MAPPING OF BROWN STEM ROT RESISTANCE, GENETIC YIELD GAINS, AND YIELD IMPACT OF SOYBEAN CYST NEMATODE

BY

KEITH FRANKLIN RINCKER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Brian Diers, Chair
Professor Randall Nelson
Professor Glen Hartman
Assistant Professor Alexander Lipka
Associate Professor Martin Bohn

ABSTRACT

Fine Mapping of Resistance Genes from Five Brown Stem Rot Resistance Sources in Soybean

Brown stem rot (BSR) of soybean [*Glycine max* (L.) Merr.] caused by *Cadophora gregata* (Allington & D.W. Chamb.) T.C. Harr. & McNew, can be controlled effectively with genetic host resistance. Three BSR resistance genes *Rbs1*, *Rbs2*, and *Rbs3* have been identified and mapped to a large region on chromosome 16. Marker-assisted selection (MAS) will be more efficient and gene cloning will be facilitated with a narrowed genomic interval containing an *Rbs* gene. The objective of this study was to fine map the positions of *Rbs* genes from five sources. Mapping populations were developed by crossing the resistant sources 'Bell', PI 84946-2, PI 437833, PI 437970, L84-5873, and PI 86150 with either the susceptible cultivar Colfax or Century 84. Plants identified as having a recombination event near *Rbs* genes were selected and individually harvested to create recombinant lines. Progeny from recombinant lines were tested in a *C. gregata* root-dip assay and evaluated for foliar and stem BSR symptom development. Overall, 4,878 plants were screened for recombination and progeny from 52 recombinant plants were evaluated with simple sequence repeat (SSR) genetic markers and assessed for symptom development. Brown stem rot resistance was mapped to intervals ranging from 0.34 to 0.04 Mb in the different sources. In all sources, resistance was fine mapped to intervals inclusive of BARCSOYSSR_16_1114 and BARCSOYSSR_16_1115, which provides further evidence that one locus provides BSR resistance in soybean.

Genome-Wide Association Study of Brown Stem Rot Resistance in Soybean across Multiple Populations

Genetic resistance to BSR of soybean, has been identified and mapped with biparental populations. Although nearly 400 accessions have been identified with BSR resistance, this trait has been mapped in only 12 sources, and just two, PI84946-2 and PI88788, have been utilized to develop BSR resistant cultivars. Thus, there is a serious need to improve our knowledge of the genetic basis of BSR resistance in soybean so that resistance genes in cultivars can be diversified and markers close to resistance genes can be identified and used in marker-assisted selection (MAS). To this end, we conducted a genome-wide association study (GWAS) to identify novel

genomic loci associated with BSR and to gain further insight into a previously-reported chromosome 16 region containing *Rbs* genes. A total of 52,041 single nucleotide polymorphisms (SNPs) were tested for association with BSR in a set of 4,735 accessions from four diversity panels evaluated for resistance from 1989 to 2003. Using a unified mixed linear model and stepwise model selection, we refined the signals within the *Rbs* interval on chromosome 16 by finding associations that explain a substantial proportion of the total variation of BSR resistance. In combination with significant GWAS signals found elsewhere in the genome, our study will aid efforts to improve BSR resistance by providing new targets for MAS.

Genetic Improvement of U.S. Soybean in Maturity Groups II, III, IV

Soybean improvement via plant breeding has been critical for the success of the crop. The objective of this study was to quantify genetic change in yield and other traits that occurred during the past 80 yr of North American soybean breeding in maturity groups (MGs) II, III, and IV. Historic sets of 60 MG II, 59 MG III, and 49 MG IV soybean cultivars, released from 1923 to 2008, were evaluated in field trials conducted in 17 U.S. states and one Canadian province during 2010 to 2011. Averaged over 27 MG II and MG IV and 26 MG III site-years of data, the estimated rates of yield improvement during the 80 yr were 23 kg ha^{-1} yr^{-1} for MGs II and III, and 20 kg ha^{-1} yr^{-1} for MG IV cultivars. However, a two-segment linear regression model provided a better fit to the data and indicated that the average current rate of genetic yield gain across MGs is 29 kg ha^{-1} yr^{-1}. Modern cultivars yielded more than old cultivars in all environments, but particularly in high-yielding environments. New cultivars in the historic sets used in this study are shorter in height, mature later, lodge less, and have seeds with less protein and greater oil concentration. Given that on-farm soybean yields in the United States are also increasing at a rate of 29 kg ha^{-1} yr^{-1}, it can be inferred that continual release of greater-yielding cultivars has been a substantive driver of the U.S. on-farm realized yield increases.

