 Gm05_35518408
105.6 Gm05_35724027
106.0 Gm05_35501960
106.1 Gm05_35690382
106.4 Gm05_35650389
106.5 Gm05_35788006
107.2 Gm05_35754878
112.1 Gm05_36558442
113.6 Gm05_36671535
119.6 Gm05_38036821
120.1 Gm05_38073855
Ch 6
0.0 Gm06_489058
6.8 Gm06_1514202
20.3 Gm06_3370143
20.8 Gm06_3252579
21.4 Gm06_3608127
25.3 Gm06_3941524
27.9 Gm06_4212222
28.4 Gm06_4298418
28.7 Gm06_4302100
29.4 Gm06_4820133
30.1 Gm06_5058345
30.5 Gm06_5135025
34.5 Gm06_5375696
34.6 Gm06_5493202
34.7 Gm06_5591484
37.1 Gm06_6094090
37.5 Gm06_6141525
38.5 Gm06_6399055
39.0 Gm06_6576054
40.0 Gm06_6797764
40.4 Gm06_6914278
41.5 GSM1029
42.1 Gm06_7323345
42.4 Gm06_7259462
43.4 GSM1030
44.0 GSM1028
45.5 GSM1031
46.1 Gm06_7683418
46.6 Gm06_7756576
47.1 Gm06_7523395
52.4 Gm06_8979504
53.3 Gm06_9397134
58.5 Gm06_10781518
59.3 Gm06_10823424
59.5 Gm06_10891060
61.4 Gm06_11333769
61.8 Gm06_11479872
67.0 Gm06_12963850
69.0 Gm06_13008773
78.0 Gm06_14924961
78.6 Gm06_14972027
79.4 Gm06_15032691
82.8 Gm06_15189169
83.6 Gm06_15288970
83.7 Gm06_15571070
83.8 Gm06_15453344
85.4 Gm06_15702934
86.8 Gm06_15847484
88.1 Gm06_16207402
92.7 Gm06_16669594
93.8 Gm06_16745340
99.4 Gm06_17617727
110.4 Gm06_46321637
111.0 Gm06_46271407
116.9 Gm06_47063453
128.6 Gm06_47823144
130.6 Gm06_47987101
131.2 Gm06_48037370
139.7 Gm06_48988513
141.0 Gm06_49106015
141.8 Gm06_49209907
143.1 Gm06_49330690
144.6 Gm06_49643463
150.6 Gm06_50217611
Ch 7
0.0 Gm07_439872
0.3 Gm07_51132
9.0 Gm07_1630335
21.5 Gm07_3318193
22.9 Gm07_3440069
39.2 Gm07_4921108
41.4 Gm07_5097877
42.3 Gm07_5152231
43.2 Gm07_5417760
43.4 Gm07_5490895
44.2 Gm07_5763368
45.0 Gm07_6254353
45.4 Gm07_6781309
46.2 Gm07_6986674
46.8 Gm07_7160924
50.9 Gm07_7505701
51.3 Gm07_7732175
53.0 Gm07_7984980
53.4 Gm07_8018299
53.7 Gm07_8112122
53.8 Gm07_8166605
Gm07_8327392
55.1 Gm07_8385653
Gm07_8270118
Gm07_8212204
55.9 Gm07_8488086
56.3 Gm07_8535939
56.5 Gm07_8591371
56.9 Gm07_8657083
Gm07_8698310
58.6 Gm07_8966079
Gm07_8887938
58.7 Gm07_8928954
58.8 Gm07_9535678
58.9 Gm07_9452257
59.0 Gm07_9372460
59.1 Gm07_9402863
62.8 Gm07_9774687
64.4 Gm07_10009107
65.1 Gm07_10117575
68.7 Gm07_10193199
69.4 Gm07_10451670
69.6 Gm07_10236359
70.6 Gm07_10806729
71.2 Gm07_11762657
71.4 Gm07_11153441
71.5 Gm07_11218843
71.7 Gm07_11956773
71.9 Gm07_11517990
72.0 Gm07_12858110
72.2 Gm07_11707828
72.5 Gm07_10967802
74.0 Gm07_14862607
78.2 Gm07_15085661
78.4 Gm07_15017830
78.8 Gm07_15137575
80.5 Gm07_15394475
81.0 Gm07_15590266
82.0 Gm07_15717199
85.6 Gm07_16155436
88.8 Gm07_16489886
89.6 Gm07_16937302
90.0 Gm07_16809664
90.1 Gm07_16438076
90.7 Gm07_16702758
91.6 Gm07_17615932
106.0 Gm07_36782624
106.9 Gm07_36886320
107.5 Gm07_37096617
119.8 Gm07_38690419
126.7 Gm07_39819949
129.8 Gm07_40621014
135.2 Gm07_42694461
135.7 Gm07_42851461
137.5 Gm07_43182856
139.6 Gm07_43878840
Ch 8
0.0 Gm08_1187744
0.7 Gm08_1284606
1.4 Gm08_1370837
1.9 Gm08_1504393
2.4 Gm08_1622175
4.6 Gm08_1817703
5.0 Gm08_1970033
5.2 Gm08_2254106
6.1 Gm08_2394647
6.5 Gm08_2467547
14.0 Gm08_3600353
17.8 Gm08_3985008
24.0 Gm08_8268861
24.2 Gm08_8305965
35.1 Gm08_7512536
35.8 Gm08_7876754
36.1 Gm08_7989547
37.4 Gm08_8396392
37.6 Gm08_8627848
37.8 Gm08_8577294
38.1 Gm08_9111316
39.1 Gm08_9528526
39.3 Gm08_9484695
39.9 Gm08_11359720
40.3 Gm08_11215154
40.4 Gm08_11052135
40.6 Gm08_10116360
40.8 Gm08_10048490
40.9 Gm08_10172600
41.1 Gm08_10330658
41.5 Gm08_10735245
42.1 Gm08_10670190
42.3 Gm08_10621107
44.2 Gm08_9274750
44.9 Gm08_9597333
46.4 Gm08_10947242
46.5 Gm08_11073425
47.8 Gm08_11451417
48.1 Gm08_11560988
48.2 Gm08_11501419
49.7 Gm08_12055669
50.2 Gm08_12475844
50.9 Gm08_11971416
52.2 Gm08_12392887
53.0 Gm08_12251183
55.1 Gm08_13040141
55.2 Gm08_12971742
55.9 Gm08_13236059
58.2 Gm08_13640846
60.6 Gm08_13782412
67.5 Gm08_15386438
67.8 Gm08_15617476
72.7 Gm08_16683505
75.1 GSM1039
78.2 Gm08_17083426
78.3 Gm08_17171212
80.2 Gm08_17273651
80.6 Gm08_17307109
80.9 Gm08_17399376
81.8 Gm08_18225805
84.4 Gm08_19189540
88.1 Gm08_19525771
88.3 Gm08_19292385
88.7 Gm08_19226069
88.8 Gm08_19635121
88.9 Gm08_19793146
89.1 Gm08_19887196
89.2 Gm08_20123542
90.8 Gm08_20536534
92.1 Gm08_20851904
92.6 Gm08_21170988
92.8 Gm08_21172458
93.9 Gm08_21362406
95.6 Gm08_16823465
98.8 GSM1042
99.7 GSM1040
100.8 Gm08_35664370
100.9 Gm08_35467148
101.1 Gm08_35856368
101.6 Gm08_25219400
101.8 Gm08_26551849
101.9 Gm08_25518395
102.0 Gm08_31401151
102.1 Gm08_30079644
102.3 Gm08_23639196
102.5 Gm08_35102127
103.1 Gm08_36324664
103.5 Gm08_28027433
104.5 GSM1034
106.2 GSM1038
107.1 GSM1041
109.3 GSM1037
111.8 Gm08_38965668
113.8 Gm08_39378435
116.3 Gm08_39819141
116.8 Gm08_39969061
117.6 Gm08_40012659
118.3 Gm08_40066964
122.7 Gm08_40303964
125.3 Gm08_40345085
125.7 Gm08_40401190
126.1 Gm08_40882335
126.2 Gm08_40735987
126.5 Gm08_41126169
127.5 Gm08_41318101
127.6 Gm08_41254477
127.9 Gm08_41798055
128.0 Gm08_41693863
128.1 Gm08_41504420
128.2 Gm08_41427473
128.3 Gm08_41615045
131.6 Gm08_41844881
132.2 Gm08_42210099
133.4 Gm08_42292068
135.8 Gm08_42757127
136.4 Gm08_43146859
136.8 Gm08_42937574
137.1 Gm08_43212289
137.3 Gm08_43306749
142.2 Gm08_44087748
156.1 Gm08_45893888
156.2 Gm08_45590964
156.3 Gm08_45541906
156.4 Gm08_45846181


Ch 9
Ch 10
Ch 11
Ch 12
0.0 Gm09_342047
1.4 Gm09_2784336
2.5 Gm09_104404
3.2 Gm09_439265
4.1 Gm09_789163
5.2 Gm09_611923
6.7 Gm09_886505
7.5 Gm09_1162832
7.8 Gm09_1251128
8.2 Gm09_1320051
8.6 Gm09_1397467
9.0 Gm09_1888876
9.1 Gm09_1723633
9.5 Gm09_1791363
10.5 Gm09_1963703
10.6 Gm09_2022443
10.8 Gm09_2327785
11.6 Gm09_2175703
37.5 Gm09_3570585
39.0 Gm09_3658288
39.3 Gm09_3708901
40.6 Gm09_3886229
41.9 Gm09_5038550
42.4 Gm09_5016500
42.9 Gm09_5345445
44.3 Gm09_3807440
45.7 Gm09_4799335
46.1 Gm09_4991159
46.9 Gm09_4663088
52.1 Gm09_5906687
60.4 Gm09_18598782
60.7 Gm09_18223900
60.8 Gm09_14078748
60.9 Gm09_13615987
61.0 Gm09_12093815
61.1 Gm09_12240541
61.5 Gm09_10520180
62.3 Gm09_7983364
69.1 Gm09_33894091
70.5 Gm09_34830637
71.3 Gm09_35112134
75.9 Gm09_36586870
80.2 Gm09_37232958
80.3 Gm09_37128030
87.9 Gm09_37913679
91.4 Gm09_38353043
94.2 Gm09_38797525
94.7 Gm09_38908716
114.5 Gm09_41045296
115.5 Gm09_40535562
115.8 Gm09_40978410
118.5 Gm09_41548661
118.9 Gm09_41321164
120.2 Gm09_41861265
122.0 Gm09_42392784
122.3 Gm09_42452071
122.4 Gm09_42578079
130.8 Gm09_43508261
131.7 Gm09_43613729
131.8 Gm09_43538127
138.3 Gm09_46243163
138.5 Gm09_45906491
138.7 Gm09_45659797
139.1 Gm09_46139114
139.2 Gm09_45961226
139.5 Gm09_45548496
139.6 Gm09_45482083
141.5 Gm09_46570461
141.6 Gm09_46437772
141.8 Gm09_46319164
141.9 Gm09_46694731
142.1 Gm09_46803312
0.0 Gm10_227864
0.4 Gm10_121493
3.0 Gm10_14714
3.6 Gm10_406427
5.8 Gm10_1109654
6.7 Gm10_1268065
9.2 Gm10_1667248
16.2 Gm10_2069041
21.9 Gm10_2482570
23.2 Gm10_2583564
25.3 Gm10_2973699
26.0 Gm10_3235061
28.0 Gm10_3636895
29.0 Gm10_3518283
29.2 Gm10_3590079
32.9 Gm10_4035277
34.3 Gm10_38900522
35.2 Gm10_39086213
35.4 Gm10_38593847
36.0 Gm10_38661148
39.1 Gm10_39501335
39.4 Gm10_4568148
39.9 Gm10_4670275
40.5 Gm10_4722679
40.9 Gm10_40022334
41.2 Gm10_4775496
42.0 GSM1025
43.2 Gm10_4972351
43.7 Gm10_4892612
43.8 Gm10_40258740
44.9 GSM947
45.3 Gm10_40077126
46.5 Gm10_41032294
47.4 GSM1022
47.6 Gm10_37762178
48.2 GSM849
48.6 GSM846
49.1 GSM844
50.0 Gm10_41958155
50.3 GSM843
51.0 Gm10_6524119
51.1 Gm10_40784262
51.2 GSM800
51.4 Gm10_40875883
52.1 GSM845
52.5 GSM879
53.0 GSM876
54.0 Gm10_6779054
54.1 Gm10_7074398
54.3 Gm10_6735414
54.4 GSM878
54.7 Gm10_7246248
54.8 Gm10_41567802
55.1 Gm10_7687163
55.2 Gm10_8404574
55.4 Gm10_9459822
55.5 GSM823
56.2 GSM824
56.6 GSM847
57.4 Gm10_36276152
57.6 GSM848
57.8 Gm10_33347216
57.9 Gm10_33484651
58.2 Gm10_14376726
58.3 GSM798
58.7 GSM1019
59.3 GSM1021
59.5 GSM825
59.9 Gm10_42845506
60.2 GSM802
60.8 GSM803
61.2 GSM826
61.8 Gm10_43354539
62.4 GSM877
63.1 GSM827
63.6 GSM842
64.0 GSM1020
65.1 GSM1018
65.8 GSM1023
67.2 GSM799
67.8 GSM850
68.3 Gm10_37618173
68.5 GSM1024
69.9 GSM953
71.1 GSM951
72.8 GSM840
73.8 Gm10_44885874
76.3 Gm10_46069887
76.4 Gm10_45863169
76.5 Gm10_46177554
76.9 Gm10_46351117
77.0 Gm10_46269434
78.6 GSM841
80.5 Gm10_46688003
81.7 Gm10_45608453
83.5 Gm10_46528941
84.0 Gm10_46748011
84.3 GSM1026
87.6 Gm10_47120869
87.8 Gm10_47246802
88.8 Gm10_47733469
89.0 Gm10_47656484
89.7 Gm10_47946637
90.9 Gm10_48111404
91.2 Gm10_48170958
91.4 Gm10_48331152
94.3 Gm10_48695057
96.0 Gm10_49034509
96.1 Gm10_48966589
96.5 Gm10_49228687
97.2 Gm10_49359424
97.7 Gm10_49630664
98.1 Gm10_49872106
98.3 Gm10_49784470
0.0 GSM856
4.9 GSM831
8.1 Gm11_614313
10.5 Gm11_705392
10.9 Gm11_63293
11.5 Gm11_767028
11.7 Gm11_916305
14.4 GSM805
15.7 Gm11_1267809
15.9 Gm11_1434561
16.2 Gm11_1395042
17.0 Gm11_1521896
19.3 GSM875
20.7 Gm11_2167393
22.4 GSM874
24.1 GSM857
25.7 Gm11_2092125
29.9 GSM832
32.5 Gm11_3384878
32.6 Gm11_3432665
34.7 Gm11_3277472
35.2 Gm11_3323629
38.5 Gm11_4345198
38.7 Gm11_4216279
38.8 Gm11_4329480
39.0 Gm11_4135246
40.7 Gm11_4702578
41.0 Gm11_4961902
41.3 Gm11_5001720
41.8 Gm11_4796577
44.7 Gm11_4408645
45.3 Gm11_4866335
45.8 Gm11_4566780
46.3 Gm11_5065170
51.1 Gm11_6232705
51.6 Gm11_6383207
53.4 Gm11_6901726
59.7 Gm11_37462158
59.8 Gm11_37408299
61.4 Gm11_7843684
61.5 Gm11_7661182
61.8 Gm11_7883763
67.7 Gm11_8556791
68.9 Gm11_37978746
70.2 Gm11_38183607
71.1 Gm11_8095797
72.0 Gm11_38299330
76.4 Gm11_38983074
76.5 Gm11_39108822
77.6 Gm11_37626857
80.9 Gm11_38424068
82.4 Gm11_10446473
83.3 Gm11_10892564
83.6 Gm11_10926986
84.1 Gm11_10999596
84.2 Gm11_38942197
85.3 Gm11_11269310
86.1 Gm11_37749863
87.8 Gm11_37237023
88.8 Gm11_11572077
89.1 Gm11_11633102
89.5 Gm11_14500437
89.6 Gm11_14571958
90.7 Gm11_14704799
91.6 Gm11_14896456
91.7 Gm11_15000789
93.3 Gm11_15558504
93.7 Gm11_15739122
95.8 Gm11_16202963
95.9 Gm11_16277043
96.6 Gm11_35886927
96.9 Gm11_35925243
98.9 Gm11_15788688
101.6 Gm11_17417230
102.1 Gm11_18517100
103.8 Gm11_17607681
104.0 Gm11_18083338
104.2 Gm11_17645767
105.1 Gm11_17318773
106.2 Gm11_17237725
114.2 Gm11_34928092
0.0 Gm12_3175818
0.4 Gm12_2975576
1.2 Gm12_3225421
1.8 Gm12_3133148
2.3 Gm12_35111865
2.5 Gm12_2675412
3.1 Gm12_34979415
3.3 Gm12_2552317
4.4 Gm12_3416958
5.4 Gm12_2379195
7.0 Gm12_2331255
7.9 Gm12_1064727
8.2 Gm12_975837
9.0 Gm12_1433336
9.4 Gm12_1460019
10.1 Gm12_405621
12.0 Gm12_3843573
13.5 Gm12_34102452
14.5 Gm12_34217043
15.1 Gm12_34380492
16.1 Gm12_34591501
16.4 Gm12_34619316
16.8 Gm12_34661871
17.3 Gm12_4025840
19.7 Gm12_33536846
20.3 Gm12_33539324
21.7 Gm12_33110081
24.4 Gm12_35311208
25.5 Gm12_35525603
25.6 Gm12_4925203
25.7 Gm12_5610878
25.8 Gm12_5475411
26.1 Gm12_5392662
26.5 Gm12_4883456
26.7 Gm12_4592200
27.0 Gm12_35626081
27.6 Gm12_6195964
27.9 Gm12_6352144
28.0 Gm12_6407458
28.3 Gm12_6295017
29.2 Gm12_6420304
30.0 Gm12_6621471
31.0 Gm12_35876375
31.3 Gm12_6892713
32.7 Gm12_36222407
34.3 Gm12_36432404
35.0 Gm12_6932978
35.5 Gm12_6943211
40.4 Gm12_36839430
41.8 Gm12_37075152
42.5 Gm12_37315664
43.5 Gm12_37455120
46.3 Gm12_9023940
46.7 Gm12_8891405
46.9 Gm12_37761495
47.0 Gm12_8844839
47.9 Gm12_38053931
48.4 Gm12_39240113
48.7 Gm12_8435100
48.9 Gm12_38183927
49.2 Gm12_38597359
49.7 Gm12_38202617
50.4 Gm12_38932416
51.1 Gm12_39042297
53.5 Gm12_39378464
54.2 Gm12_39420263
56.5 Gm12_39672572


Ch 13
Ch 14
Ch 15
Ch 16
0.0 Gm13_8168052
0.6 Gm13_8264628
0.9 Gm13_13892231
1.2 Gm13_12045535
1.3 Gm13_16096547
1.5 Gm13_10330320
1.6 Gm13_11185388
1.8 Gm13_16365275
1.9 Gm13_16721293
2.1 Gm13_12158356
3.4 Gm13_12461649
3.5 Gm13_13159950
3.8 Gm13_17652991
3.9 Gm13_19040172
4.0 Gm13_17191857
8.2 Gm13_8046044
8.8 Gm13_7935638
9.7 Gm13_7733811
9.9 Gm13_7532164
10.6 Gm13_7559523
11.0 Gm13_6925859
11.4 Gm13_6811934
11.6 Gm13_7234018
12.4 Gm13_6761450
14.5 Gm13_6626930
16.7 Gm13_6207590
26.1 Gm13_4204489
26.4 Gm13_3822639
27.8 Gm13_4187551
27.9 Gm13_4319851
28.1 Gm13_3620512
28.4 Gm13_3541218
30.1 Gm13_2911344
30.7 Gm13_2758183
30.9 Gm13_2833623
31.0 Gm13_2444492
31.8 Gm13_2270752
32.4 Gm13_2336776
32.9 Gm13_1766981
33.1 Gm13_1739314
34.4 Gm13_1528565
34.7 Gm13_1267672
35.2 Gm13_1435626
35.8 Gm13_1467585
42.8 Gm13_898111
46.3 Gm13_24847753
46.7 Gm13_717318
47.7 Gm13_25224091
48.0 Gm13_25174854
48.6 Gm13_25271849
49.5 Gm13_25871969
50.1 Gm13_25707705
51.2 Gm13_26447438
52.3 Gm13_26291720
53.8 Gm13_26573887
54.0 Gm13_43641774
54.1 Gm13_43496306
57.2 Gm13_24036579
57.4 Gm13_40623113
57.5 Gm13_23981077
59.1 Gm13_41424200
60.0 Gm13_41886328
60.1 Gm13_42150006
60.4 Gm13_41818065
60.6 Gm13_42192487
60.8 Gm13_41464063
61.5 Gm13_42337548
62.1 Gm13_22446130
62.2 Gm13_22764767
62.4 Gm13_42537117
63.0 Gm13_42868706
63.1 Gm13_23403102
68.5 Gm13_21507487
69.7 Gm13_39054715
70.5 Gm13_38813781
71.8 Gm13_38734719
87.5 Gm13_36822760
89.2 Gm13_37144714
95.7 Gm13_36031702
97.3 Gm13_35909612
97.8 Gm13_35854636
100.7 Gm13_35370448
103.0 Gm13_34465720
0.0 Gm14_1436509
0.4 Gm14_656104
1.2 Gm14_923444
1.3 Gm14_587754
1.4 Gm14_837938
2.3 Gm14_303453
2.4 Gm14_344009
3.3 Gm14_1340473
8.0 Gm14_3470438
9.5 Gm14_2523580
9.8 Gm14_2597934
10.5 Gm14_2480875
10.6 Gm14_2404878
10.8 Gm14_2201279
10.9 Gm14_2131407
23.1 Gm14_5655487
32.6 Gm14_7064274
34.1 Gm14_7302299
34.2 Gm14_7567627
35.3 Gm14_7637964
35.4 Gm14_7690766
35.7 Gm14_7733148
40.6 Gm14_8096329
41.2 Gm14_8288373
43.6 Gm14_8918021
52.5 Gm14_28262597
53.0 Gm14_28361034
57.0 Gm14_30760829
57.2 Gm14_34265654
62.4 GSM807
65.2 Gm14_42703536
65.4 Gm14_42974017
65.5 Gm14_43053930
66.0 Gm14_43270248
66.7 Gm14_43348217
67.0 Gm14_43417417
67.9 Gm14_43744309
68.0 Gm14_44146640
68.3 Gm14_44393026
68.5 Gm14_44410612
69.1 Gm14_44495960
70.0 Gm14_44697544
71.8 Gm14_44977165
75.8 GSM808
79.5 Gm14_45901019
84.4 Gm14_46703244
84.7 Gm14_46635262
85.1 Gm14_46513248
85.3 Gm14_46392454
85.5 Gm14_46205549
88.3 Gm14_46845195
91.5 Gm14_46968410
91.7 Gm14_47081002
92.6 Gm14_47178286
92.8 Gm14_47299310
93.2 Gm14_47354923
95.0 Gm14_47643269
95.3 Gm14_47475993
97.1 GSM833
101.9 Gm14_48026115
102.3 Gm14_48219705
105.3 Gm14_48425707
106.4 Gm14_48535790
107.5 Gm14_49680135
107.7 Gm14_49636532
108.6 Gm14_49564304
109.1 Gm14_49357738
110.4 Gm14_49503348
119.1 Gm14_48836596
0.0 Gm15_1265753
0.5 Gm15_1209819
2.5 Gm15_1463586
6.4 Gm15_3468596
6.6 Gm15_3621773
7.4 Gm15_2391075
7.7 Gm15_2486485
8.2 Gm15_2565153
8.3 Gm15_2624165
8.8 Gm15_2877025
8.9 Gm15_2972064
9.7 Gm15_3160423
10.3 Gm15_3281983
11.4 Gm15_3727108
11.6 Gm15_3573558
16.3 Gm15_4973977
33.3 Gm15_8934427
38.5 Gm15_9520804
44.1 Gm15_9948537
46.3 Gm15_10199062
46.6 Gm15_10467095
47.3 Gm15_10261760
48.7 Gm15_10665966
51.8 Gm15_11172677
61.1 Gm15_11551488
62.4 Gm15_11712082
64.3 Gm15_11855585
70.9 Gm15_12475289
71.0 Gm15_12531884
77.7 Gm15_12906302
78.9 Gm15_13257987
79.3 Gm15_13801080
79.4 Gm15_13828058
79.5 Gm15_13427310
79.6 Gm15_13595237
88.4 Gm15_16928891
90.7 Gm15_17653104
95.9 Gm15_44059851
97.3 Gm15_28916718
97.4 Gm15_37128520
97.7 Gm15_40878937
102.0 Gm15_47679988
102.1 Gm15_47744421
103.4 Gm15_47984642
103.5 Gm15_47871831
104.1 Gm15_48117938
106.8 Gm15_49560894
109.7 Gm15_49088291
110.2 Gm15_49543373
110.6 Gm15_49458992
116.5 Gm15_49779326
117.0 Gm15_49976772
121.2 Gm15_50195113
122.0 Gm15_50277228
123.6 Gm15_50752900
123.8 Gm15_50536934
129.9 Gm15_50626980
130.4 Gm15_50864537
0.0 Gm16_486857
19.0 Gm16_2405914
19.4 Gm16_2513346
20.8 Gm16_2707467
21.8 Gm16_2800650
22.7 Gm16_2950818
23.1 Gm16_3036177
23.4 Gm16_3099668
25.0 Gm16_3541782
25.5 Gm16_3962328
25.8 Gm16_3917294
29.9 Gm16_4399162
31.9 Gm16_4707461
46.2 Gm16_6061510
47.4 Gm16_6559259
48.8 Gm16_6702694
51.1 Gm16_7364708
51.8 Gm16_7947809
52.3 Gm16_7407374
53.0 Gm16_7460429
53.2 Gm16_7743486
53.4 Gm16_7583918
54.3 Gm16_11091505
54.8 Gm16_7990029
54.9 Gm16_9356529
55.5 Gm16_22474964
57.1 Gm16_20054798
57.3 Gm16_20344286
57.6 Gm16_19031157
58.4 Gm16_26824305
58.8 Gm16_27092931
61.1 Gm16_27978857
61.5 Gm16_28154253
61.8 Gm16_28232079
62.2 Gm16_28443553
64.1 Gm16_27750410
73.4 Gm16_30991544
74.7 Gm16_31332752
75.6 Gm16_31454423
76.1 Gm16_31292709
78.7 Gm16_31547380
82.3 Gm16_32017661
83.3 Gm16_30418735
85.5 Gm16_30308287
90.7 Gm16_31491620
91.3 Gm16_31397286
93.8 Gm16_32534697
95.2 Gm16_32484340
95.9 Gm16_32665742
96.1 Gm16_32876100
96.4 Gm16_32274336
100.7 Gm16_33608003
101.1 Gm16_33549185
101.9 Gm16_33742275
103.9 Gm16_34685922
104.5 Gm16_35148803
106.3 Gm16_35218386
107.7 Gm16_35643452
108.4 Gm16_35700223
108.7 Gm16_35738081
108.9 Gm16_35973543
109.0 Gm16_35933600
109.1 Gm16_35816475
109.2 Gm16_36013043
113.3 Gm16_37026443
113.7 Gm16_37078478
116.0 Gm16_37319334

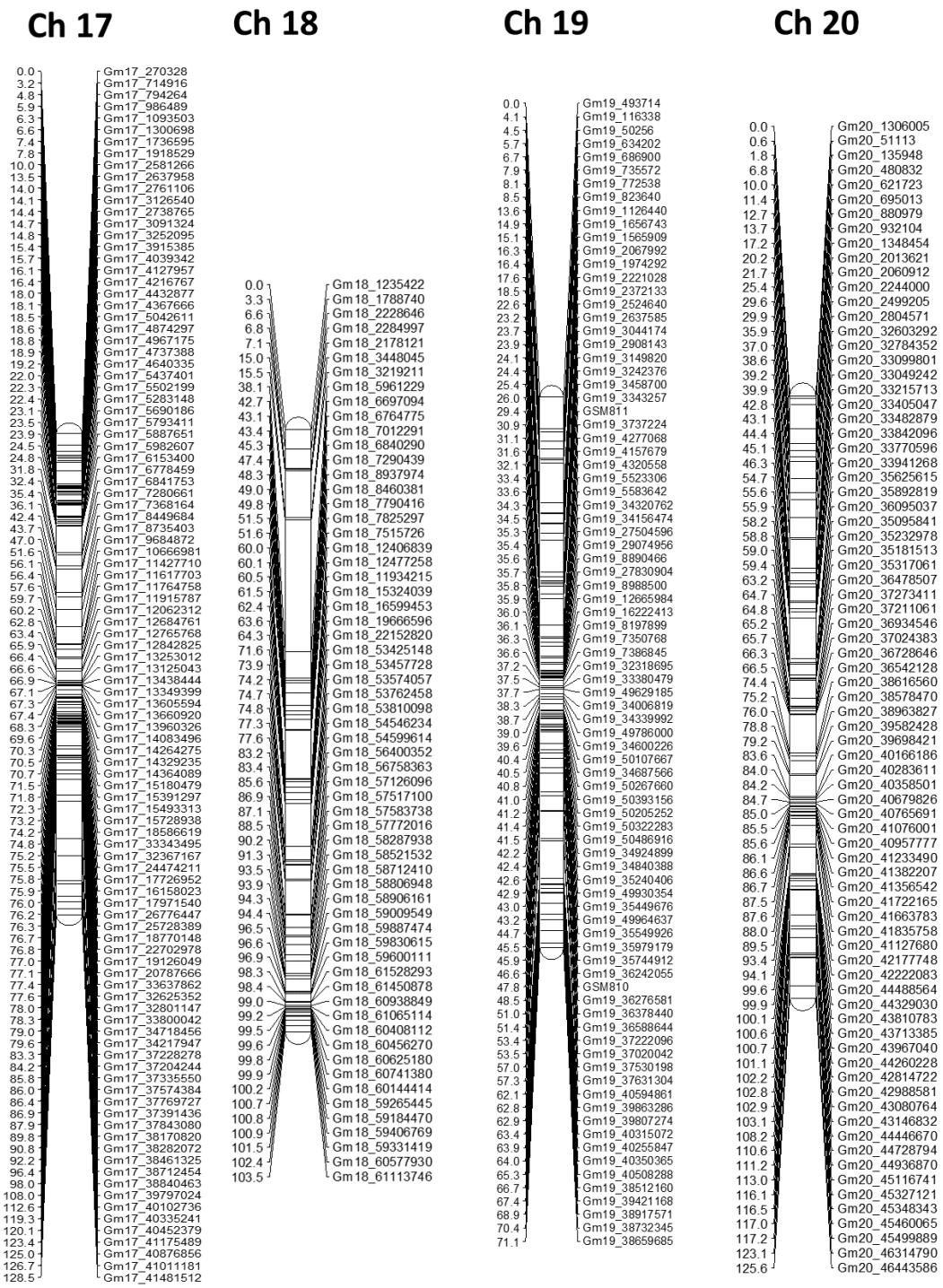


Fig S2.2. Linkage map of the recombinant inbred line population derived from 'Bossier' × PI 567295 using SoySNP6K Infinium Chip and KASP SNP markers and positions of QTLs on Chromosomes (Chrs) 6, 8, and 10 controlling HG type 0 (red) and 1.2.- (green) resistance. Positions of QTLs and confidence intervals were displayed with boxes.

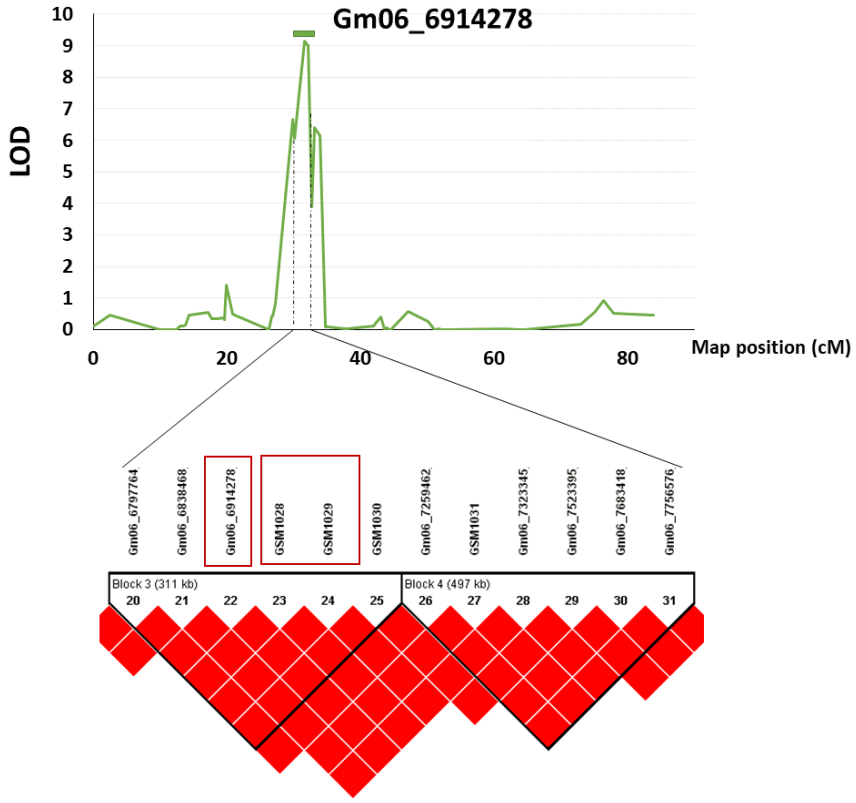


Fig S2.3 Haplotype block of the QTL detected on chromosome 6 in the interval of Gm06_6797764 and Gm06_7323345 in an $F_{5:6}$ RIL population derived from 'Bossier' × PI 567295. The plot was generated using Haploview 4.1. The red block color indicates strong linkage with D' = 1. Two KASP SNP markers GSM1028 and GSM1029 were located within the exon of *Glyma.0690300* that was linked by Gm06_6914278.

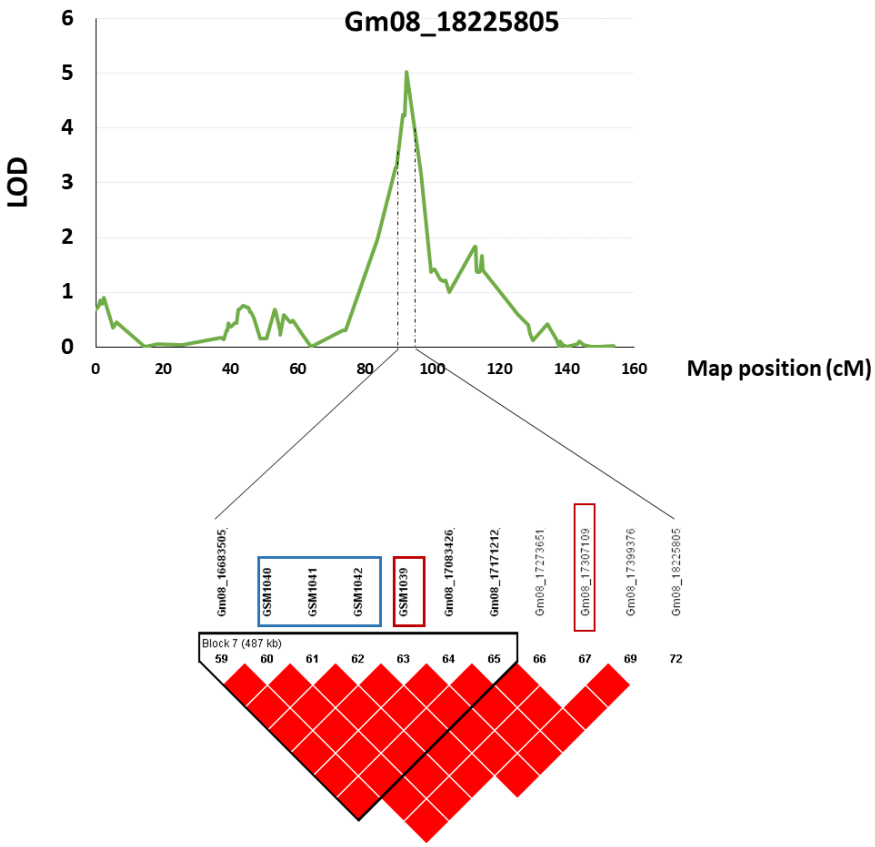


Fig S2.4 Haplotype block of the QTL on chromosome 8 associated with resistance to HG type 1.2.-. The dark red block indicates high linkage disequilibrium value (D'). GSM1040, GSM1041 and GSM1042 are non- synonymous SNPs of *Glyma.08207600*; the GSM1039 is located within the 6th exon of *Glyma.08g210000*.

CHAPTER 3

SOYBEAN MICROBIOME COMPOSITION AND ITS IMPACT ON THE HOST PLANT RESISTANCE TO SOYBEAN CYST NEMATODE[2]

[2] D. T. Tran, M. Mitchum, J. Wallace, and Z. Li. To be submitted to *Scientific Reports*.

Abstract

Microbial communities play a significant role in the growth and development of plants, including plant immunity and decomposition of complex substances to absorbable nutrients. Hence, utilizing beneficial microbes becomes a promising strategy for optimization of plant cultivation. The objective of this research was to explore the root bacterial profile across different soybean genotypes using 16S rRNA sequencing. Soybean genotypes with soybean cyst nematode (SCN) susceptible and resistant phenotypes were grown at Plains, GA in two years and in a greenhouse with three replicates per genotype. Bulked soil, rhizosphere and root samples were collected from each replicate. Sequencing of the bacterial 16S gene indicated that the bacterial profile of soybean root and soil samples overlapped but were different in bacterial communities. Bacterial phyla *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* dominate the soybean root enriched microbiota. Samples collected in the field condition over two years and different sampling time points in the greenhouse have large effects on the structure of the microbiome. However, the host genotypes have small, but significant effects on the diversity of root microbiome under SCN pressure at 10 days after inoculation, possibly representing the beneficial bacteria related to SCN resistance.

Keywords: soybean microbiome, soybean cyst nematode, amplicon sequencing, 16S rRNA

Introduction

It is well known that plants and microbes can possess a mutual relationship. Nitrogen fixation bacteria belonging to the *Bradyrhizobium* genus are well known examples of beneficial bacteria that convert nitrogen to useful ammonia to support legumes' growth and development (Siqueira et al., 2014). Numerous bacteria in the *Actinobacteria* phylum can synthesize indole-3-acetic acid (IAA) or auxin- mimicking modules that stimulate plant growth and development,

especially under drought stress (Yandigeri et al., 2012; Duca et al., 2014). Beneficial bacteria might also act as protectors against pathogens. Numerous bacteria have been developed as biocontrol agents for pathogen management such as *Bacillus licheniformis* for fungal species causing leaf spot and blight diseases, and *Pseudomonas chlororaphis* strain 63-28 for wilt diseases and root rots (Choudhary and Johri, 2009). The known pathogen defense mechanism called induced systemic resistance (ISR) was shown as being triggered by root microbiomes. One of these bacteria was the *Pseudomonas aeruginosa* PM12 that was found in healthy tomato roots helping to combat *Fusarium* wilt (Fatima and Anjum, 2017). Bacteria can colonize every part of the plant including the exterior of above-ground parts (phyllosphere); the interior parts (endosphere); root surface (rhizoplane); and rhizosphere region (defined as a narrow region of adjacent soil that interacts with the plant root) (Compant et al., 2019). Root microbiomes have been characterized for plants including *Arabidopsis thaliana* (Lundberg et al., 2012; Bergelson et al., 2019; Beilsmith et al., 2021), barley (Bulgarelli et al., 2015; Yang et al., 2020a), canola (Lay et al., 2018; Floc'h et al., 2020), rice (Edwards et al., 2015; Wang et al., 2020), maize (Wang et al., 2017; Beirinckx et al., 2020), wheat (Rascovan et al., 2016; Chen et al., 2019; Si et al., 2021) and soybean (Rascovan et al., 2016; Wang et al., 2017; Yang et al., 2020b). These studies have demonstrated that the distribution of the microbiomes differs between plant species and can change with different environmental conditions and plant health. One study using the model plant *Arabidopsis thailiana* that was grown under a controlled condition with natural soils reported that the host genotype had a small effect on the root microbiomes (Lundberg et al., 2012). Furthermore, differences in the abundance of microbiomes on three potato genotypes were detected mainly for species belonging to *Pseudomonales, Streptomycetaceae,* and *Micromonosporaceae,* which might be related to plant-pathogen management (Weinert et al.,

2011). Understanding interactions between the microbiomes and their host plants and identifying the host plant genes associated with these interactions could be transformative in plant breeding and biotechnology.