Impact of Soybean Cyst Nematode Resistance on Soybean Yield

Soybean cyst nematode (SCN) (*Heterodera glycines* Ichinohe) of soybean [*Glycine max* (L.) Merr.] causes extensive yield loss, and host resistance has been an effective strategy to minimize this loss. However, shifts in SCN population virulence compatibility have resulted from the

extensive use of PI 88788 as a source of resistance in the Northern U.S. and has the potential to reduce the effectiveness of this resistance source. The Northern Regional Soybean Cyst Nematode Tests offer a vast amount of yield testing combined with entry resistance screening and characterization of nematode host compatibility. The objectives of this study were to utilize previous yield testing to (i) quantify the impact of resistance as the initial field SCN egg counts increases, (ii) explore effects of maturity group (MG) on the impact of resistance, and (iii) gain insights into the effects of host compatibility on PI 88788 on yield. Yield testing from over 11 years was combined into a single dataset with over 1247 test-environment combinations. The yield advantage of SCN resistant entries increased as initial egg counts increased and a larger advantage was found in early MGs (00-II) than later MGs (III-IV). A yield advantage was documented at environments with an initial egg count of 100 eggs 100 cm^{-3} soil. At all levels of virulence on PI 88788, breeding lines with resistance from PI 88788 yielded more than susceptible entries. Predictions from this dataset offer a unique view of the impact SCN resistance provides in soybean and relationships among differing levels of virulence on PI 88788.

ACKNOWLEDGEMENTS

I thank my advisor Dr. Brian Diers for his guidance and support through my doctoral studies. I would also like to thank Dr. Donald Bullock, Dr. Glen Hartman, and Dr. Randall Nelson, Dr. Alex Lipka, and Dr. Martin Bohn for their guidance and thoughtful comments. In addition, I thank Dr. James Haudenshield for his guidance in plant pathology. Furthermore, I thank undergraduate research assistants for their help with my greenhouse tests in addition to current and previous graduate students Lillian Brzostowski, Russell Ward, Neil Yu, Abraham Akpertey, Tim Pruski and Carol Fox for their advice and fruitful discussions. I thank Troy Cary, John Meharry, and Sarah Schultz for the time they spent teaching me field, greenhouse, or lab activities.

Lastly, I thank my parents Dave and Elma for their support. Most importantly, I thank my wife Amanda and children for their support and sacrifices for me to explore my ambitions in plant science.

TABLE OF CONTENTS

CHAPTER ONE: Literature Review 1

CHAPTER TWO: Fine Mapping of Resistance Genes from Five Brown Stem Rot Resistance Sources in Soybean 23

CHAPTER THREE: Genome-Wide Association Study of Brown Stem Rot Resistance in Soybean across Multiple Populations 57

CHAPTER FOUR: Genetic Improvement of U.S. Soybean in Maturity Groups II, III, IV 87

CHAPTER FIVE: Impact of Soybean Cyst Nematode Resistance on Soybean Yield 118

CHAPTER ONE

Literature Review

BROWN STEM ROT RESISTANCE IN SOYBEAN

Brown stem rot (BSR) of soybean (*Glycine max* (L.) Merr.) caused by the soilborne fungus *Cadophora gregata* (Allington & Chamb.) T.C. Harr. & McNew, was first discovered in 1944 in Illinois and within four years had become important in Indiana, Iowa, Missouri, and Ohio (Allington and Chamberlain, 1948). Today this disease causes yield loss in the Midwestern states of the USA as well as Canada, Brazil, Argentina, and China (Wrather et al., 2010). Symptoms first appear in July or early August in Illinois as brown discoloration of stem pith and vascular elements (Allington and Chamberlain, 1948). As the disease progresses, the brown pith extends up the stem and foliar symptoms develop as interveinal chlorosis and necrosis. Damage to the stem as a result of fungal infection can lead to an increased amount of lodging (Allington and Chamberlain, 1948). In some environments, no foliar symptoms occur, or leaves may suddenly turn brown, dry, and remain attached at the petiole (Gray, 1974). Yield losses of up to 38% have been reported (Gray, 1972) and yield losses in the United States averaged 423,000 metric tons per year from 2006 to 2009 (Koenning and Wrather, 2010). Wrather et al. (2010) found yield loss in the top eight soybean producing countries to total 1,562,000 metric tons in 2006.