Soybean [*Glycine max* (L.)] is one of the most economically significant crops in the world with important usage in livestock feed. It had the production of 112.5 Million Metric Tons in 2020 (Association, 2021). It is well known that soybean forms a symbiotic relationship with nitrogen-fixing rhizobia. Previous studies have found that soybean roots selectively release compounds into root exudates including isoflavones and soyasaponin Bb which might attract microbiomes (Fujimatsu et al., 2020). The soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is one of the most damaging pests in soybean production, causing yield loss of ~ 3.6 Million Metric Tons in 28 soybean producing states in the U.S. and Ontario, Canada in 2014 (Allen et al., 2017). A report in 2021 showed that SCN has been found in all soybeans-producing states across the US (Tylka and Marett, 2021). Planting resistant cultivars is the common strategy to reduce SCN damage. Currently, there are two known resources of resistance: 'Peking' and PI 88788 (Mitchum, 2016). 'Peking' carries two major quantitative trait locus (QTL) alleles named *rhg1-a* (Reaction to *Heterodera Glycines*) on Chromosome (Chr) 18 and *Rhg4* on Chr 8 while PI 88788 requires only *rhg1-b* on Chr 18. 'Peking' and PI 88788 have been proved to have different responses to SCN due to different resistance alleles and copy numbers at the *Rhg1 locus* (Cook et al., 2012). 'Peking' reaction to SCN is faster than PI 88788 (Mitchum, 2016).

Seed treatment for controlling SCN has been used since late 1990s (Munkvold, 2009) in soybean production and is supplemented in SCN management with crop rotation and use of resistant cultivars (Thapa et al., 2022). Due to the cost and effects on environment, nematicides have not been used commonly (Desaeger et al., 2020) *Pasteuria nishizawae*, an obligate parasite

in SCN was firstly reported that could reduce the SCN population density (Noel and Stanger, 1994). This bacterium infects the nematode body and takes nutrients from the nematode while it reproduces its cells (Noel and Stanger, 1994). This bacterium was developed as a commercial product, "Clariva Complete Beans (CCB)" for seed treatment by Syngenta Crop protection (Syngenta Corporation, Wilmington, DE, U.S.). CCB could significantly increase yield of 0.34 tonnes/ha in a two-year experiment in Iowa (Bissonnette et al., 2018). The bacterial strain, *Bacillus firmus* I-1582, has been shown to possess a great potential in reducing SCN damage (Beeman and Tylka, 2017). This strain, originally isolated from soil that was collected in Israel, and has been exploited in a commercial seed treatment, Poncho/ VOTiVO by BASF Corporation (Research Triangle Park, NC, U.S). Besides, *Pseudomonas spp.*, *Paenibacillus spp.*, and *Streptomyces spp.* were also potential for SCN management in a greenhouse screening test (Tian and Riggs, 2000).

It even has been postulated that plants actively recruit beneficial bacteria in the soil to combat pathogen's attacks. Hence, we hypothesized that soybean roots might recruit beneficial bacteria from the soil to suppress the soybean cyst nematode. The knowledge about bacterial microbiota related to SCN is still limited. Therefore, determining the structure and assembly of the bacterial communities in soil, rhizosphere, and soybean roots that are related to soybean cyst nematodes can provide a better understanding of the promising beneficial bacteria and factors affecting bacterial assembly. Two field tests in Plains, GA in 2018 and 2019 and a greenhouse test in 2020 with and without SCN suppression were conducted and combined with high throughput sequencing of 16S rRNA gene amplicon to characterize the bacterial communities in soil, rhizosphere, roots, and factors affecting to assembly.

Materials and methods

Materials and experimental design

Fourteen soybean genotypes were selected to represent three SCN reaction groups: 'Peking'-type resistant, PI 88788-type resistant, and susceptible groups (Table 3.1). These genotypes are in the maturity groups (MG) III to VIII. Plants were grown in the Southwest Georgia Research and Education Center in Plains, GA in 2018 and 2019. Each genotype was planted in a plot consisting of four rows with 76.2 cm row spacing and 4.9 m row length, and plots were arranged in a randomized complete block design with three replicates per genotype.

After approximately two months of growth, until the beginning of flowering, roots were carefully removed from the soil using a shovel. The roots of five random plants were sampled within two middle rows of each plot to avoid border effects. All samples were chilled on ice immediately after collection and then stored at -20^{o}C until samples were processed within 48 hours. In addition, the soil was sampled at multiple sites in the field and mixed well into one bulked sample. It was analyzed in the University of Georgia (UGA) Plant Disease Clinics to diagnose nematodes.

To determine whether resistant and susceptible soybean genotypes have different microbiome profiles after SCN attack, a time series test was performed at UGA Nematology Greenhouse in Athens, GA. The greenhouse experiment was arranged in a randomized complete block design with two inoculation treatments: SCN HG type 0 (race 3) and water. For SCN treatment, each plant was inoculated with 1,500 eggs on the same day of planting. Topsoil from the UGA Iron Horse Farm in Watkinsville, GA was collected using shovels. The field soil was mixed with sterile sand with a ratio of 1:2. The soil-sand mixture was placed into 6-inch length pipes that were then placed into a 10- inch crock (19 pipes each). Crocks were placed and

maintained in water baths with a controlled temperature of 27^0C. The soil was also sent to the UGA Extension Nematology to diagnose nematodes. Three soybean genotypes including 'Woodruff' ('Peking'-type resistant), G00-3080 (PI 88788-type resistant), and 'Lee 74' (susceptible) were selected and used in this study (Table 3.2). Seeds from these genotypes were pre-germinated two days before planting. Samples (bulked soil, rhizosphere, and root) were collected from three replicates of each genotype within each treatment at three-time points: 10, 28, and 42 days after inoculation (DAI), respectively. The 10 DAI was the estimated time for early development of SCN including establishing a feeding site and a starting point for plants to respond; 28 DAI was the finish date of 1^{st} life cycle of SCN, and 42 DAI was expected in the second life cycle. The roots of three plants for each genotype were combined, representing one replicate of tissue for downstream sample processing. One blank control (only soil) was also included in each crock. In addition, three genotypes were also planted with three replications per genotype to verify SCN phenotypes. Roots from these plants were washed at 28 DAI and cysts on the roots for each genotype were then counted under a microscope. Four HG type indicator lines were also included to verify that the HG type 0 population was used in this test.

Sampling and processing

The protocol of rice microbiomes reported by Edwards et al. (2018) was followed with minor modifications in collecting and processing samples. Three sample types--bulked soil, rhizosphere, and roots--were collected. Bulked soil was excess soil collected from roots by manually shaking the roots, leaving approximately 1 mm of soil still attached to the roots which were considered as the rhizosphere. Then, 0.25 g of bulked soil was transferred to a 2 mL tube and stored at -80^{o}C for DNA extraction. The rhizosphere soil was collected after shaking the roots to remove excess soil. During the sampling process of the rhizosphere, approximately 1 g

of each root sample was placed in a 50 mL sterile Falcon tube containing 20 mL of autoclaved Phosphate Buffered Saline (PBS) 1X solution. The tubes were then vortexed for 10 s to wash out the rhizosphere from the root. The root was then removed from the tube using flame sterilized forceps and placed in a new sterile Falcon tube for further use. Then, the tube containing the extracted rhizosphere was centrifuged at 3,600 rpm for 30 min until the soil formed a pellet of rhizosphere in the bottom of the tube and supernatant was then removed. The rhizosphere pellets were stored at -80^{o}C until DNA extraction. After separation from the rhizosphere, any remaining soil on the root samples was removed by washing with autoclaved phosphate-buffered saline (PBS) until the buffer solution ran clear. After washing the roots were cut into pieces approximately 3 mm in size. To further process the root samples, 0.25 g of roots was placed into a 2 mL tube and stored in a -80^{o}C freezer for overnight to let roots become brittle, and then roots were grounded with beads using a SPEX Sample Prep Geno Grinder (Metuchen, New Jersey, U.S) for DNA extraction.

DNA extraction and library preparation

Pellets from the rhizosphere soil samples were dissolved in 500 µL of PBS buffer in Falcon tubes and then broken down using a pipette tip. The tube was vortexed to suspend the rhizosphere soil in the PBS buffer, and a 500 µL mix of soil and PBS buffer was transferred into a 2 mL tube for DNA extraction. DNA from all samples including bulked soil, rhizosphere, and grinded roots were extracted using a QIAGEN DNEasy PowerSoil kit (QIAGEN, Germantown, MD, U.S) and eluted in 50 µL of elution buffer. The concentration of DNA samples was measured using Qubit **(**Thermo Fisher Scientific, Inc., Waltham, MA, U.S**)** and then diluted to 2 ng/µL for library preparation.

The 16S rRNA gene consists of nine hypervariable regions flanked by nine conservative regions. The nine hypervariable regions are named as V1 through V9. Previous study reported that the V4 region was one of the most reliable and common regions for assigning taxonomy (Peiffer et al., 2013; Wasimuddin et al., 2020).Therefore, targeted profiling of the samples was carried out by sequencing the V4 region of the 16S rRNA. Amplification of the V4 region was performed using the universal primer pair (515F and 806R) and a two-step PCR method with a modified protocol based on work by Caporaso et al. (2011) and Tinker et al. (2016). PCR reaction mixes were made using Phusion High Fidelity (HF) DNA Polymerase (Thermo Fisher Scientific, Inc., Waltham, MA, U.S). Two PCRs were performed based on the manufacturer's protocol of HF Phusion and Tinker et al. (2016). The 1st PCR mix had a total volume of 13 µL consisting of 7.6 µL of dH20, 3 µL of HF Phusion, 0.75 µL of 515F and 806R primers, 0.45 µL of MgCl2, 0.3 µL of 10 nM dNTP, 0.15 µL of HF polymerase and 2 µL of DNA. The condition of the 1st PCR was as follow: 98°C for 30 s, followed by 15 cycles at 98°C for 10 s, 52°C for 30 s, and 72°C for 30 s, with a final extension step at 72°C for 5 min for the initial V4 region amplification. The 2nd PCR mix was performed in a reaction volume of 21 µL containing 10.2 µL of dH2O, 6 µL of HF Phusion, 1.5 µL of each forward and reverse primers including dual barcodes and Illumina adapters, 0.9 µL of MgCl2, 0.6 µL of 10nM dNTP, 0.3 µL of HF polymerase and 9 µL of the 1st PCR product. The 2nd PCR condition consisted of 98°C for 30 s, followed by 4 cycles at 98°C for 10 s, 52°C for 30 s, and 72°C for 30 s, followed by 6 cycles at 98°C for 10 s and 72°C for 1 min, concluding with a final extension at 72°C for 5 min.

The final products were run on a 2% agarose gel using TAE buffer to ensure the amplification (~400 bp). PCR products were then purified with a EZ Cycle Pure kit (Omega Bio-tek, Inc., Norcross, GA, U.S) according to the manufacturer's protocol. Samples were eluted in a

30 µL of elution buffer, and the concentration was measured with Qubit (Invitrogen. Life Technologies, CA, U.S), and A260/280 with Nanodrop (Thermo Fisher Scientific, Wilmington, DE, U.S). Prior to sequencing, extra quantification, normalization, and pooling were performed. Forward and reverse dual barcodes were used to pool up to 210 samples together into one sequencing run (Table S3.1 & S3.2)

Sequencing

Paired end (2 x 250 bp) sequencing was carried out on an Illumina Miseq platform (San Diego, CA, U.S) at the UGA Genomics and Bioinformatics Core (GGBC). The demultiplexed sequencing reads were processed with MOTHUR v 1.45.3 (Schloss et al., 2009) commands following the MiseqSOP (Schloss et al., 2009) (Table S3.3). Briefly, paired end sequencing reads were merged and only those with a maximum homopolymer length of 8 bp that is a sequence of consecutive and identical nucleotides and a maximum of 275 bp were kept in order to remove any reads significantly longer than the V4 region of 16S gene (250 bp) due to bad assembly. Then, sequences were aligned to the SILVA reference database (Release 132, (Quast et al., 2013). Chimeras were detected and removed by the uchime algorithm using the chimera.uchime command (Edgar et al., 2011). After chimera removal, taxonomic classification of samples was run using the Ribosomal Database Project (RDP) reference database (Cole et al., 2014) version 18 containing 20,712 bacterial 16S rRNA gene sequences. Sequences that were unclassified or identified as chloroplast, mitochondria, and eukaryote were removed using the remove.lineage command (Schloss et al., 2009). Sequences that had greater than 97% similarity were clustered into the same Operational Taxonomic Unit (OTUs).

Data analysis and visualization

Data analysis and visualization were performed in R using phyloseq Version 1.36.0 (McMurdie and Holmes, 2013), dplyr Version 1.0.6 (Wickham et al., 2021), tidyr Version 1.1.3 (Wickham, 2021) and ggplot2 packages Version 3.3.3 (Wickham, 2016). Samples that produced <1,000 reads and taxa that had low abundance OTUs (<5 total counts) were removed. Bar plots for phylum having 1% relative abundance was created in ggplot2 package Version 3.3.3 (Wickham, 2016). Prior to alpha and beta diversity analysis, the dataset was normalized by rarefying to the smallest size with 1,000 reads to lessen the variation in the number of reads among samples. Alpha diversity was calculated using four metrics: Observed OTUs, Chao1, Shannon, and Inverse Simpson. Chao1 index was used to estimate species richness, which was a measurement of the expected OTUs in the sample given all the species are identified (Chao, 1984). Shannon Index estimates the overall richness and the uniformity between the taxa present in samples (Spellerberg and Fedor, 2003), while the Inverse Simpson focuses more on evenness (Simpson, 1949). Beta diversity was used to evaluate the difference in microbial communities among samples. It was calculated using relative abundance (Bray-Curtis) measures (Bray and Curtis, 1957). Bray Curtis provides a measure of community composition differences among samples based on their OTU counts, regardless of taxonomic assignment.

Data was visualized using non-metric multidimensional scaling (NMDS) (Kruskal, 1964) and the Canonical Analysis of Principal coordinates (CAP) (Anderson and Willis, 2003). Each sample was displayed in the result plots as a point and the clusters indicated the similarity and dissimilarity of samples. The NMDS relies on the rank order distance for ordination with the stress value that represents the difference between distance in the reduced dimension compared to the complete multidimensional space (Kruskal, 1964). The stress values < 0.05, < 0.1, > 0.1, >

0.2 indicate an excellent, great, good and fair representation in reduced dimensions, respectively, while stress value > 0.3 means a poor representation, which requires more dimensions to visualize data (Kruskal, 1964).

To evaluate the effects of years, locations, genotypes, and phenotypes, data was analyzed with a permutational multivariate analysis of variance (PERMANOVA) using a vegan package (Oksanen et al., 2020). Lefse (Linear Discriminant Analysis Effect Size) was used to determine the OTUs most significantly explaining the difference between groups by using a Kruskal-Wallis rank sum test with a significant alpha of 0.05 and an effect size threshold of 2 (Segata et al., 2011).

Results

16S rRNA sequences

To investigate the structure of soybean root bacterial microbiomes, bulked soil, rhizosphere, and roots samples were collected from 13 and 14 soybean genotypes in Plains, GA in 2018 and 2019, respectively (Table 3.1) for a total of 247 samples. The V4 region of 16S rRNA was sequenced by paired end 250 bp sequencing on Illumina MiSeq. A total of 10,322,031 high quality sequences were obtained after quality control and clustered into OTUs defined by 97% similarity using MOTHUR (Schloss et al., 2009). Samples with less than 1,000 reads were discarded, so a total of 234 samples with 9,210,821 reads were used for further data analyses. Low abundant OTUs with less than five reads, unclassified OTUs, and chloroplasts were removed, resulting in 2,532 OTUs for further analysis.

In the greenhouse test in 2020, 210 samples were collected including bulked soil, rhizosphere, and roots of three genotypes at three time points and unplanted soils from both treatments water and HG type 0 (SCN race 3). A total of 5,946,972 high quality sequences were

obtained after quality control and clustered into OTUs defined by 97% similarity using MOTHUR commands. After removing samples with <1,000 reads, a total of 195 samples with 5,918,834 reads were used for further data analyses. A total of 8,561 OTUs were available for data analysis after removing low abundance OTUs, unclassified OTUs and chloroplasts.

Core bacterial taxa in different root compartments of soybean, grown in field and greenhouse conditions

Both alpha (α) and beta (β) diversity analyses revealed that microbiomes were variable among three compartments in both field and greenhouse conditions. Alpha diversity which measures within sample diversity showed a diversity gradient from the root to both the rhizosphere and bulk soil (Fig 3.1 A). In field conditions, root communities had the lowest α-diversity and the bulk soil had the highest α- diversity. The mean α- diversity was highest in bulk soil (Observed OTUs: 820; Shannon: 4.9; and InvSimpson: 56.8); and was lowest in roots (Observed OTUs: 270; Shannon: 3.4; and InvSimpson: 13.8). The difference in α- diversity among the three compartments were statistically significant (ANOVA: $P < 2e\text{-}16$). Moreover, NMDS coupled with PERMANOVA (Bray Curtis distance) showed that bacterial community structures in bulked soil, rhizosphere and roots were significantly different from each other ($P = 4.07e\text{-}10$) (Fig 3.1 B).

Both alpha and beta diversity patterns in the greenhouse test were similar to those in the field tests. The mean α- diversity of bulked soil was the highest while the mean of the root samples was the lowest. The difference in α- diversity among three compartments was also significant (ANOVA: $P < 2e\text{-}16$) (Fig.3.2 A). NMDS with Bray Curtis distance showed three clear clusters for each compartment across three sampling times and were strongly separated from each other ($P < 3.04e\text{-}06$) (Fig. 3.2 B). Different sampling times also revealed the

difference in bacterial compositions based on the β diversity. Although bulked soil samples did not show significant separations among three time points in greenhouse, rhizosphere and root samples that were collected at 10 DAI were separated from those collected at 28 and 42 DAIs. However, these two sample types did not have significant differences in bacterial composition at both 28 and 42 DAIs.

Core soybean root microbiome

Field: To discover taxonomic structure, core bulked soil, rhizosphere and root samples were defined as those OTUs that were present in at least 80% of samples in both years. Across all samples, 404, 373, and 82 OTUs of bulked soil, rhizosphere, and root samples of field tests, respectively fit within this category. Six core phyla were identified in root samples including three predominant phyla: *Proteobacteria, Actinobacteria* and *Bacteroidetes* (Fig. 3.3). Together these three phyla accounted for 98% and 99% of all bacteria in 2018 and 2019, respectively. Three other phyla consisted of *Firmicutes, Nitrospirae* and *Armatimonadetes*, presenting in smaller amounts of less than 1% of the total dataset but presenting in more than 80% of all root samples in field tests. The *Novosphingobium* genus was the largest genus in *Proteobacteria* phylum in 2018 (11%) and the second most abundance in 2019 (15%), after *Bradyrhizobium.* The high relative abundance of *Bradyrhizobium* across root samples was observed in both years because it is the best known as root nodulating bacteria of legumes in *Proteobacteria* phylum. In 2018, this genus accounts for 8% of all bacteria, while it accounted for 17% in 2019. Another N-fixing bacterium, *Rhizobium* genus in *Proteobacteria* phylum, was also identified in root samples in both years with smaller amounts (4% in 2018 and 3% in 2019).

The *Actinobacteria* was the second most dominant phylum in both years with over 30% of total bacteria represented. Within the *Actinobacteria* phylum, the *Streptomyces* genus had the

largest representation, accounting for 16% and 21% of whole root samples in 2018 and 2019 respectively. The *Bacteroidetes* was the third most abundant phylum in both years with 10% and 7% of all bacteria. Two genera: *Niastella* and *Chitinophaga* were the largest in *Bacteroidetes* phylum in 2018. These two genera together accounted for 50% of whole bacteria of *Bacteroidetes* phylum. In 2019, *Niastella* was also highly present in root samples with 22% of all bacteria in Bacteroidetes while the proportion of *Chitinophaga* was reduced. The *Sediminibacterium* genus was the largest genus, accounting for 30% of *Bacteroidetes* in 2019.

Greenhouse: The bacterial sequences were affiliated with three core dominant phyla including *Actinobacteria, Bacteroidetes* and *Proteobacteria* (Fig S3.2). At all sampling timepoints, *Proteobacteria* was the predominant phylum, accounting for over 80% of all observed bacteria. Within the *Proteobacteria* phylum, bacteria from the *Bradyrhizobium* genus were especially prevalent, accounting for over 30% of total bacteria at all three timepoints. The result agreed well with root samples collected in field conditions. The *Novosphingobium*, which was the most prevalent genus in root samples collected in the field, had been less present under a greenhouse condition, accounting for 1 to 5% of the whole root dataset at different time points and treatments. Different sampling times showed a difference in abundance of another genus. The *Massilia* genus was the second most abundant in root samples collected at 10 DAI which belonged to *Betaproteobacteria* class (*Proteobacteria* phylum) with 10 and 15% under HG type 0 and water treatment, respectively. At 28 DAI, the *Streptomyces* genus from *Actinobacteria* phylum was the second most after the *Bradyrhizobium* genus with 11% of total bacteria. At 42 DAI, the *Burkholderia* genus from *Proteobacteria* phylum was the second most (8%) under both water and HG type 0 treatments after the *Bradyrhizobium* genus. NMDS analysis with Bray Curtis distance also showed the difference among sampling time points. Root samples that were

collected at 10 DAI separated strongly from samples collected at 28 and 42 DAI (Fig. 3.2 B). It indicated that microbiota shift in the root with different sampling times, and the change in microbiome might have minor change at the later stages (28 and 42 DAI) because the two samples collected at these time points did not separate strongly in NMDS analysis.

Core rhizosphere microbiome

Field: Seventy-six rhizosphere samples from both years had 373 core taxa, which represented more than 80% of samples. Five phyla including *Proteobacteria, Nitrospirae*, *Firmicutes, Bacteroidetes* and *Actinobacteria* were highly present in the core microbiome (Fig. 3.4). Similar to root samples, *Proteobacteria* was the most predominant phylum present in the rhizosphere, accounting for over 60% of all bacteria in the rhizosphere each year. The second most abundant phylum was *Actinobacteria* with 17% and 14% in 2018 and 2019, respectively; and the third most abundant phylum was *Bacteroidetes*, which was most abundant in root samples. At genus level, the *Sphingomonas* genus in *Proteobacteria* phylum was the most predominant with 16% of bacterial communities in 2018 and 17% in 2019. Following this genus *Novosphingobium* from the same phylum had high abundance in the rhizosphere with 9% in 2018 and 10% in 2019. Within *Actinobacteria* phylum, two most abundant genus were *Gaiella* and *Streptomyces* in both years, while *Gaiella* had lower abundance in root samples with less than 1%. Within *Bacteroidetes* phylum, three genera: *Chrysobacterium*, *Niastella* and *Chitinophaga* had the highest proportion. NMDS plot revealed two groups of rhizosphere samples in two years were separated (Fig. 3.1 B). Two genera: *Bradyrhizobium* and *Rhizobium* were shown to be significantly and differentially represented between two years by the LEfSe analysis with a larger abundance in 2019.

Greenhouse: Core taxa of rhizosphere included 202 taxa in 49 samples with six phyla: *Actinobacteria, Bacteroidetes, Firmicutes, Nitrospirae, Proteobacteria, Verrucomicrobia* that were more diverse than those from root samples (Fig S3.2). Across sampling time (10, 28 and 42 DAIs), the *Proteobacteria* phylum was the most represented in rhizosphere samples with over 80% of bacterial communities. Within *Proteobacteria* phylum, *Betaproteobacteria* had a higher abundance across three-time samplings (31- 57%), than in root samples (20-25%) while *Actinobacteria* had a higher abundance in root samples than in the rhizosphere. This result differs with rhizosphere samples collected from field conditions. In the field conditions, *Betaproteobacteria* had a lower proportion than *Alphaproteobacteria* with only 12% in 2018 and 8% in 2019. The genus *Pseudoduganella* from *Betaproteobacteria* was the most predominant genus at 10 DAI, while *Burkolderia* genus from the same class was highest at 28 and 42 DAI. Based on the LEfSE analysis, the genus *Pseudoduganella* was significantly enriched in 10 DAI compared with other two sampling times. The NMDS plot also showed microbial communities in the rhizosphere at 10 DAI were strongly separated from 28 and 42 DAI (Fig. 3.2 B).

Core bulked soil microbiome

Field: Three hundred and sixty-nine taxa, representing over 80% of bulked soil were core taxa. They were from 10 phyla and *Proteobacteria, Actinobacteria, Bacteroidetes and Firmicutes* were the most represented phyla in bulked soil samples across two years. Two genera: *Sphingomonas* and *Bradyrhizobium* had high proportions in both years from 8 to 12%. The *Proteobacteria* was the most abundant with higher than 50% of whole bacteria in bulked soil, but a lower proportion than root and rhizosphere samples. The *Actinobacteria* phylum had higher abundance in soil than in rhizosphere samples but lower than in root samples. However, some phyla including *Firmicutes, Nitrospiraceae* had a higher proportion in bulked soil samples

than root or rhizosphere samples, so the number of core phyla in bulked soil were more than other two types of samples. Based on NMDS analysis, bulked soil samples collected from two years were separated (Fig. 3.1 B), and the LEfSE analysis revealed a list of significant taxa significantly associated with the difference in structure of bacterial communities (significant level). Two most significant genera with LOD score > 5 were *Microvirga* in 2018 and *Sphingomonas* in 2019. Both of them were from *Proteobacteria* phylum.

Greenhouse: Evaluating taxa showed 201 core taxa from seven phyla presenting over 80% of samples, including: *Acidobacteria, Actinobacteria, Bacteroidetes, Firmicutes, Nitrospirae, Proteobacteria,* and *Verrucomicrobia* (Fig S3.2). Of these seven phyla, the *Proteobacteria* phylum was most abundant, which was similar to the root and rhizosphere samples from 49 to 58% of all bacterial communities across DAIs. The *Actinobacteria* phylum ranked second in abundance with 24 to 29%, higher than rhizosphere and root samples from greenhouse. The *Bacteroidetes* phylum was the third most abundant with less than 10% of all bacterial communities. At genus level, the *Sphingomonas* from *Proteobacteria* phylum was highly presented and constant at all three time points with 13 and 14%. The result agreed with bulk soil samples collected from the field. After the *Sphingomonas*, the *Gaiella* genera from *Actinobacteria* phylum was also highly represented across all sampling time points with the highest proportion at 10 DAI and the lowest at 42 DAI. The LEfSE analysis revealed significant differences in taxa abundance in bulked soil samples among three time points. At genus level, the *Pseudoduganella* was more abundant in soil samples at 10 DAI while *Nitrospira* was found to be more abundant at 28 DAI and *Ralstonia* at 42 DAI (LOD score > 5).

Microbial communities of resistant and susceptible genotypes under SCN pressure

In field conditions, all three sample types including bulked soil, rhizosphere and root samples did not show significant differences between SCN resistant and susceptible genotypes based on alpha and beta diversity. The soil sample was diagnosed for nematode profile, but no SCN was detected and only two lesion nematodes were found in tested soil. It might be a reason why no significant difference in bacterial composition was found between two contrasting genotypes. Therefore, a greenhouse test was set up with two treatments: mock control (water inoculation) and HG type 0 inoculation; and samples were collected at three different time points. At 28 DAI, roots were collected for evaluating SCN response. 'Woodruff' possessing *rhg1-a Rhg4* resistance genes *(*'Peking'-type resistance*)* had a female index of 0.5%, while 'G00-3880' carrying only *rhg1-b* resistance gene *(*PI 88788-type resistance had a female index of 1.33% when using 'Lee 74' as a susceptible check.

No difference in microbial communities in root samples was observed among three DAIs (*P = 0.3, 0.7* and *0.6* for 10, 28, and 42 DAIs, respectively) for the water treatment based on PERMANOVA. Similar to root samples with water inoculation, no difference in microbial communities was observed among 25 rhizosphere and 25 bulked soil samples from these genotypes (*P > 0.2*). With the HG type 0 treatment, no difference in microbial communities in bulked soil samples was found at three timepoints (*P = 0.1, 0.17*, and 0.29 for 10, 28 and 42 DAIs, respectively) but significant difference among three genotypes was observed in rhizosphere samples at 10 DAI and 42 DAI (*P = 0.02* and 0.009, respectively). No statistical difference in rhizosphere samples was observed among three genotypes *(P = 0.3*) at 28 DAI.

Analyzing the root samples revealed a significant difference in microbial communities among three genotypes at 10 DAI (*P = 0.01*); but no significant difference at 28 and 42 DAIs (*P*

= *0.7* and *0.4*) based on PERMANOVA analysis. Therefore, the 10 DAI was a good time point to explore the difference in microbial communities in both rhizosphere and root samples related to SCN response. At 10 DAI, the LEfse analysis revealed a list of taxa that contributed to significant differences in the microbial communities in rhizosphere samples (Table 3.3). The relative abundance of *Bradyrhizobium* and *Neochlamydia* genera were enriched in 'G00-3880'; while *Armatimonas* genus from the *Armatimobadaceae* family was increased in 'Woodruff'. No genus was found to be significantly differentially enriched in 'Lee 74'.

In total, eight root samples at 10 DAI had 1,990 taxa that mainly came from four phylum: *Actinobacteria, Bacteroidetes, Firmicutes* and *Proteobacteria*. The *Proteobacteria* was the most prevalent phylum in all three genotypes. However, within *Proteobacteria*, the *Alphaproteobacteria* was more abundant in 'Lee 74' with 74% of whole microbial communities in root while 26 and 34% in 'G00-3880' and 'Woodruff', respectively. Relative abundance analysis at genus level compared the composition and proportion of classified bacteria genus of root samples between two treatments: Hg Type 0 and water treatment (Fig. 3.5). Sixteen dominant phyla (each with <2% of total number sequences in all samples) varied in relative sequence abundance in three soybean genotypes under two treatments. The *Bradyrhizobium* genus had the greatest relative sequence abundance in 'Lee 74' root under Hg type 0 treatment than water treatment with 61.6 and 33.5%, respectively. In 'Woodruff', the change of relative abundance of the *Bradyrhizobium* genus was small with 7 and 10% in water and HG type 0 treatments, respectively. However, this genus showed a trend of decreasing relative sequence abundance from 30.1% under water to 5% under HG type 0 treatment in 'G00-3880' roots.

The *Pseudoduganella* genus was shown being dominant in both SCN resistance genotypes: 'Woodruff' and 'G00-3880' under SCN pressure (15 and 16%) than water control

(5and 7%) while this genus decreased from 5% to nearly 0 between HG type 0 and water treatments in ‘Lee 74’ (Fig. 3.5). At the genus level, six dominant phyla were shown in ‘Lee 74’ while 14 and 13 phyla were indicated in ‘G00-3880’ and ‘Woodruff’ roots under HG type 0 treatment (Fig. 3.5). Similarly, α diversity using Shannon and InvSimpson measurement indicated that species richness of ‘Woodruff’ and ‘G00-3880’ were significantly higher than those of ‘Lee 74’ (*P = 0.01)* (Fig. 3.6).

However, based on α diversity, the difference in richness and diversity of bacteria between ‘Woodruff’ and ‘G00-3880’ were not significant (Fig. 3.6), which indicated that two types of resistance sources did not reveal different diversity levels of bacteria. LEfSe analysis revealed that four most abundant genera responded differently to these genotypes using a threshold LDA score of 3.5 and *P value* < 0.05 (Fig. 3.7). The abundance of *Ramlibacter* and *Mycobacterium* were increasing in ‘Woodruff’, and *Dyella* were enriched in ‘G00-3880’. The *Dyella* genus, from the *Proteobacteria* showed an increasing trend from bulked soil (0.02%) to rhizosphere (0.9%) and more enriched in root samples of ‘G00-3880’ (3%) under HG type 0 pressure. This genus was found with small amounts (0.4%) in raw soil without soybean plants. Under water treatment at 10 DAI, the *Dyella* genus was more dominant in ‘G00-3880’ root samples than ‘Woodruff’ and ‘Lee 74’.

The *Ramlibacter* genus from *Proteobacteria* phyla were more enriched in ‘Woodruff’ compared to other two genotypes in root samples. At 10 DAI, this genus was dominant in both two treatments: water and HG type 0 comprising 7% of total bacteria while small proportions at other DAIs were observed in both treatments (< 1% of all classified bacteria). Based on LefSe analysis, the *Mycobacterium* genus from *Actinobacteria* phyla was more enriched in ‘Woodruff’ than ‘G00-3880’ and ‘Lee 74’ in root samples. This genus also showed the different enrichment

between two treatments in ‘Woodruff’ roots. Under water treatment, this genus comprised of only 0.03% of total bacteria while it comprised of more than 2% of all bacteria under HG type 0 treatment

In both bulked soil and rhizosphere samples, ‘Woodruff’ did not show significant difference in amount of *Mycobacterium* between two treatments. Whereas in ‘Lee 74’, only *Sphingomonas* genus were more enriched with 3% of total bacteria, in contrast, it only comprised a small amount with nearly 0% of whole bacteria in root samples of ‘Woodruff’ and ‘G00-3880’. Under water treatment, this genus was not observed in ‘Lee 74’ roots. In bulked soil and rhizosphere samples of ‘Lee 74’ under both treatments, this genus also showed less enrichment with the proportion of less than 0.5% of total bacteria at 10 DAI.

Discussion

Microbial communities differed consistently among the sample types

The data presented here provides characterization of the microbiomes of soybean roots using 16S rRNA sequencing from field and controlled greenhouse studies. The compositions of three distinct compartments- bulked soil, rhizosphere and endosphere of soybean have been characterized to gain insights into the effects of environmental factors on each of these compartments especially associated with SCN resistance. The rarefaction curve plots represent different microbiota from different sample types (Fig. S3.1 and S3.2). The rhizosphere and bulked soil samples were more diverse than root samples as the number of species (OTUs) increased. However, most of the curves had not reached the plateau although they showed a rapid increase in the beginning and then slowing down. This suggested that sequencing depth, sampling efforts, and sample size could be further improved for a better representation of the microbial community by capturing some rare species.

In this study, three sample types were evaluated in both greenhouse and field conditions. Beta diversity analysis revealed that root microbiomes were variable among sample types (Fig. 3.1 B and 3.2 B). Similar to the reports for many other plant species such as *Arabidopsis* (Lundberg et al., 2012), microbiomes in three sample types bulked soil, rhizosphere and root in this study exhibited overlapping and also distinct. The enrichment and depletion effects within compartments from bulked soil to rhizosphere and roots have been observed in this study (Fig. 3.3; 3.4 and S3.2) and also reported in a large number of plant microbiome studies (Edwards et al., 2015; Liu et al., 2019).

Bacteria that were associated with SCN resistance

Thirteen and 14 soybean cultivars were planted in fields in two years: 2018 and 2019. These cultivars were divided into two groups based on their genotypes related to SCN response: SCN resistance and susceptible groups. Beta diversity showed SCN resistant and susceptible genotypes did not have significant difference in bacterial communities in all three sample types in field tests ($P = 0.8$ and 0.7 for 2018 and 2019 respectively) based on PERMANOVA and α diversity also indicated no significant difference in density level of bacteria. The reason might be that no SCN appeared in the field, so the plants might not recruit beneficial bacteria for combating SCN attack.

Therefore, to discover bacterial communities related to SCN response, the composition of bacteria of soybean with two inoculation treatments: HG type 0 and water was investigated in a greenhouse condition. It has been reported that SCN carries a rich variety of bacteria which could play a role in the ecology of SCN (Nour et al., 2003). Cysts collected from fields contained up to 30 phylotypes with the enrichment of *Lysobacter* and *Variovorax spp.* were observed under a microscope (Nour et al., 2003). Comparison between two treatments at 10, 28, and 42 DAIs

indicated a shift of microbial communities. Based on LefSe analysis. 28 genera were more enriched in three sample types under the HG type 0 treatment (Table S3.4). Genera from *Proteobacteria* phylum was abundant in all sample types inoculated with SCN (Table S3.4). The *Lysobacter* genus from *Proteobacteria* was highly abundant in rhizosphere samples at 28 DAI under the HG type 0 treatment based on LefsE analysis, The *Proteobacteria* phylum with *Alphaproteobacteria, Betaproteobacteria, Deltaproteobacteria and Gammaproteobacteria* classes has been reported as the dominant taxa in SCN cysts, which is consistent with the previous studies based on DGGE and cultivar-based approaches (Nour et al., 2003). A recent study also reported that this phylum was the dominant bacteria accumulating in cysts in SCN suppressive soil, followed by *Actinobacteria* and *Bacteroidetes* (Hussain et al., 2018). However, with unplanted pots being included in two treatments, no significant difference between two treatments were found ($P = 0.5$) in this study (Fig S3.4). This suggested that the shift of microbial community associated with SCN in soil might initiate in the presence of soybean plants and SCN as an obligate parasite might require a living host to maintain its growth, otherwise it might die due to lacking the host.

Fully differentiated bacterial communities were assembled in this study in the rhizosphere and root. No enriched taxa was found in 'Lee 74', the susceptible genotype in the rhizosphere compared to other SCN resistant genotypes at 10 DAI under HG type 0 treatment. The result agreed with α diversity analysis that revealed 'Woodruff' and 'G00-3880' were more diverse in bacteria than 'Lee 74' at 10 DAI (Fig. 3.6*).* The enrichment of some taxa in the rhizosphere and root samples suggested that soybeans might actively select a specific set of bacterial microbiomes. *Bradyrhizobium, Massilia, Pseudoduganella* genera were found to be more prevalent in roots with HG type 0 inoculation than those with water inoculation at 10 DAI.

Of them, the *Bradyrhizobium* was more represented in 'Lee 74' while *Massilia* and *Pseudoduganella* were highly abundant in 'G00-3880' and 'Woodruff' (Fig. 3.5). Both *Massilia* and *Pseudoduganella* belonged to the *Oxalobacteraceae* family and the role of these families for SCN resistance was unknown. The most widely studied bacteria that have biological control potential of plant parasitic nematodes are *Pasteuria, Pseudomonas, Rhizobium* and *Bacillus* (Tian et al., 2007). In our study, the prevalence of *Pseudomonas, Rhizobium* and *Bacillus* was significantly higher in resistant genotypes at 10 DAI, and *Pseudomonas and Rhizobium* at 28 DAI. The *Pseudomonas* from *Pseudomonadaceae* family has been known for antagonistic activities against ring nematodes and it might inhibit nematode infection at the juvenile stage (Kluepfel et al., 2002). The *Rhizobium* from *Rhizobiaceae* has also been shown to be associated with the suppression of *H. schachtii* (Yin et al., 2003).

At 10 DAI, the enriched taxa in two types of resistant plant: 'Peking' and PI 88788 were observed. The genus *Dyella* was increased in 'G00-3880' root than 'Woodruff' and 'Lee 74' (Fig. 3.7). *Dyella* was reported as one of dominant bacteria in pines infected with nematodes, but there was no report related to nematode resistance (Guo et al., 2020). Two genera: *Ramlibacter* and *Mycobacterium* were more abundant in 'Woodruff' that carries 'Peking'-type resistance (Fig. 3.7). The relationship between these genera with nematode resistance has not been reported. Comparing the proportion of these bacteria between two treatments of each genotype, the *Ramlibacter* genus was constant at 7% of total bacteria in both treatments while the *Mycobacterium* increased from 0.03 to 3% between water to HG type 0 treatments. In 'G00-3880', the *Dyella* genus was more enriched in root samples under HG type 0 (3%) than under water (0.9% of total classified bacteria). The *Bradyrhizobium* genus that is related to nodulation in soybean was found to be more enriched in 'Lee 74', the SCN susceptible genotypes (62%)

than in SCN resistance genotypes under HG type 0 treatment (only 6 to 10%). This genus was also reported to be more enriched in SCN susceptible soybean rhizosphere soil and root endosphere in SCN inoculated suppressive soil (Hussain et al., 2018). At 28 and 42 DAIs, the number of enriched taxa in the resistant genotypes were less than the samples collected at 10 DAI. Besides, there was no difference in 'Woodruff' at 28 and 42 DAI. Hence, we hypothesized that the life cycle of SCN also affected the structure and variation of bacterial communities in root samples. Around 10 DAI, resistant genotypes responded to SCN attacks. Based on the previous review, the cytoplasmic degeneration happens in 8 to 10 DAI with PI 88788-type resistance while 'Peking' has faster response with a rapid degeneration around 2 DAI (Mitchum, 2016). Hence, at 10 DAI in our test, resistant genotypes actively recruited bacteria from a soil pool to respond to SCN. After that, SCN juveniles cannot maintain feeding sites and then do not have nutrients to survive.