Cadophora gregata was placed in the Helotiales (discomycetes) of the Ascomycota through the use of ribosomal DNA sequences (Harrington and McNew, 2003). No sexual reproductive stages have been identified but asexual conidia are produced on phialides (Gray and Grau, 1999). Isolates have been classified by two methods. First, Gray (1971) described type I isolates as producing both stem pith browning and interveinal chlorotic symptoms in the leaves resulting from a toxin (Gray, 1974). Type II isolates are described as only having stem pith browning. A newer method to classify isolates uses fungal DNA polymorphisms (Chen et al., 2000) which is based on Gray's classification but uses primers to develop PCR products that determine the presence or absence of an insertion/deletion in the intergenic spacer region. Isolates with the larger PCR product are classified as genotype A, which has been found to be the same as type I

isolates and the smaller PCR product is genotype B, which is the same as type II classification (Hughes et al., 2002). Isolate Oh_2 is a type I (A) isolate from Ohio collected by Dr. Gray and is used frequently in greenhouse studies to evaluate both foliar and stem symptoms.

Cadophora gregata has been found to be non-pathogenic to maize (*Zea mays*), but adzuki bean (*Vigna angularis*) and mung bean (*Vigna radiata*) have been found to have a vascular disease similar to BSR. Overwintering occurs in crop residues as mycelium (Adee et al., 1997), and after the crop has been planted, infection occurs as mycelium enters the roots and slowly grows through the xylem and later the leaf vascular system (Gray and Grau, 1999). Disease development is greatest at 15-28° C and can be reduced at temperatures approaching 32° C (Gray, 1974; Gray and Grau, 1999; and Phillips, 1971). Disease symptoms develop faster after flowering and the pathogen can be isolated from stem vascular tissue before stem or leaf symptoms develop (Gray, 1974). During the R3 to R4 growth stages of a field planting, leaf symptoms become visible, and at the R5 growth stage assessments can be made to determine yield loss (Gray, 1974; Mengistu and Grau, 1987).

Brown stem rot disease development is not only influenced by temperature but also cultural factors and interactions with other diseases. Waller et al. (1992) found that low soil fertility increases disease development, and Mengistu and Grau (1987) studied the effects of irrigation and found that plants receiving post flowering irrigation had more foliar symptoms. Tillage has been found to reduce disease (Adee et al., 1997; Workneh et al., 1999). However, *C. gregata* has been found at significant levels in crop residue 16 months post-harvest (Impullitti and Grau, 2006). Therefore, crop rotations of 3 or more years are needed to reduce the impacts of this disease (Gray and Grau, 1999). Nonetheless, tillage and adequate fertility levels can be used to reduce the impact of this disease. Interactions with other diseases also influence BSR development. Tachibana and Card (1972) found that plants infected with soybean mosaic virus had half as much BSR symptoms compared to plants not infected with soybean mosaic virus. Tabor et al. (2003; 2006b) studied the interaction between soybean cyst nematode, *Heterodera glycines*, infection and BSR, and found that the incidence and severity of BSR are increased when plants are infected with *H. glycines*.

Considering the cultural practices available to producers, further control measures are needed. Some plant diseases can be controlled through the use of fungicides, however currently no fungicides offer protection to BSR. The most promising and cost effective control option for producers is the planting of resistant cultivars which has shown to prevent yield loss (Bachman et al., 1997b). Methods of screening plant introductions to identify BSR resistance have been developed in the field and greenhouse. Greenhouse screening is typically preferred over field tests because there is greater consistency of results from greenhouse tests compared to field tests. In the greenhouse, plants are primarily inoculated using a root dip method. With this method, plants are grown to the V2 stage (Fehr et al., 1971), the roots are trimmed, dipped into an inoculum, transplanted into a pot, and the remaining inoculum is poured into the soil (Patzoldt et al., 2003; Sebastian et al., 1985). Other greenhouse methods inoculate the plants by cutting the taproot and inserting mycelium into the wound (Gray, 1971), or using a needle to inject a conidia solution into the stem (Tabor et al., 2003, 2006a).

Multiple screens of the USDA Soybean Germplasm Collection for BSR resistance have identified almost 400 accessions with resistance similar to current resistant cultivars (Bachman et al., 1997a; Bachman and Nickell, 2000a; Chamberlain and Bernard, 1968; Hughes et al., 2004; Nelson et al., 1989; and Patzoldt et al., 2003). Nelson et al. (1989) tested over 3,400 accessions ranging from maturity group (MG) 000 to IV in the field and later tested putatively resistant entries in the greenhouse. Only three late maturing entries were resistant in all tests and several more were resistant at two or more locations. Bachman et al. (1997a) evaluated 559 accessions from central China ranging from MG II to IV in field and greenhouse tests. They found 13 accessions with resistance equal to resistant checks. More accessions were found to be resistant to three isolates of *C. gregata* f. sp. *sojae* in greenhouse conditions by Bachman and Nickell (2000a). Of the 829 accessions tested 241 were found to be resistant. These accessions ranged from MG IV to VIII from central and southern China. Patzoldt et al. evaluated 624 accessions from south central China in greenhouse tests (2003). Eighty-five accessions were identified as resistant after testing with one isolate. Eight more isolates were then used to identify ten resistant accessions ranging in MG IV to IX. Although, many accessions have been screened for disease resistance, public cultivars primarily incorporate resistance from plant introduction (PI) 84946-2. However, germplasm has been developed that contain BSR resistance from the

accessions PI 86150, PI 90138, PI 437821, and PI 437833 (Nickell and Bernard, 1992; Nickell et al., 1990; Nickell et al., 1994).

Three genes conferring resistance to BSR have been identified through genetic studies. *Rbs1* was identified in the germplasm line L78-4094 (Hanson et al., 1988), *Rbs2* in PI 437833 from Japan (Hanson et al., 1988), and *Rbs3* in PI 437970 from China (Willmot and Nickell, 1989). More genetic studies identified soybean accessions that contained one or more of these reported *Rbs* genes (Bachman and Nickell, 1997; Eathington and Nickell, 1994; Eathington et al., 1995; Lohnes and Nickell, 1995). Although these studies identified previously reported *Rbs* genes, often the data were variable and definitive results were elusive. Furthermore, *Rbs1* to *Rbs3* were given unique names because the original studies showed the genes were unlinked (Hanson et al., 1988; Wilmot and Nickell, 1989); however, genetic mapping studies have placed all three resistance genes onto chromosome 16 (linkage group (LG) J) near simple sequence repeat markers (SSR) Satt215 and Satt431 (Bachman et al., 2001; Lewers et al., 1999; Patzoldt et al., 2005b). Additional BSR resistance quantitative trait loci (QTL) mapped from eight soybean accessions from central and south-central China have been mapped by single marker analysis or interval mapping to the same region on chromosome 16 near Satt431, Satt547, or Satt244 (Patzoldt et al., 2005a; Perez et al., 2010). Before breeders attempt to pyramid BSR resistance QTL, the number of different QTL in the region must be determined. Another problem presented with current QTL information is that marker-assisted selection (MAS) is not as efficient as it could be due to a large interval containing the resistance QTL, which is 10.2 Mb of Williams 82 genome sequence.

Additional loci have been proposed to interact with *Rbs* loci or act as modifiers (Bachman and Nickell, 2000b; Hughes et al., 2004; Sebastian and Nickell, 1985; Waller et al., 1991). In a cross of the susceptible cultivars Asgrow A3127 and Elf, the resistant cultivar A3733 was released. Waller et al. (1991) explained the appearance of resistance from the two susceptible parents by proposing a two gene model with one acting as a major gene and the second acting as a modifier or minor gene. In addition, the BSR resistant cultivar Fayette was developed from the susceptible parents Williams and PI 88788 (Bernard et al., 1988), and the resistant germplasm line LN89-5717 was developed from the susceptible parents Williams 82 and PI 89772 (Hughes

et al., 2004). Each of these three examples has Williams or Williams 82, its isoline, in their parentage.