Structure variation of soybean microbiomes by different sampling times

Samples from different years revealed different bacterial communities based on beta diversity. The microbial community in bulked soil was larger and more diverse, which is a pool of microbes for plants to recruit (Fig. 3.1 A and 3.2 A). Therefore, changes in bulked soil samples can affect the microbial profile in the rhizosphere and roots. Environmental conditions including pH, nutrient profile, soil type are major drivers of bulked soil bacterial community composition (Lauber et al., 2008; Edwards et al., 2015). In rhizosphere samples, although *Sphingomonas* and *Novosphingobium* genus were constantly abundant in both years, 2018 and 2019, while the *Bradyrhizobium* and *Rhizobium* genus, from *Alphaproteobacteria* class, were more present in 2019. Besides, alpha diversity showed the richness of root samples were significantly different between years (Fig. S3.3). Hence, samples collected at the same fielding

site from different years may exhibit different microbial communities, which have been shown in this study (Fig. 3.1 B) and in other plant microbiome studies (Lauber et al., 2008; Edwards et al., 2015; Rascovan et al., 2016).

In a greenhouse test, different sampling times allowed for the observation of variation in microbial communities (Fig. 3.2 B). The *Pseudoduganella* was found in higher proportion in rhizosphere samples at 10 DAI while *Ralstonia* was highly enriched at 28 DAI. At 42 DAI, when the root was more developed, the N fixation bacteria, *Rhizobium* and *Bradyrhizobium* were increased. The difference in bacterial communities between short sampling times were also reported in different studies including *Arabidopsis* (Chaparro et al., 2014), rice , sugar beet and wheat (Araujo et al., 2019). *Acidobacteria* and *Cyanobacteria* were changed in abundance in *Arabidopsis thaliana* at 17, 24, 31, and 28 days after planting (Chaparro et al., 2014). The study of time series related to rice microbiomes also showed the separation of microbiomes in root compartments by time (Edwards et al., 2015). Relative abundance of core microbiome in the rhizosphere peaked at three days after transplanting, then decreased later (Edwards et al., 2015). The endosphere showed a similar trend to the rhizosphere samples but slower decline (Edwards et al., 2015). The peak time of the endosphere samples was 13 days after transplanting (Edwards et al., 2015). It suggested that microbiome acquisition from soil was rapid, and plants began to assemble microbiomes within a short time after planting.

A study of a perennial plant, *Arabis alpine,* also indicated significantly different root microbiomes were observed at three-time points across 28 weeks in a greenhouse test (Dombrowski et al., 2017). In soybean, the rhizosphere microbial community showed a difference in diversity at different growth stages of soybean plants (Sugiyama et al., 2014). Some

genera including *Bacillus, Strepromyces* had higher relative abundance during flowering stage than mature stage (Sugiyama et al., 2014).

Changes in microbial communities by time could be due to changes in soil physiochemical properties such as moisture and nutrition availability (Zhang et al., 2011; Baskan et al., 2013). Besides, evidence demonstrated that the structure of the bacterial community at different time points could be related to root exudates that might change following a plant development (Baudoin et al., 2002; Chaparro et al., 2013). A chromatographic analysis on the root exudates of *Arabidopsis* at different development stages revealed that this plant secreted different proportions of sugar, amino acid and phenolic (Chaparro et al., 2013). Another possible reason might be associated with the ability of recruiting beneficial microbes to support pathogen resistance, help in mineral nutrition uptake across different stages of development (Pascale et al., 2020).

Beneficial bacteria related to plant growth and disease resistance

The soybean root microbiomes have a significantly greater proportion of *Bradyrhizobiaceae,* a family of *Rhizoiales* order than those in the rhizosphere and bulked soil samples in both field (Fig. 3.3) and greenhouse tests (Fig. 3.5). This family having *Bradyrhizobium* species known as symbiotic nitrogen fixers could be found in the roots of plants especially in legume plants (Sprent et al., 2017). Besides this family, the *Streptomycetaceae* family was found with the greatest proportion of root microbiomes in field tests in both 2018 (16 %) and 2019 (24 %) and a greenhouse test with SCN inoculation (27%). The genus *Streptomyces,* dominated in the *Streptomycetaceae* family, has been found to promote plant growth and be associated with increased drought tolerance (Fitzpatrick et al., 2018). This genus has also been reported as antibiotic producing bacteria that can be used as a biocontrol to

produce a large number of volatile organic compounds that support plant growth (Worsley et al., 2020). In addition, this family has been reported predominantly in endosphere compartments in other plant species such as *Arabidopsis thaliana* (Lundberg et al., 2012) and wheat (Prudence et al., 2021). In both treatments of our greenhouse test, *Burkholedriaceae* family, belonging to *Proteobacteria* phylum with main genera *Burkholderia,* and *Ralstonia* was also found to be prevalent in roots. Other studies also reported that this family was dominant in root microbiomes in wheat and rice (Kawasaki et al., 2021).

For rhizosphere samples in this study, the family *Sphingomonadaceae*, belonging to *Proteobacteria* phylum, was the most prevalent in both field tests in 2018 and 2019 (16 and 17%, respectively). This family was reported in rhizosphere samples in sweet potato (Marques et al., 2014), maize and soybean (Meier et al., 2021). The most prevalent family in *Bacteroidetes* in the rhizosphere of field tests was *Chitinophagaceae* in both years. This family contains several genera including *Balneola, Flavisolbacter, Lacibacter* and *Niastella.* Several strains of the *Chitinophagaceae* family have been isolated as plant growth promoting rhizobacteria (Madhaiyan et al., 2015). This family also has been reported as a dominant family in rhizosphere samples of soybean (Yamazaki et al., 2021). The family *Oxalobacteraceae,* the prevalent family in rhizosphere samples at 10 DAI with 40% in our greenhouse test, was also reported in maize rhizosphere being associated with plant nutrition and growth status (Yu et al., 2021). At 28 DAI, the *Burkholderiaceae* family was the most abundant with 24%. This family was found commonly in the rhizosphere of pumpkin (Kusstatscher et al., 2021), maize (Beirinckx et al., 2020), and wheat (Kawasaki et al., 2021). This family was known for its bioactive metabolite production which has antifungal activity and influences plant growth (Carrion et al., 2018).

Conclusion

This study was focused on the determination of SCN resistance to be associated with microbial communities. Detailed characterization of bulked soil, rhizosphere and root microbiome composition was performed across soybean genotypes in both field and controlled greenhouse conditions. In the field study, no significant association was found between SCN genotype groups (resistant and susceptible) and microbiomes because SCN was not detected in this field. However, the field tests gave a detailed structure and variation of bulked soil, rhizosphere and root samples and factors including years and genotypes that affected microbiota. In our greenhouse study, a list of core bacteria including *Pseudomonas, Rhizobium, Bacillus, Mycobacterium, Dyella and Ramlibacter* genera which might be associated with SCN resistance phenotypes was proposed, which might provide a base knowledge for further studies to explore beneficial bacteria to combat SCN.

References

Ahmad, I., Akhtar, M.J., Asghar, H.N., Ghafoor, U., and Shahid, M. (2016). Differential effects of plant growth-promoting rhizobacteria on maize growth and cadmium Uptake. *J. Plant Growth Regul.* 35(2)**,** 303-315. doi: 10.1007/s00344-015-9534-5.

Al-ani, R.A., Adhab, M.A., Mahdi, M.H., and Abood, H. (2018). *Rhizobium japonicum* as a biocontrol agent of soybean root rot disease caused by *Fusarium solani* and *Macrophomina phaseolina*. *Plant Prot. Sci.* 48**,** 149-155.

Ali, N. (2010). "Soybean processing and utilization," in *The soybean: botany, production and uses*, ed. G. Singh. (Oxfordshire, UK: CAB International), 345-375.

Allen, T.W., Bradley, C.A., Sisson, A.J., Byamukama, E., Chilvers, M.I., Coker, C.M., et al. (2017). Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. *Plant Health Prog.* 18(1)**,** 19-27. doi: 10.1094/php-rs-16-0066.

Anderson, M.J., and Willis, T.J. (2003). Canonical analysis of principal coordinates: A useful method of constrained ordination for ecology. *Ecology* 84(2)**,** 511-525. doi: 10.1890/0012-9658(2003)084[0511:CAOPCA]2.0.CO;2.

Araujo, R., Dunlap, C., Barnett, S., and Franco, C.M.M. (2019). Decoding wheat endosphere–rhizosphere microbiomes in *Rhizoctonia solani*–infested soils challenged by *Streptomyces* biocontrol agents. *Front. Plant Sci.* 10(1038). doi: 10.3389/fpls.2019.01038.

Association, A.S. (2021). A Reference guide to important soybean facts & figures.

Baskan, O., Kosker, Y., and Erpul, G. (2013). Spatial and temporal variation of moisture content in the soil profiles of two different agricultural fields of semi-arid region. *Environ Monit Assess* 185(12)**,** 10441-10458. doi: 10.1007/s10661-013-3343-8.

Baudoin, E., Benizri, E., and Guckert, A. (2002). Impact of growth stage on the bacterial community structure along maize roots, as determined by metabolic and genetic fingerprinting. *Appl. Soil Ecol.* 19(2)**,** 135-145. doi: 10.1016/S0929-1393(01)00185-8.

Beeman, A.Q., and Tylka, G.L. (2017). Assessing the effects of ILeVO and VOTiVO seed treatments on reproduction, hatching, motility, and root penetration of the soybean cyst nematode, *Heterodera glycines*. *Plant Dis.* 102(1)**,** 107-113. doi: 10.1094/PDIS-04-17-0585-RE.

Beilsmith, K., Perisin, M., Bergelson, J., and Martiny, J.B.H. (2021). Natural bacterial assemblages in *Arabidopsis thaliana* tissues become more distinguishable and diverse during host development. *MBio* 12(1)**,** e02723-02720. doi:10.1128/mBio.02723-20.

Beirinckx, S., Viaene, T., Haegeman, A., Debode, J., Amery, F., Vandenabeele, S., et al. (2020). Tapping into the maize root microbiome to identify bacteria that promote growth under chilling conditions. *Microbiome* 8(1)**,** 54. doi: 10.1186/s40168-020-00833-w.

Bergelson, J., Mittelstrass, J., and Horton, M.W. (2019). Characterizing both bacteria and fungi improves understanding of the *Arabidopsis* root microbiome. *Sci. Rep.* 9(1)**,** 24. doi: 10.1038/s41598-018-37208-z.

Bissonnette, K.M., Marett, C.C., Mullaney, M.P., Gebhart, G.D., Kyveryga, P., Mueller, T.A., et al. (2018). Effects of Clariva Complete Beans seed treatment on *Heterodera glycines* reproduction and soybean yield in Iowa. *Plant Health Prog.* 19(1)**,** 1-8. doi: 10.1094/PHP-08-17-0043-RS.

Bray, J.R., and Curtis, J.T. (1957). An ordination of the upland forest communities of Southern Wisconsin. *Ecol. Monogr.* 27(4)**,** 325-349. doi: 10.2307/1942268.

Bulgarelli, D., Garrido-Oter, R., Münch, P.C., Weiman, A., Dröge, J., Pan, Y., et al. (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17(3)**,** 392-403. doi: 10.1016/j.chom.2015.01.011.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* 108(Supplement 1)**,** 4516. doi: 10.1073/pnas.1000080107.

Carrion, V., Cordovez, V., Tyc, O., Etalo, D., Bruijn, I., Jager, V., et al. (2018). Involvement of *Burkholderiaceae* and sulfurous volatiles in disease-suppressive soils. *The ISME Journal* 12**,** 1. doi: 10.1038/s41396-018-0186-x.

Chao, A. (1984). Nonparametric Estimation of the Number of Classes in a Population. *Scand Stat Theory Appl* . 11(4)**,** 265-270.

Chaparro, J.M., Badri, D.V., Bakker, M.G., Sugiyama, A., Manter, D.K., and Vivanco, J.M. (2013). Root exudation of phytochemicals in *Arabidopsis* follows specific patterns that are developmentally programmed and correlate with soil microbial functions. *PLoS One* 8(2)**,** e55731. doi: 10.1371/journal.pone.0055731.

Chaparro, J.M., Badri, D.V., and Vivanco, J.M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *The ISME Journal* 8(4)**,** 790-803. doi: 10.1038/ismej.2013.196.

Chen, S., Waghmode, T.R., Sun, R., Kuramae, E.E., Hu, C., and Liu, B. (2019). Root-associated microbiomes of wheat under the combined effect of plant development and nitrogen fertilization. *Microbiome* 7(1)**,** 136. doi: 10.1186/s40168-019-0750-2.

Choudhary, D.K., and Johri, B.N. (2009). Interactions of *Bacillus spp.* and plants – with special reference to induced systemic resistance (ISR). *Microbiol. Res.* 164(5)**,** 493-513. doi: 10.1016/j.micres.2008.08.007.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42(Database issue)**,** D633-642. doi: 10.1093/nar/gkt1244.

Compant, S., Samad, A., Faist, H., and Sessitsch, A. (2019). A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *J. Adv. Res.* 19**,** 29-37. doi: 10.1016/j.jare.2019.03.004.

Cook, D.E., Lee, T.G., Guo, X., Melito, S., Wang, K., Bayless, A.M., et al. (2012). Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338(6111)**,** 1206. doi: 10.1126/science.1228746.

Desaeger, J., Wram, C., and Zasada, I. (2020). New reduced-risk agricultural nematicides - rationale and review. *J. Nematol.* 52**,** e2020-2091. doi: 10.21307/jofnem-2020-091.

Dombrowski, N., Schlaeppi, K., Agler, M.T., Hacquard, S., Kemen, E., Garrido-Oter, R., et al. (2017). Root microbiota dynamics of perennial *Arabis alpina* are dependent on soil residence time but independent of flowering time. *The ISME Journal* 11(1)**,** 43-55. doi: 10.1038/ismej.2016.109.

Duca, D., Lorv, J., Patten, C.L., Rose, D., and Glick, B.R. (2014). Indole-3-acetic acid in plant-microbe interactions. *Antonie Van Leeuwenhoek* 106(1)**,** 85-125. doi: 10.1007/s10482-013-0095-y.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16)**,** 2194-2200. doi: 10.1093/bioinformatics/btr381.

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci.* 112(8)**,** E911. doi: 10.1073/pnas.1414592112.

Edwards, J., Santos-Medellín, C., and Sundaresan, V. (2018). Extraction and 16S rRNA sequence analysis of microbiomes associated with rice roots. *Bio-protocol* 8(12)**,** e2884-e2884. doi: 10.21769/BioProtoc.2884.

Fatima, S., and Anjum, T. (2017). Identification of a potential ISR determinant from *Pseudomonas aeruginosa* PM12 against *Fusarium* Wilt in Tomato. *Front. Plant Sci.* 8(848). doi: 10.3389/fpls.2017.00848.

Fitzpatrick, C.R., Copeland, J., Wang, P.W., Guttman, D.S., Kotanen, P.M., and Johnson, M.T.J. (2018). Assembly and ecological function of the root microbiome across angiosperm plant species. *Proc. Natl. Acad. Sci.* 115(6)**,** E1157. doi: 10.1073/pnas.1717617115.

Floc'h, J.-B., Hamel, C., Lupwayi, N., Harker, K.N., Hijri, M., and St-Arnaud, M. (2020). Bacterial communities of the canola rhizosphere: Network analysis reveals a core bacterium shaping microbial interactions. *Front. Microbiol.* 11(1587). doi: 10.3389/fmicb.2020.01587.

Fujimatsu, T., Endo, K., Yazaki, K., and Sugiyama, A. (2020). Secretion dynamics of soyasaponins in soybean roots and effects to modify the bacterial composition. *Plant direct* 4(9)**,** e00259-e00259. doi: 10.1002/pld3.259.

Guo, Y., Lin, Q., Chen, L., Carballar-Lejarazú, R., Zhang, A., Shao, E., et al. (2020). Characterization of bacterial communities associated with the pinewood nematode insect vector *Monochamus alternatus Hope* and the host tree *Pinus massoniana. BMC Genom.* 21(1)**,** 337. doi: 10.1186/s12864-020-6718-6.

Hussain, M., Hamid, M.I., Tian, J., Hu, J., Zhang, X., Chen, J., et al. (2018). Bacterial community assemblages in the rhizosphere soil, root endosphere and cyst of soybean cyst nematode-suppressive soil challenged with nematodes. *FEMS Microbiol. Ecol.* 94(10). doi: 10.1093/femsec/fiy142.

Kawasaki, A., Dennis, P.G., Forstner, C., Raghavendra, A.K.H., Richardson, A.E., Watt, M., et al. (2021). The microbiomes on the roots of wheat (*Triticum aestivum L.*) and rice (*Oryza sativa L.*) exhibit significant differences in structure between root types and along root axes. *Funct. Plant Biol.* 48(9), 871-888. doi: 10.1071/FP20351.

Kluepfel, D.A., Nyczepir, A.P., Lawrence, J.E., Wechter, W.P., and Leverentz, B. (2002). Biological control of the phytoparasitic nematode *Mesocriconema xenoplax* on peach trees. *J. Nematol.* 34(2)**,** 120-123.

Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29(1)**,** 1-27. doi: 10.1007/BF02289565.

Kusstatscher, P., Adam, E., Wicaksono, W.A., Bernhart, M., Olimi, E., Müller, H., et al. (2021). Microbiome-assisted breeding to understand cultivar-dependent assembly in *Cucurbita pepo*. *Front. Plant Sci.* 12(505). doi: 10.3389/fpls.2021.642027.

Lauber, C.L., Strickland, M.S., Bradford, M.A., and Fierer, N. (2008). The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol. Biochem.* 40(9)**,** 2407-2415. doi: 10.1016/j.soilbio.2008.05.021.

Lay, C.-Y., Bell, T.H., Hamel, C., Harker, K.N., Mohr, R., Greer, C.W., et al. (2018). Canola root-associated microbiomes in the Canadian Prairies. *Front. Microbiol.* 9**,** 1188-1188. doi: 10.3389/fmicb.2018.01188.

Liu, F., Hewezi, T., Lebeis, S.L., Pantalone, V., Grewal, P.S., and Staton, M.E. (2019). Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. *BMC Microbiol.* 19(1)**,** 201. doi: 10.1186/s12866-019-1572-x.

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., et al. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488(7409)**,** 86-90. doi: 10.1038/nature11237.

Madhaiyan, M., Poonguzhali, S., Senthilkumar, M., Pragatheswari, D., Lee, J.S., and Lee, K.C. (2015). *Arachidicoccus rhizosphaerae gen. nov., sp. nov.*, a plant-growth-promoting bacterium in the family *Chitinophagaceae* isolated from rhizosphere soil. *Int J Syst Evol Microbiol* 65(Pt 2)**,** 578-586. doi: 10.1099/ijs.0.069377-0.

Marques, J.M., da Silva, T.F., Vollu, R.E., Blank, A.F., Ding, G.C., Seldin, L., et al. (2014). Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown sweet potato plants. *FEMS Microbiol Ecol* 88(2)**,** 424-435. doi: 10.1111/1574-6941.12313.

McMurdie, P.J., and Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE* 8(4)**,** e61217. doi: 10.1371/journal.pone.0061217.

Meier, M.A., Lopez-Guerrero, M.G., Guo, M., Schmer, M.R., Herr, J.R., Schnable, J.C., et al. (2021). Rhizosphere microbiomes in a historical maize-soybean rotation system respond to host species and nitrogen fertilization at the genus and subgenus levels. *Appl. Environ. Microbiol.* 87(12)**,** e03132-03120. doi:10.1128/AEM.03132-20.

Mitchum, M.G. (2016). Soybean resistance to the soybean cyst nematode *Heterodera glycines*: An update. *Phytopathology* 106(12)**,** 1444-1450. doi: 10.1094/phyto-06-16-0227-rvw.

Munkvold, G.P. (2009). Seed pathology progress in academia and industry. *Annu Rev Phytopathol* 47**,** 285-311. doi: 10.1146/annurev-phyto-080508-081916.

Noel, G.R., and Stanger, B.A. (1994). First Report of *Pasteuria sp.* attacking *Heterodera glycines* in North America. *J Nematol* 26(4 Suppl)**,** 612-615.

Nour, S.M., Lawrence, J.R., Zhu, H., Swerhone, G.D., Welsh, M., Welacky, T.W., et al. (2003). Bacteria associated with cysts of the soybean cyst nematode (*Heterodera glycines*). *Appl Environ Microbiol* 69(1)**,** 607-615. doi: 10.1128/aem.69.1.607-615.2003.

Pascale, A., Proietti, S., Pantelides, I.S., and Stringlis, I.A. (2020). Modulation of the root microbiome by plant molecules: The basis for targeted disease suppression and plant growth promotion. *Front. Plant Sci.* 10**,** 1741-1741. doi: 10.3389/fpls.2019.01741.

Peiffer, J.A., Spor, A., Koren, O., Jin, Z., Tringe, S.G., Dangl, J.L., et al. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci.* 110(16)**,** 6548. doi: 10.1073/pnas.1302837110.

Prudence, S., Newitt, J., Worsley, S.F., Macey, M.C., Murrell, J.C., Lehtovirta-Morley, L.E., et al. (2021). Soil, senescence and exudate utilisation: Characterisation of the Paragon var. spring bread wheat root microbiome. *Environ. Microbiol.* 16**,** 12 (2021). doi: 10.1186/s40793-021-00381-2

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41(D1)**,** D590-D596. doi: 10.1093/nar/gks1219.

Rascovan, N., Carbonetto, B., Perrig, D., Díaz, M., Canciani, W., Abalo, M., et al. (2016). Integrated analysis of root microbiomes of soybean and wheat from agricultural fields. *Sci. Rep.* 6(1)**,** 28084. doi: 10.1038/srep28084.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75(23)**,** 7537-7541. doi: 10.1128/AEM.01541-09.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol* 12(6)**,** R60. doi: 10.1186/gb-2011-12-6-r60.

Si, J., Froussart, E., Viaene, T., Vázquez-Castellanos, J.F., Hamonts, K., Tang, L., et al. (2021). Interactions between soil compositions and the wheat root microbiome under drought stress: From an in silico to in planta perspective. *Comput. Struct. Biotechnol. J.* 19**,** 4235-4247. doi: 10.1016/j.csbj.2021.07.027.

Simpson, E.H. (1949). Measurement of Diversity. *Nature* 163(4148)**,** 688-688. doi: 10.1038/163688a0.

Siqueira, A.F., Ormeño-Orrillo, E., Souza, R.C., Rodrigues, E.P., Almeida, L.G.P., Barcellos, F.G., et al. (2014). Comparative genomics of *Bradyrhizobium japonicum CPAC 15* and *Bradyrhizobium diazoefficiens CPAC 7*: elite model strains for understanding symbiotic performance with soybean. *BMC Genom.* 15(1)**,** 420. doi: 10.1186/1471-2164-15-420.

Spellerberg, I.F., and Fedor, P.J. (2003). A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener' Index. *Global Ecology and Biogeography* 12(3)**,** 177-179. doi: 10.1046/j.1466-822X.2003.00015.x.

Sprent, J.I., Ardley, J., and James, E.K. (2017). Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytologist* 215(1)**,** 40-56. doi: 10.1111/nph.14474.

Sugiyama, A., Ueda, Y., Zushi, T., Takase, H., and Yazaki, K. (2014). Changes in the bacterial community of soybean rhizospheres during growth in the field. *PLOS ONE* 9(6)**,** e100709. doi: 10.1371/journal.pone.0100709.

Thapa, S., Cole, E., Howland, A.D., Levene, B., and Quintanilla, M. (2022). Soybean cyst nematode (*Heterodera glycines*) resistant cultivar rotation system impacts nematode population density, virulence, and yield. *Crop Protection* 153**,** 105864. doi: 10.1016/j.cropro.2021.105864.

Tian, B., Yang, J., and Zhang, K.-Q. (2007). Bacteria used in the biological control of plant-parasitic nematodes: populations, mechanisms of action, and future prospects. *FEMS Microbiol. Ecol.* 61(2)**,** 197-213. doi: 10.1111/j.1574-6941.2007.00349.x.

Tian, H., and Riggs, R.D. (2000). Effects of rhizobacteria on soybean cyst nematode, *Heterodera glycines. J Nematol* 32(4)**,** 377-388.

Tinker, K.A., Ottesen, E.A., and Schloss, P.D. (2016). The core gut microbiome of the American cockroach, *Periplaneta americana*, is stable and resilient to dietary shifts. *Appl. Environ. Microbiol.* 82(22)**,** 6603-6610. doi: doi:10.1128/AEM.01837-16.

Tylka, G.L., and Marett, C.C. (2021). Known distribution of the soybean cyst nematode, *Heterodera glycines*, in the United States and Canada in 2020. *Plant Health Prog.* 22(1)**,** 72-74. doi: 10.1094/php-10-20-0094-br.

Wang, M., Eyre, A.W., Thon, M.R., Oh, Y., and Dean, R.A. (2020). Dynamic changes in the Microbiome of rice during shoot and root growth derived from seeds. *Front. Microbiol.* 11(2183). doi: 10.3389/fmicb.2020.559728.

Wang, P., Marsh, E.L., Ainsworth, E.A., Leakey, A.D.B., Sheflin, A.M., and Schachtman, D.P. (2017). Shifts in microbial communities in soil, rhizosphere and roots of two major crop systems under elevated CO2 and O3. *Sci. Rep.* 7(1)**,** 15019. doi: 10.1038/s41598-017-14936-2.

Wasimuddin, Schlaeppi, K., Ronchi, F., Leib, S.L., Erb, M., and Ramette, A. (2020). Evaluation of primer pairs for microbiome profiling from soils to humans within the One Health framework. *Mol Ecol Resour* 20(6)**,** 1558-1571. doi: 10.1111/1755-0998.13215.

Weinert, N., Piceno, Y., Ding, G.-C., Meincke, R., Heuer, H., Berg, G., et al. (2011). PhyloChip hybridization uncovered an enormous bacterial diversity in the rhizosphere of different potato cultivars: many common and few cultivar-dependent taxa. *FEMS Microbiol. Ecol.* 75(3)**,** 497-506. doi: 10.1111/j.1574-6941.2010.01025.x.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. Springer International Publishing.

Wickham, H. (2021). tidyr: tidy messy data. *R package version 1.1.3*.

Wickham, H., François, R., Henry, L., and Müller, K. (2021). dplyr: A grammar of data manipulation. *R package version 1.0.6*.

Worsley, S.F., Newitt, J., Rassbach, J., Batey, S.F.D., Holmes, N.A., Murrell, J.C., et al. (2020). *Streptomyces* endophytes promote host health and enhance growth across plant species. *Appl. Environ. Microbiol.* 86(16)**,** e01053-01020. doi: doi:10.1128/AEM.01053-20.

Yamazaki, S., Mardani-korrani, H., Kaida, R., Ochiai, K., Kobayashi, M., Nagano, A.J., et al. (2021). Field multi-omics analysis reveals a close association between bacterial communities and mineral properties in the soybean rhizosphere. *Sci. Rep.* 11(1)**,** 8878. doi: 10.1038/s41598-021-87384-8.

Yandigeri, M.S., Meena, K.K., Singh, D., Malviya, N., Singh, D.P., Solanki, M.K., et al. (2012). Drought-tolerant endophytic actinobacteria promote growth of wheat (Triticum aestivum) under water stress conditions. *Plant Growth Regul.* 68(3)**,** 411-420. doi: 10.1007/s10725-012-9730-2.

Yang, L., Schröder, P., Vestergaard, G., Schloter, M., and Radl, V. (2020a). Response of barley plants to drought might be associated with the recruiting of soil-borne endophytes. *Microorganisms* 8(9)**,** 1414.

Yang, Y., Liu, L., Singh, R.P., Meng, C., Ma, S., Jing, C., et al. (2020b). Nodule and root zone microbiota of salt-tolerant wild soybean in coastal sand and saline-alkali soil. *Front. Microbiol.* 11(2178). doi: 10.3389/fmicb.2020.523142.

Yin, B.-B., Valinsky, L., Gao, X., Becker, J., and Borneman, J. (2003). Bacterial rRNA genes associated with soil suppressiveness against the plant-parasitic nematode *Heterodera schachtii*. *Appl. Environ. Microbiol.* 69**,** 1573-1580. doi: 10.1128/AEM.69.3.1573-1580.2003.

Yu, P., He, X., Baer, M., Beirinckx, S., Tian, T., Moya, Y.A.T., et al. (2021). Plant flavones enrich rhizosphere *Oxalobacteraceae* to improve maize performance under nitrogen deprivation. *Nature Plants* 7(4)**,** 481-499. doi: 10.1038/s41477-021-00897-y.

Zhang, C.-h., Wang, Z.-m., Ju, W.-m., and Ren, C. (2011). Spatial and temporal variability of soil C/N ratio in Songnen Plain maize belt. *Huan jing ke xue* 32**,** 1407-1414.

Tables and Figures

Table 3.1 Soybean genotypes included in the field experiments conducted in Plains, GA in 2018 and 2019

Genotype	Maturity group (MG)	Type of Resistance	No of samples [a]
PI 548402	IV	'Peking'-type resistance	9
PI 90763	IV	'Peking'-type resistance	9
PI 89772	IV	'Peking'-type resistance	9
PI 437654	III	'Peking'-type resistance	9
G93-9009	VI	'Peking'-type resistance	9
PI 88788	III	PI 88788-type resistance	9
PI 209332	IV	PI 88788-type resistance	9
PI 548316	III	PI 88788-type resistance	9
G93-9106	VII	PI 88788-type resistance	9
'Bossier'	VIII	Susceptible	9
'CNS'	VII	Susceptible	9
'Hutcheson'	V	Susceptible	9
'Lee'	VI	Susceptible	9
'Lee 74'	VI	Susceptible	9

[a] included bulked soil, rhizosphere, and root samples from three replications.

Table 3.2 HG type 0 female index of soybean genotypes included in a greenhouse experiment in 2020 at 28 days after inoculation.

Soybean genotypes	Maturity group (MG)	Type of SCN Resistance	No of samples [a]	Number of cysts [b]	FI (%)[c]	Phenotype
'Woodruff'	VIII	'Peking'-type resistant	27	0.7	0.5	Resistant
'G00-3880'	VII	PI 88788-type resistant	27	1.9	1.3	Resistant
'Lee 74'	VI	Susceptible	27	142.8	100	Susceptible

[a] included bulked soil, rhizosphere, and root samples from three replicates at each of three time points: 10, 28, and 42 days after inoculation.
[b] Number of cysts: average number of cysts on the roots of each soybean genotype.
[c] Female index (FI) = (average number of cysts) x 100/ (average number of cysts on 'Lee 74').

Table 3.3 Four enriched genera with a LDA score of 3.0 among rhizosphere samples of three genotypes at 10 DAI under HG type 0 treatment based on LefSe analysis. No significantly enriched genera was found in 'Lee 74'.

Phylum	Class	Family	Genus	p. value	LDA score	Direction
Proteobacteria	*Alphaproteobacteria*	*Bradyrhizobiaceae*	*Bradyrhizobium*	0.04	4.3	'G00-3880'
Armatimonadetes	*Armatimonadia*	*Armatimonadaceae*	*Armatimonas/ Armatimonadetes*	0.04	3.96	'Woodruff'
Chlamydiae	*Chlamydiia*	*Parachlamydiaceae*	*Neochlamydia*	0.03	3.37	'G00-3880'
Proteobacteria	*Alphaproteobacteria*	*Rhizobiaceae*	*Shinella*	0.02	3.05	'G00-3880'

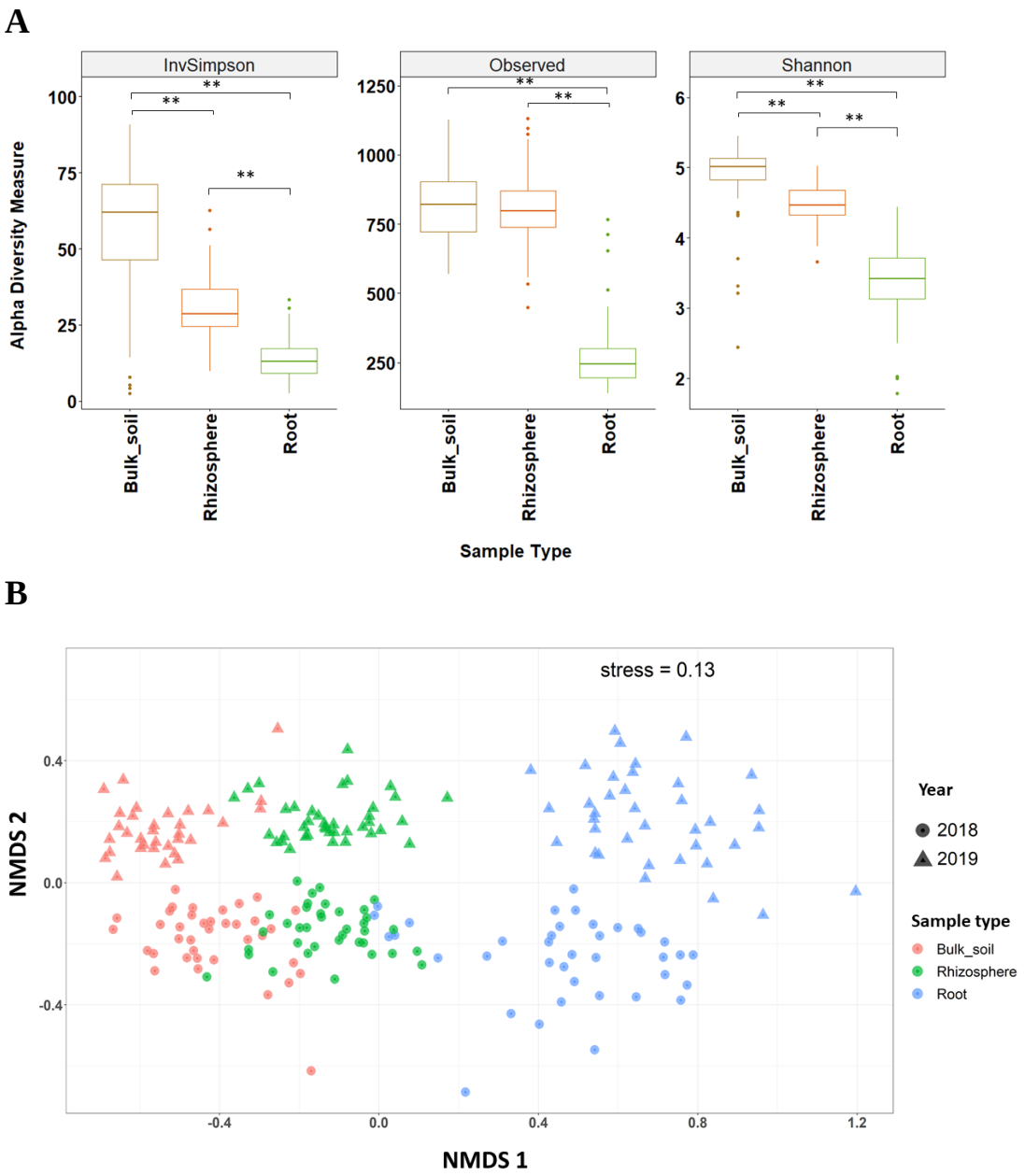
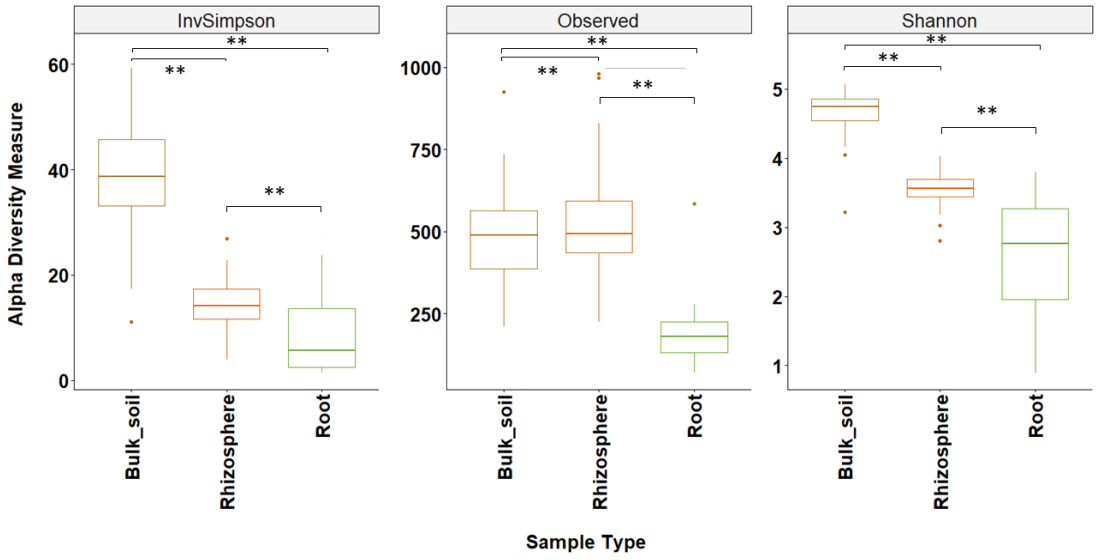


Fig 3.1 Alpha diversity and NMDS plot of field tests in 2018 and 2019
A) Alpha diversity analysis across all three types of samples from the field tests in 2018 and 2019. Alpha diversity measures including Inverse Simpson (InvSimpson), Observed and Shannon index revealed significant differences in richness among three types of samples. Pairwise comparisons were tested using a Tukey's method. ** denotes $P < 0.01$.
B) NMDS plot based on the Bray Curtis distance comparing the β diversity among three types of samples and between two years. Microbial communities in root showed separation from bulk soil and rhizosphere samples (PERMANOVA: $P = 4.07e\text{-}10$).

A



B

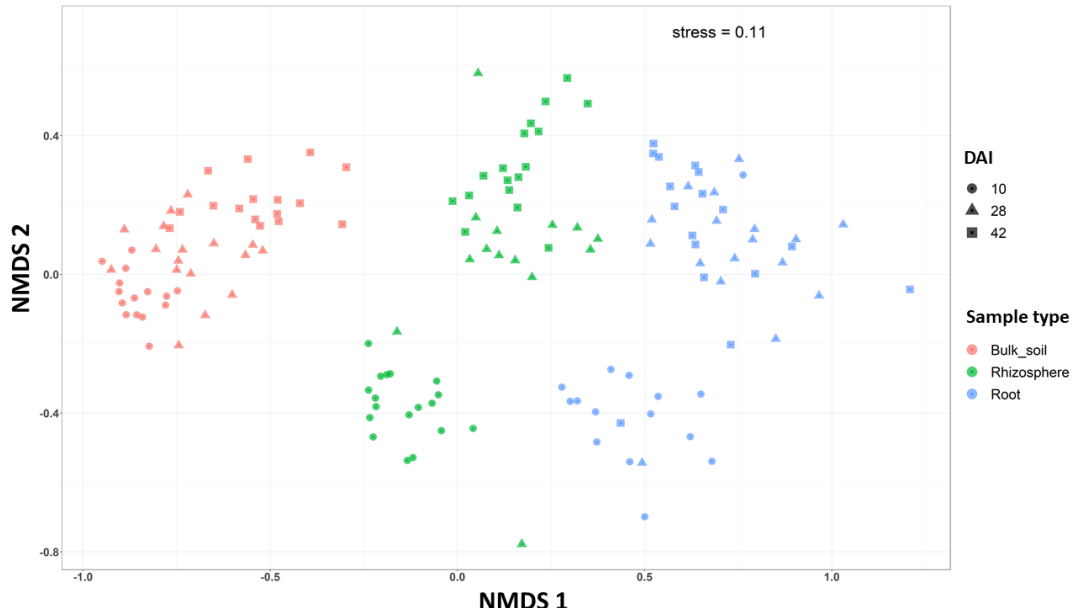


Fig 3.2 Alpha diversity of bacterial communities and NMDS plot of a greenhouse test
A) Alpha diversity of bacterial communities across three types of samples in a greenhouse test using InvSimpson, Observed OTUs, and Shannon diversity. Turkey tests revealed significant differences in alpha diversity with $P < 0.01$ (**).
B) NMDS plot of Bray Curtis distance among three types of samples at three time points. Rhizosphere and root samples that were collected at 10 DAI (days after inoculation) showed a significant separation from those collected at 28 and 42 DAIs.