An additional example of BSR resistance likely controlled by gene interactions was described by Patzoldt et al. (2005b). They mapped a BSR resistance allele onto chromosome 16 in their mapping population, and they showed that this allele originated from PI 88788. However, they found that PI 88788 was significantly more susceptible than susceptible checks. Furthermore, a minor soybean cyst nematode (SCN) resistance QTL can be traced back to PI 88788 in the same region of chromosome 16 (Glover et al., 2004). Other researchers have noticed cultivars with SCN resistance from PI 88788 also have BSR resistance (Hughes et al., 2004; MacGuidwin et al., 1995). The preponderant use of PI 88788 as a source of SCN resistance in current soybean cultivars is exemplified by variety testing program entries such as those at the University of Illinois which list only ten of 336 SCN resistant MG II, III, and IV varieties contained SCN resistance from a source other than PI 88788 (Joos, et al., 2013). This presents PI 88788 as a possible major source of BSR resistance in SCN resistant cultivars. Genetic mapping of resistance loci continues and will provide further evidence to support or contradict multiple Rbs loci or modifiers.

The resistance mechanism to BSR is not known. The region of chromosome 16 where BSR resistance QTL has been located also contains the resistance QTL, *Rcs3*, which provides resistance to frogeye leaf spot (Mian et al., 1999), the minor QTL, cqSCN-003, which provides resistance to SCN (Glover et al., 2004), and *Rpp2*, which provides resistance to soybean rust (Yu et al., 2015). This suggests that a complex of disease resistance genes or a single gene capable of providing resistance to multiple pests is located on chromosome 16. To elucidate possible resistance mechanisms, Impullitti and Malvick (2014) used microscopy and fluorescently labeled *C. gregata* isolates. They suggested that resistant plants were able to produce more vascular vessels to compensate for the loss of vessel function. In addition, resistant plants were able to restrict the pathogen from the vascular system.

Linkage mapping in biparental populations has been frequently used to map traits; however these methods come with limitations. These include the time required to develop populations, high

linkage disequalibrium (LD), and small population sizes of most studies which all lead to low precision of QTL mapping (Cardon and Bell, 2001). Genome-wide association studies (GWAS) offers an alternative to linkage mapping for identifying QTL. Using hundreds or even thousands of plant accessions that capture many generations of recombination, association mapping typically has greater resolution for locating QTL compared to linkage mapping (Cardon and Bell, 2001). Furthermore, these accessions usually include more segregating loci than what is typically found in a biparental cross used in a linkage mapping study.

The use of GWAS to identify agronomically important QTL does have its concerns, mainly false discoveries (Complex Trait Consortium, 2003). These can arise due to population structure and unequal relatedness among accessions. Mixed linear models (MLM) have been demonstrated to account for population structure and unequal relatedness among accessions better than general linear model-based methods (Yu et al., 2006). In MLM models, population structure is fit as a fixed effect and kinship among individuals is incorporated as the variance-covariance structure of the random effect for the individuals (Yu et al., 2006). In an evaluation of mixed model approaches using the autogamous crop of wheat, Stich et al. (2008) found that mixed models including kinship were more appropriate than models with both population structure and kinship. Segura et al. (2012) further improved GWAS methods by including cofactors to control for large-effect loci. The use of multiple cofactors to control for large-effect loci is standard in traditional linkage mapping and has been found to improve QTL estimates compared to models without covariates (Jansen, 1993 and Zeng, 1994). A multi-locus mixed linear model (MLMM) was developed by Segura et al. (2012) to control large-effect loci and was found to increase power and reduce false discovery rate.

New genotyping technologies are making it possible to leverage phenotypic data and dissect the genetic architecture of agronomically important soybean traits, including disease resistance, with increased resolution. For example, the genotyping of 19,652 accessions in the USDA Soybean Germplasm Collection with the SoySNP50K Illumina Infinium chip, which contains 52,041 single nucleotide polymorphisms (SNP) (Song et al., 2013) is now complete. Phenotypic data on many of these accessions for BSR resistance and other traits are also available through the United States Department of Agriculture-Germplasm Resources Information Network (USDA-

GRIN) and from previous accession screening studies (Bachman et al., 1997a; Bachman and Nickell, 2000a; Nelson et al., 1989; and Patzoldt et al., 2003).