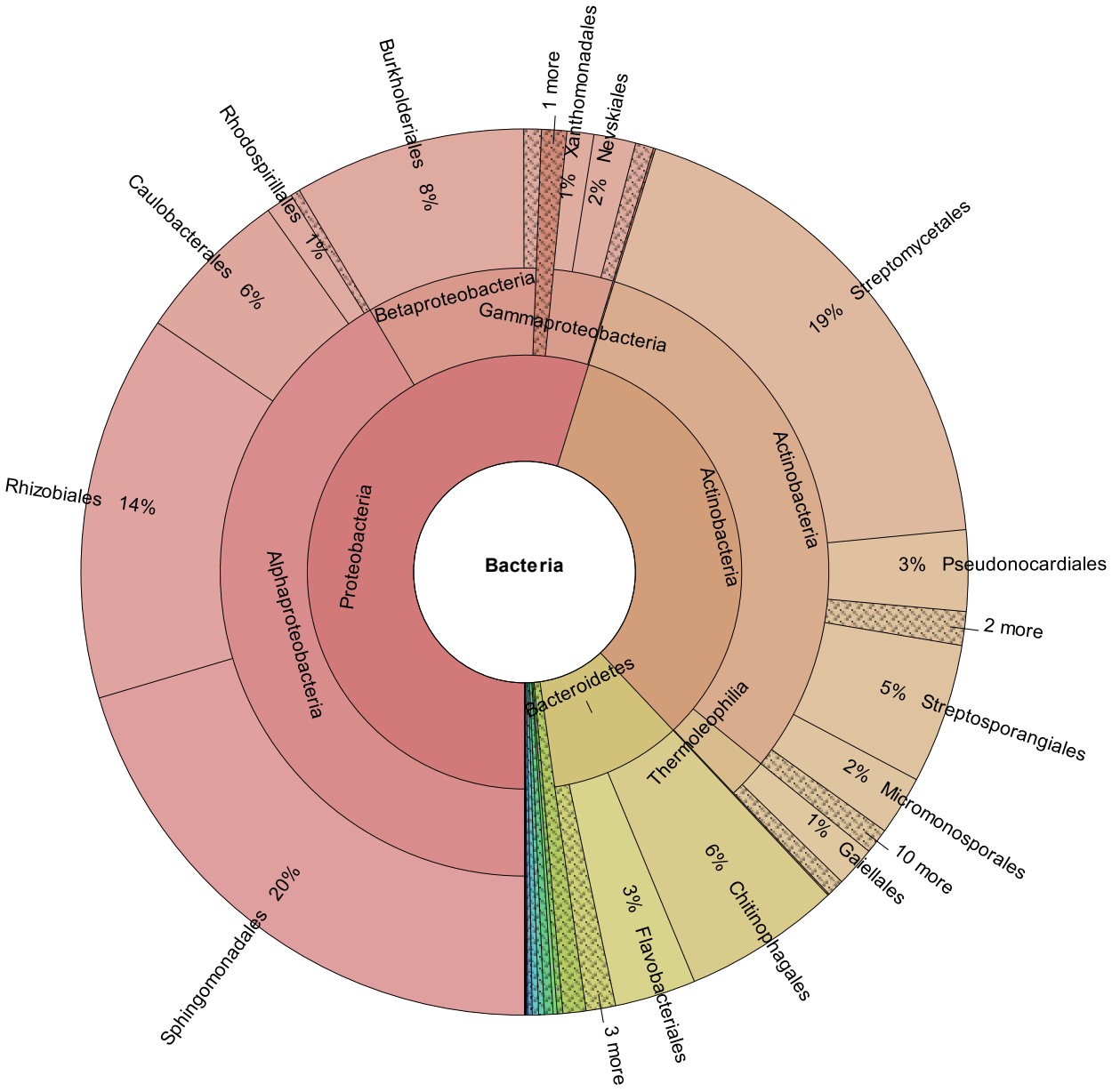


Fig 3.3 Bacterial diversity of root samples from the field test in 2018.
The distribution of 16S reads among taxa is shown as a Krona plot with the kingdom Bacteria at the center and lower taxonomic divisions toward the edge: Phylum, Class, Order.

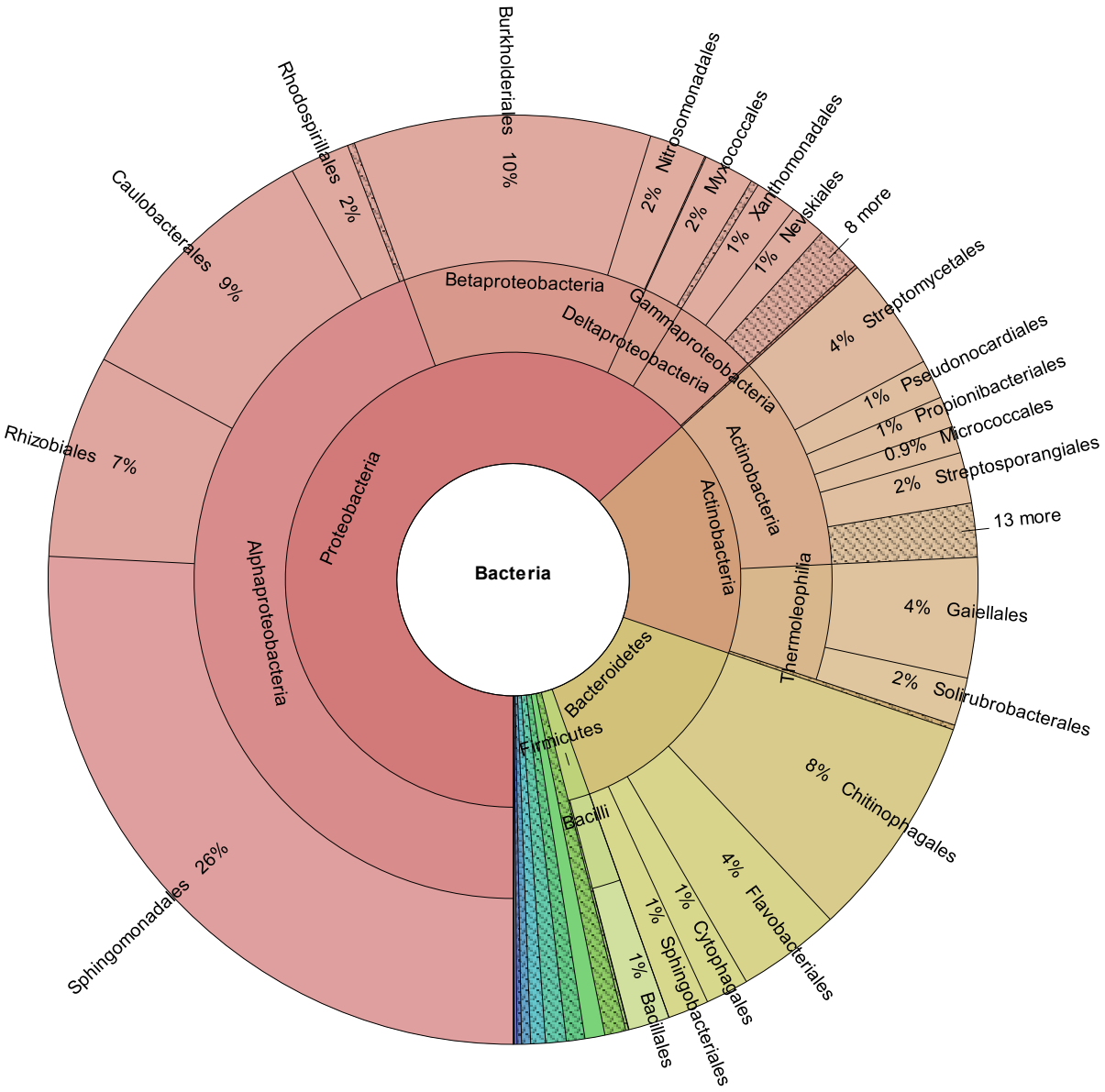


Fig 3.4 Krona graph showing the relative abundance of taxa in the rhizosphere samples from the field test in 2018.
The order of taxonomic levels are Kingdom, Phylum, Class and Order from the center to the edge of the plot.

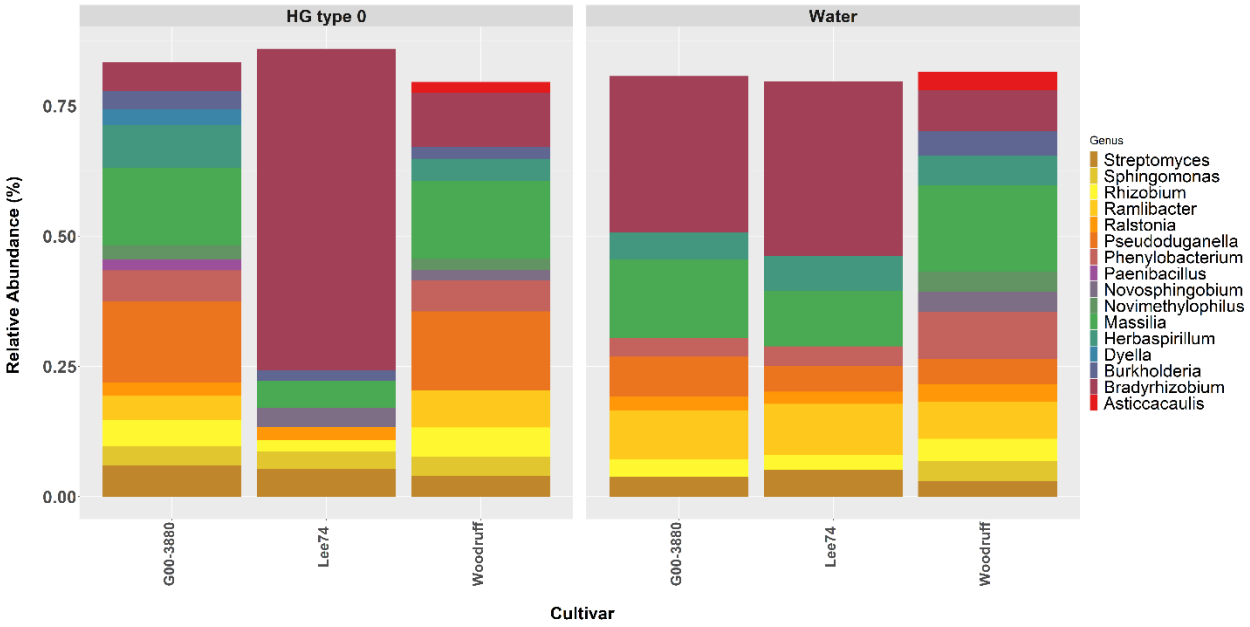


Fig 3.5 Relative abundance (> 2%) of genus level OTUs in root samples of three genotypes at 10 days after HG type 0 and water inoculations. Two HG type 0 resistant genotypes: 'G00-3880' and 'Woodruff' had higher proportions of some genera including *Pseudoduganella*, *Ramlibacter* and *Massilia* than the susceptible genotype, 'Lee 74'

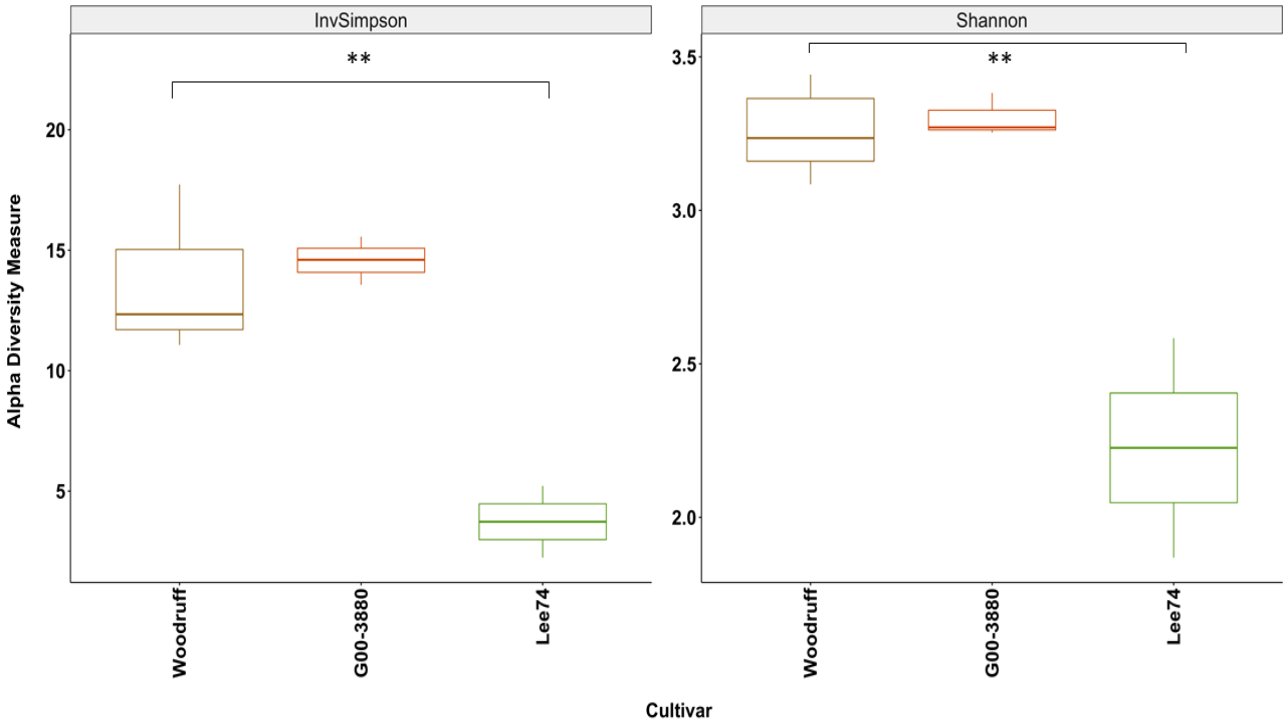


Fig 3.6 Alpha diversity using InvSimpson and Shannon index measurements 'Woodruff' and 'G00-3880' (resistant) had higher density levels than 'Lee 74' (susceptible) in root samples at 10 DAI under a HG type 0 treatment. ** denotes *P = 0.01* in the Tukey's pairwise test.

Fig 3.7 LefSe result graph of seven enriched taxa of root samples
They were identified with a LDA score of 3.5 among three genotypes at 10 DAI under HG type 0 treatment. Only the taxa meeting a significant LDA threshold was shown.

Supplementary tables and figures

Table S3.1 Primers and sequences for two-step PCR

PCR	Forward primer (5'-3')	Reverse primer (5'-3')
1st PCR [a]	GTGCCAGCMGCCGCGGTAA [b]	GGACTACHVGGGTWTCTAAT [b]
2nd PCR [c]	AATGATACGGCGACCACCGA GATCTACAC [d] TATGGTAATT [e] CA[f]	CAAGCAGAAGACGGCATACGA GAT [d] AGCAGCTCAGG[e] AC [f]

[a] Primer sequences were based on Caporaso et al. (2011) and Tinker et al. (2016)
[b] Universal primer, 515F and 806R to amplify the V4 region
[c] Primers for 2nd PCR: Forward primer = Illumina adapter + barcode + pad + linker + 515F. Reverse primer = Illumina adaptor + barcode + pad + linker + 806R. Each sample has a unique combination of forward and reverse barcodes.
[d] Illuminia adaptors
[e] 10-nt pad: prevent hairpin formation or boost the sequencing primer melting temperature.
[f] 2-nt linker: anti- complementary to the known sequences.

Table S3.2 Barcodes for 16S sequencing

No	Forward barcode [a]	Reverse barcode
1	AACCAACC	GTGAGACA
2	CCAACCAA	TACGTTGG
3	GGTTGGTT	ATGCTACG
4	TTGGTTGG	TTCGTTCG
5	AGTCGACT	ACACAGTC
6	CCATCCTA	GAGTCAGA
7	GTCAAGAG	CGATGGTT
8	TAGGTTGC	ATCGTTGG
9	AAGCAAGC	TAGCAACC
10	CGTTCGTT	ACGTCATG
11	GCAAGCAA	GATGTCGT
12	TTCGTTCG	GTGACTCA
13	AGGTGAAC	TCACAGAC
14	CTACAGCA	AGTGTCTG
15	GACACTGT	CACTGAGT
16	TCTGTCTC	GTGTGTGT
17	TCTGTGTC	AACGAACG

[a] Barcode sequences were based on Caporaso et al. (2011)

Table S3.3 MOTHUR commands used for sequence analysis

MOTHUR command [a]	**Function**
make.contig(file.txt, processors=60)	Join forward and reverse reads
screen.seqs(fasta=current, group=current, maxambig=0, maxlength=275)	Screen reads which do not assemble well by removing any sequences with ambiguous bases and longer than 275 bp
unique.seqs(fasta=current)	Select only unique sequences
count.seqs(name=current, group=current)	Generate a table where rows are name of unique sequences
pcr.seqs(fasta=silva.nr_v132.align, oligos=v4oligos.txt)	Customize database to V4 region because we use primers to amplify V4 region
align.seqs(fasta=file.unique.fasta, reference=silva.nr_v132.pcr.align, flip=true)	Align sequences to a customized reference alignment
screen.seqs(fasta=current, count=current, summary=current, start=13862, end=23444, maxhomop=8, minlength=200)	Screen reads to remove any poor alignments to make sure everything overlaps the same region.
filter.seqs(fasta=current, vertical=T, trump=.)	Filter to remove the overhands at both end
unique.seqs(fasta=current, count=current)	Check to make sure only unique sequences
pre.cluster(fasta=current, count=current, diffs=2)	De noise sequences to divide the sequences by group and sort them by abundance
chimera.uchime(fasta=current, count=current, dereplicate=t)	Check for chimeras
remove.seqs(fasta=current, accnos= current)	Remove chimeric sequences
classify.seqs(fasta=current, count=current, reference=trainset18_062020.rdp.fasta, taxonomy=trainset18_062020.rdp.tax,cutoff=80)	Assign sequences to the RDP taxonomy reference version 18
remove.lineage(fasta=current, count=current, taxonomy=current,taxon=Chloroplast-Mitochondria-unknown-Archaea-Eukaryota)	Remove unclassified sequences, and chloroplast, mitochondria
dist.seqs(fasta=current, cutoff=0.03)	Calculate pairwise distance between aligned DNA sequences
cluster(column=file.dist, count=current)	Cluster sequences into OTUs
make.shared(list=current, count=current, label=0.03)[b]	Read a list and group files to create shared file
classify.otu(list=current, count=current, taxonomy=current,label=0.03)[b]	Get consensus taxonomy for each OTU

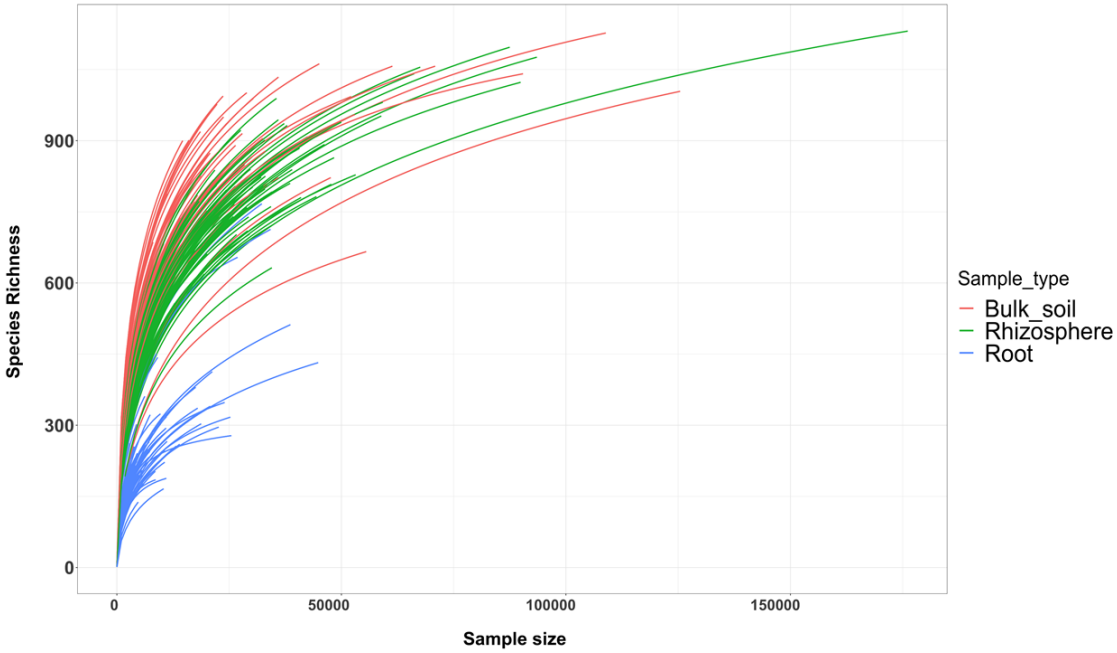
[a] The workflow was following the MiseqSOP

[b] Output files from make.shared and classify.otu were used as an input file in the phyloseq package running in R

Table S3.4 List of enriched genera in three sample types with a LDA score of 3.5 between two treatments at 10 DAI