Reported values of LD in soybean have varied across genomic regions. Across a population of 52 Asian landraces, Hyten et al. (2007) found LD to decay to an r^2 of 0.1 in 90 kb and 300 kb in two separate genomic regions but not to decay past an r^2 of 0.1 in another genomic region. In addition, linkage disequilibrium was found to be greater in a population of elite U.S. cultivars which was expected due to the relatively short timeframe of breeding compared to many years of recombination after domestication to form current landraces. Hwang et al. (2014) utilized the SoySNP50K array and found coverage of one SNP every 17 kb in euchromatic regions while 100 kb in heterochromatic regions. With this coverage across the genome, Hwang et al. (2014) and Vaughn et al. (2014) identified QTL associated with soybean seed protein and oil concentration. In addition, Vaughn et al. (2014) relied on historical phenotypic data available in the USDA-GRIN. Both studies identified previously reported QTL and refined their genomic locations. Wen et al. (2014) also utilized the SoySNP50K array to identify novel loci and further narrow the genomic regions containing resistance QTL for sudden death syndrome of soybean (*Fusarium virguliforme*). In addition, associations with resistance to Sclerotinia stem rot of soybean (*Sclerotinia sclerotiorum*) were identified by Bastien et al. (2014). These studies demonstrate the adaptability of GWAS to soybean traits including disease resistance, the use of USDA-GRIN phenotypic data, and the utilization of the SoySNP50K array as a source of genotypic data.

GENETIC IMPROVEMENT OF U.S. SOYBEAN CULTIVARS

Soybean [*Glycine max* (L.) Merr.] production in the United States has evolved from a minor forage crop in the early 1900's to a crop with a harvest of 108 million metric tons in 2014 (USDA-NASS, 2015). Soybean is valued for its protein and oil content and currently an ingredient in numerous food, feed, and industrial products (Wilson, 2008). As our global population grows, the need for agricultural products will increase as well. The United Nations projects that the global population will be 9.7 billion by 2050, an increase from 7.3 billion in 2015 (United Nations, 2015). In addition, the increased consumption of meat in emerging

markets will create more demand for feed (Pingali, 2007). The impact of global soybean demand will place strain on all aspects of production including land area harvested, agronomic inputs, and genetic improvements of cultivars. In fact, the global rate of yield increase in soybean will need to almost double by 2050 to meet predicted demand for the crop (Ray et al., 2013). In addition, obtaining higher yields on current cropland is needed to reduce the loss of native habitat and the subsequent increase in greenhouse gas emissions that accompany further land clearing for new cropland (Tilman et al., 2011).

Soybean was domesticated in China (ca. 1700-1100 B.C.) and brought to the U.S. by Samual Bowen in 1765 (Hymowitz, 1990; Hymowitz and Harlan, 1983). In the early 1900s, farmers grew either PIs from East Asia or selections from these PIs. Later when demand for soybean grain increased, breeding for productivity and agronomic adaptability began. Cultivars from these breeding efforts were first released in the 1940s. As soybean production increased, public and proprietary breeding efforts expanded and began to include selection for pest resistance along with selection for yield (Carter et al., 2004). Proprietary breeding programs have continued to expand in number and in size, and these programs currently provide most of the soybean seed sold to farmers in the USA (Specht et al., 2014).

On-farm yield gains arise from the combined impact of grower adoption of new cultivars, improved cultural practices, interactions between new cultivars and improved cultural practices, and environmental factors such as increased atmospheric CO_2 levels (Long et al., 2006; Rowntree et al., 2013a, 2013b; Specht et al., 2014; Ziska and Bunce, 2007). Average on-farm soybean yields in the U.S. have increased from 738 kg ha^{-1} in 1924 to 3208 kg ha^{-1} in 2014 (USDA-NASS, 2015). Recent estimates of yearly on-farm yield gains have ranged from 22 kg ha^{-1} yr^{-1} to 31 kg ha^{-1} yr^{-1}(Specht et al., 2014 and Wilcox, 2001).