Phylum	Class	Family	Genus	P-value	LDA score	Direction
			Bulked soil			
Proteobacteria	*Betaproteobacteria*	*Oxalobacteraceae*	*Pseudoduganella*	0.03	5.24	HG0
Proteobacteria	*Alphaproteobacteria*	*Hyphomicrobiaceae*	*Pedomicrobium*	0.02	4.35	HG0
Proteobacteria	*Gammaproteobacteria*	*Steroidobacteraceae*	*Povalibacter*	0.02	4.22	HG0
Proteobacteria	*Deltaproteobacteria*	*Cystobacteraceae*	*Archangium*	0.05	4.13	HG0
Bacteroidetes	*Chitinophagia*	*Chitinophagaceae*	*Terrimonas*	0.00	4.08	HG0
Proteobacteria	*Gammaproteobacteria*	*Moraxellaceae*	*Cavicella*	0.01	4.05	HG0
Proteobacteria	*Betaproteobacteria*	*Methylophilaceae*	*Methylophilus*	0.04	3.92	HG0
Actinobacteria	*Actinobacteria*	*Micrococcaceae*	*Pseudarthrobacter*	0.00	4.64	Water
Bacteroidetes	*Chitinophagia*	*Chitinophagaceae*	*Segetibacter*	0.00	4.17	Water
Proteobacteria	*Alphaproteobacteria*	*Phyllobacteriaceae*	*Mesorhizobium*	0.01	4.13	Water
Proteobacteria	*Gammaproteobacteria*	*Legionellaceae*	*Legionella*	0.00	4.11	Water
Bacteroidetes	*Sphingobacteriia*	*Sphingobacteriaceae*	*Mucilaginibacter*	0.05	4.1	Water
Proteobacteria	*Gammaproteobacteria*	*Coxiellaceae*	*Aquicella*	0.03	4.07	Water
Proteobacteria	*Deltaproteobacteria*	*Polyangiaceae*	*Sorangium*	0.05	4.06	Water
Actinobacteria	*Actinobacteria*	*Intrasporangiaceae*	*Terrabacter*	0.02	3.79	Water
Bacteroidetes	*Flavobacteriia*	*Flavobacteriaceae*	*Flavobacterium*	0.03	3.53	Water
			Rhizosphere			
Proteobacteria	*Betaproteobacteria*	*Oxalobacteraceae*	*Pseudoduganella*	0.00	5.71	HG0
Bacteroidetes	*Cytophagia*	*Fulvivirgaceae*	*Ohtaekwangia*	0.04	3.71	HG0
Bacteroidetes	*Sphingobacteriia*	*Sphingobacteriaceae*	*Mucilaginibacter*	0.02	4.77	Water
Verrucomicrobia	*Opitutae*	*Opitutaceae*	*Opitutus*	0.01	4.71	Water
Proteobacteria	*Alphaproteobacteria*	*Phyllobacteriaceae*	*Mesorhizobium*	0.01	4.7	Water
Proteobacteria	*Alphaproteobacteria*	*Bradyrhizobiaceae*	*Bradyrhizobium*	0.00	4.51	Water
Proteobacteria	*Alphaproteobacteria*	*Caulobacteraceae*	*Asticcacaulis*	0.01	4.49	Water
Bacteroidetes	*Chitinophagia*	*Chitinophagaceae*	*Flavisolibacter*	0.01	3.97	Water
Proteobacteria	*Gammaproteobacteria*	*Nevskiaceae*	*Stenotrophobium*	0.00	3.69	Water
Actinobacteria	*Actinobacteria*	*Micromonosporaceae*	*Actinoplanes*	0.04	3.68	Water
Proteobacteria	*Alphaproteobacteria*	*Xanthobacteraceae*	*Labrys*	0.04	3.58	Water
Bacteroidetes	*Chitinophagia*	*Chitinophagaceae*	*Segetibacter*	0.00	3.55	Water
			Root			
Proteobacteria	*Alphaproteobacteria*	*Sphingomonadaceae*	*Sphingomonas*	0.04	4.8	HG0
Proteobacteria	*Betaproteobacteria*	*Burkholderiaceae*	*Burkholderia*	0.04	4.36	HG0
Proteobacteria	*Alphaproteobacteria*	*Caulobacteraceae*	*Caulobacter*	0.05	4.02	HG0
Proteobacteria	*Alphaproteobacteria*	*Rhodospirillaceae*	*Dongia*	0.01	3.93	HG0
Proteobacteria	*Betaproteobacteria*	*Comamonadaceae*	*Ramlibacter*	0.01	5.31	Water

A



B

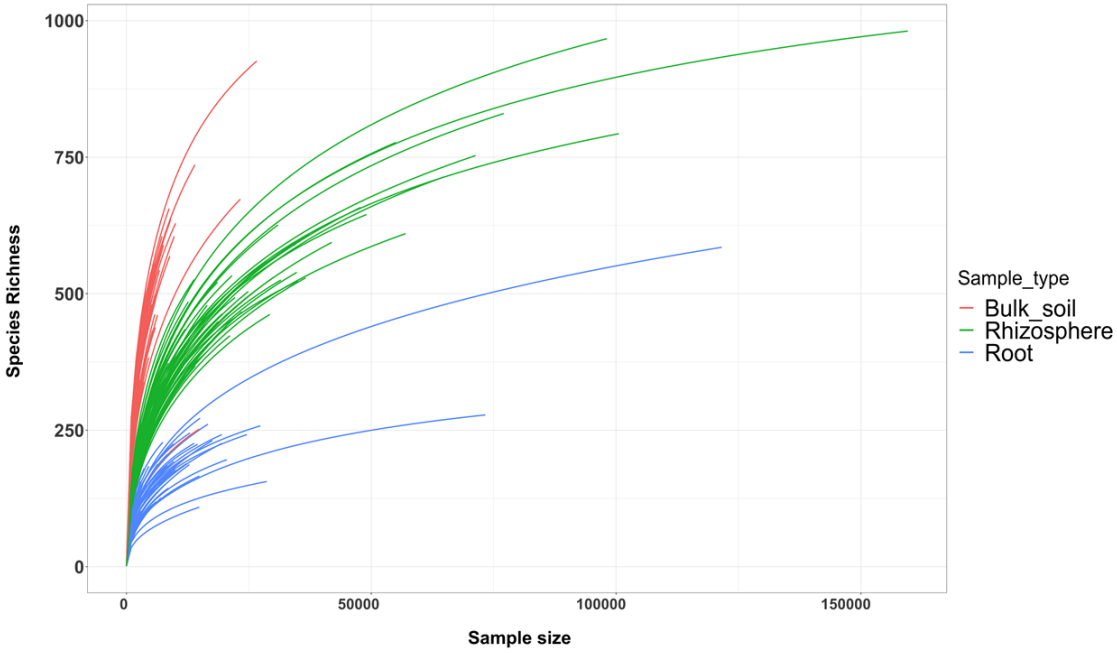


Fig S3.1 Rarefaction curves of the number of observed OTUs from 16S rRNA sequencing data. **A)** 247 samples collected in Plains, GA in 2018 and 2019; **B)** 210 samples collected in a greenhouse test in 2020

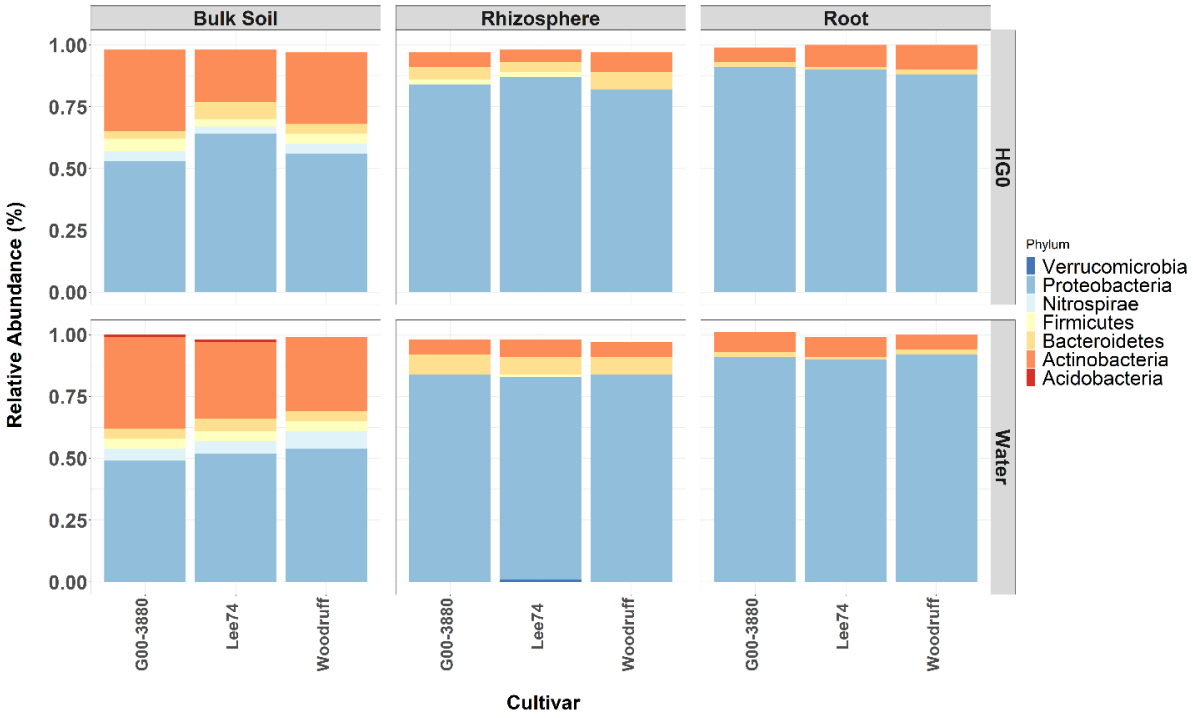


Fig S3.2 The distribution of core phyla of three types of samples: bulked soil, rhizosphere, and root from two treatments: water and HG type 0 (SCN race 3) in a greenhouse test in 2020

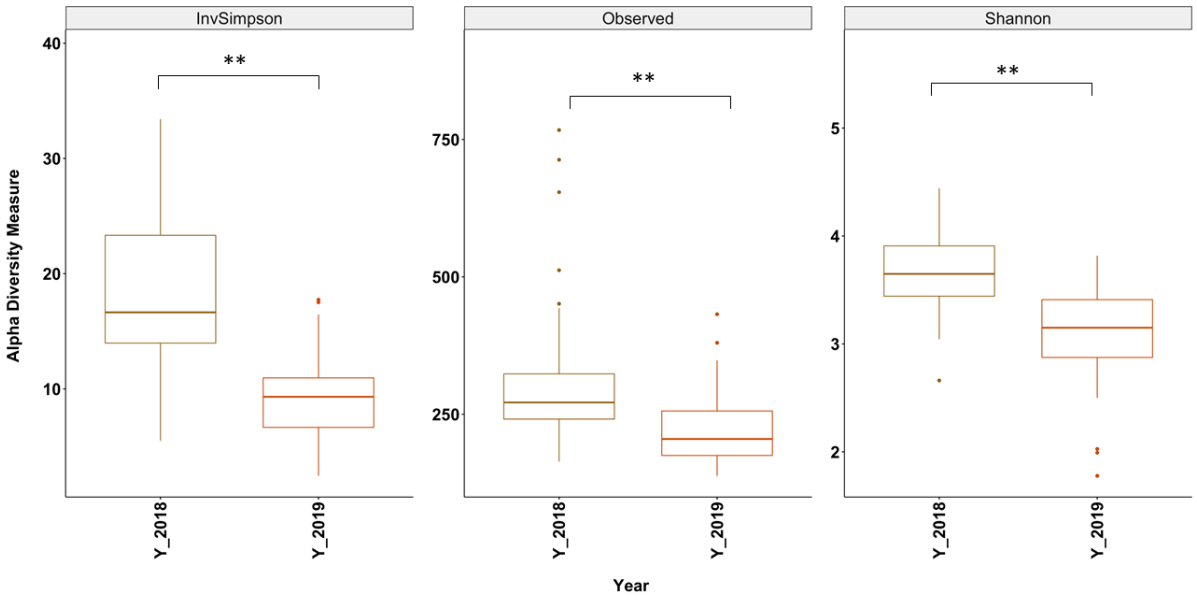


Fig S3.3 Alpha diversity using three types of measurements
Microbial diversity in root samples collected in 2018 had higher diversity than those collected in 2019. Turkey comparison was performed with P value < 0.01 (**).

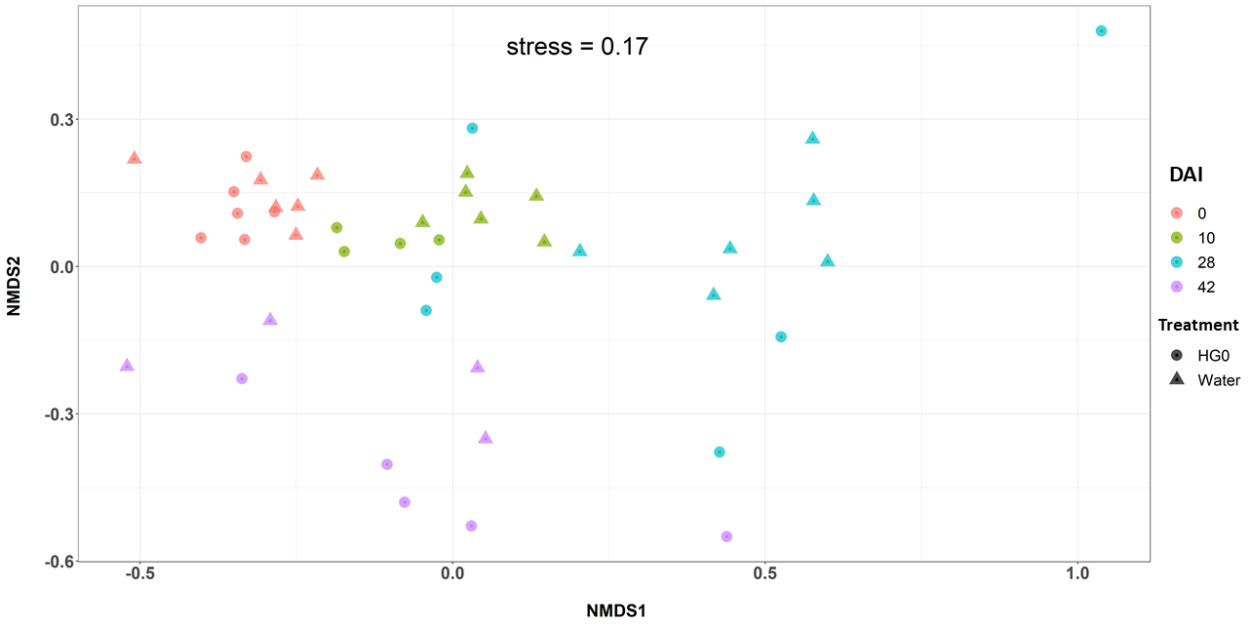


Fig S3.4 NMDS plot of unplanted pots in a greenhouse test in 2020.
NMDS plot based on the Bray Curtis distance comparing the β diversity between two treatments and four different sampling time points. Microbial communities in soil without soybean plants did not show strong separation between two treatments at all four time points with PERMANOVA: *P = 0.5*

CHAPTER 4

SUMMARY

Soybean cyst nematode (SCN) is the most destructive pest, affecting soybean production, therefore, breeding for SCN resistance is one of important goals in soybean breeding programs. However, due to the limited resources of resistance that carry two major QTLs *Rhg1* and *Rhg4*, SCN has adapted and overcome the resistance. Discovery of new sources of resistance will provide additional resources to combat SCN damages in soybean production. The objectives of this research were to: 1) map, fine map, and characterize the QTLs controlling for SCN resistance using bi-parental populations derived from a resistant source identified in previous study; and 2) determine composition of microbiome on the roots of soybean genotypes with resistance and susceptibility to SCN and investigate the impact of root microbiome on host plant resistance to SCN.

PI 567295, a plant introduction in maturity group VIII, is resistant to multiple SCN populations including HG type 0 (SCN race 3) and HG type 1.2.- (SCN race 2). $F_{2:3}$ and RIL populations derived from the same pedigree 'Bossier' × PI 567295 were developed for discovering genomic regions controlling for SCN resistance. The $F_{2:3}$ population was used for bulked segregant analysis (BSA). One hundred of twenty $F_{2:3}$ families were screened with HG type 0 and based on phenotypes of these families, four bulks including two resistant bulks and two susceptible bulks were created. These bulks along with two parents, 'Bossier' and PI 567295 were genotyped with SoySNP50K Infinium BreadChip. Six putative genomic regions on Chrs 2, 5, 10, 11, 14, and 19 were found, and the genomic regions on Chr 10 were promising due to the

high number of informative SNPs. Based on BSA results, informative SNPs were selected to design KASP markers that were included in mapping and fine mapping studies. One hundred and forty-one $F_{5:6}$ RILs were phenotyped with two SCN populations: HG type 0 and HG type 1.2.- with six replicates for each RIL per SCN population. The RIL population was genotyped with SoySNP6K Infinium Chip and ~100 designed KASP SNP markers. QTL analysis identified one major QTL on Chr 10 significantly associated with resistance to both SCN populations, and two minor QTLs on Chrs 6 and 8, respectively, controlling the resistance to HG type 1.2.-. All favorable alleles came from the resistant parent, PI 567295. Based on QTL results, two RILs, G19-SPT1748, highly resistant, and G19-SPT1719, highly susceptible to both SCN populations along with two parents were included in whole genome resequencing (WGRS) to identify candidate genes at the QTL regions. Based on WGRS results, 18, 5, and 6 candidate genes on Chrs 10, 6, and 8, respectively, were identified. Based on annotation and gene ontology from SoyBase databases, candidate genes that related to nematode and disease resistance were selected. SNPs within exons of these genes were selected to design KASP SNP markers to validate the association between genotypes and phenotypes. SNP marker GSM1021. Located within the 2nd exon of *Glyma.10g,199300* on Chr 10 was the most significant marker associated with resistance to both SCN populations. The second marker, GSM1019 that was located within the exon of *Glyma.10g201900* was significantly associated with SCN resistance. Both genes have functions related to intracellular transport. The most significant SNP marker, GSM1021 could be used as the functional marker for marker-assisted selection. Combining favorable alleles of multiple QTL have been shown to enhance SCN resistance in this study. Pyramiding favorable alleles of QTLs on Chrs 6, 8, and 10 could reduce FI down to 15.2%. The result will bring additional resources of resistance and QTLs from a different source of resistances that is

different from ‘Peking’ and PI 88788 at *Rhg1* and *Rhg4* loci to enable plant breeders to develop SCN cultivars with broad resistance to multiple SCN populations.

Interaction between roots and microbes for disease suppression, nutrition acquisition is one of key drivers for plant growth and development. Hence, identifying beneficial microbiomes that improves plant productivity is a promising approach in plant cultivation. In this study, we focused on investigating root-associated bacteria across different soybean genotypes related to SCN resistance and susceptibility using 16S rRNA sequencing. Thirteen and 14 soybean genotypes with contrasting traits of resistance and susceptibility to SCN were grown at Plains, GA with three replicates per genotype per year. Three types of samples, bulked soil, rhizosphere, and root were taken from each plot. Based on the sequencing analysis, both overlapping and distinct microbiomes were found in each type of these samples. *Proteobacteria, Actinobacteria* and *Bacteroidetes* were found to be the most abundant phylum in all three sample types. However, there was a transition pattern in some bacterial communities between root and soil samples, revealing that soybean had an ability to select bacteria from soil and combine with bacteria in roots to have their bacteria niche. Some potential microbiomes related to plant growth and disease resistance including *Balneola, Flavisolbacter, Lacibacter* and *Niastella* genera were found. To evaluate structure of bacterial communities related to SCN suppression, a greenhouse test with two treatments, mock control (water) and HG type 0 inoculation was performed. Three soybean genotypes including ‘Woodruff’ (‘Peking’-type resistance), ‘G00-3880’ (PI 88788-type resistance) and ‘Lee 74’ (susceptible) were included in this test. Three sample types, bulked soil, rhizosphere, and root of each genotype were collected at three time points, 10, 28 and 42 days after inoculation (DAI). Structure of bacterial communities of soybean microbiomes showed a variation by different sampling times. Of three time points, 10 DAI was the time that resistant

genotypes actively select bacteria from the soil to respond to SCN. Enrichment of some taxa *Dyella, Ramlibacter* and *Mycobacterium* genera was observed in ‘Woodruff’ and ‘G00-3880’. Based on α diversity, both ‘Woodruff’ and ‘G00-3880’ demonstrated a higher density level of bacteria than the susceptible, ‘Lee 74’ in the root samples. This result supports that root microbiome assembly plays a role in resistance to SCN.