The biological limit of soybean yield potential is not known. Specht et al. (1999) suggested a biological maximum yield of 8,000 kg ha^{-1} and they based their projections of future yield potential with consideration of this maximum. However, this maximum yield obtained has been broken in multiple years by Kip Cullers in southeast Missouri. His record stands at 10,800 kg ha^{-1} produced in 2010 (Cubbage, 2010). While the level of yield potential realized by

yield contest winners is not expected across all production acres, the opportunity remains to increase yields with improved cultivars and agronomic inputs. Genetic yield gain has been estimated in a number of studies by growing soybean cultivars with different release years in common environments. These estimates have ranged from 11 kg ha^{-1} to 25 kg ha^{-1} from different North American producing regions (Boerma, 1979; De Bruin and Pedersen, 2008b; Salado-Navarro et al., 1993; Specht and Williams, 1984; Ustun et al., 2001; Voldeng et al., 1997; Wilcox et al., 1979). In addition to these studies showing that soybean genetic yield potential has improved, no evidence has been published of a yield plateau (De Bruin and Pedersen, 2008b; Specht et al., 1999; Ustun et al., 2001; Wilcox, 2001). The stability of yields across low and high yielding environments is another important aspect to evaluate in cultivars. De Bruin and Pedersen (2008b), Voldeng et al. (1997), and Wilcox et al. (1979) all determined that soybean yields have increased over generations of breeding without detectable reductions in yield stability.

As breeding programs continually select for improved yield potential and better adaptability to current agronomic practices and climate change, a number of factors can change the rate of genetic yield gain over time. For instance, the Plant Variety Protection (PVP) Act of 1970 has stimulated proprietary investment in soybean breeding programs by giving plant breeders better intellectual property protection of their cultivars and allowing a greater return on investment (Fehr, 1991; USDA, 2006). Other factors that can favorably affect the rate of gain are the increased capacity to manage field plots (Eathington et al., 2007) and an increased knowledge of DNA and RNA which allows greater genotyping capacity and understanding of gene function (Poland and Rife, 2012; Schmutz et al., 2010). Some changes in our environment are also expected to have an advantageous effect on yield potential such as a continued increase in atmospheric carbon dioxide concentration and warmer springs which enable earlier planting.

On the other hand, increases in atmospheric ozone concentrations and erratic weather patterns caused by climate change have the potential to decrease yield gains due to increased abiotic stresses (Betzelberger et al., 2012 and Hassol, 2009). An increase in selection criteria in addition to yield potential within a breeding program can also have a negative impact on yield gains as limited resources are diverted away from selection for yield. Examples include pest resistance

and transgenes. Gains in yield potential are also impacted by the available genetic diversity present in breeding programs. The narrowing of the genetic base of North American cultivars has been documented by Gizlice et al. (1994) and Sneller (1994). For example, 35 genotypes have accounted for 95% of the genes in the North American germplasm as of 1988 (Gizlice et al., 1994).

YIELD IMPACT OF SOYBEAN CYST NEMATODE RESISTANCE

Soybean cyst nematode (*Heterodera glycines* Ichinohe) is estimated to be the most damaging pest to soybean [*Glycine max* (L.) Merr.] production in the United States due to its widespread distribution and the yield losses it causes. Losses in the United States averaged 3,170,000 Mg annually during 2003 – 2009 (Wrather and Koenning, 2006 and Koenning and Wrather, 2010). Growers are often slow to respond to SCN infestations in fields because aboveground symptoms such as plant stunting often do not occur while yield impacting infections are present (Wang et al., 2003 and Young, 1996). Damage to soybean plants occurs when second stage juvenile nematodes hatch from eggs, penetrate roots, and feed from vascular tissue (Koenning, 2004). This results in structural changes in host plant cells which begin with the fusion of adjacent cells (Endo, 1991). The fusion of cells continues until a large multinucleate feeding site, called a syncytium forms (Niblack et al., 2004). As the lifecycle of SCN continues, females become swollen and remain attached to the roots but exposed on the surface. Males leave the root and then mate with exposed females. After mating, eggs are produced in cysts that offer protection for several years until optimum conditions are present for hatching (Koenning, 2004). Lauritis et al. (1983) found the life cycle of *H. glycines* to take 21 d to complete at 25 C in laboratory conditions.

Soybean cyst nematode was first described in 1915 in Japan by Hori, (1915) and first found in the U.S. in North Carolina during 1954 (Winstead et al., 1955). Its wide-spread introduction into the U.S. is believed to have occurred from movement of soil from Asia to soybean fields which served as inoculum for *Bradyrhizobium japonicum* (Kirchner) Jordan (Noel, 1992). Soybean cyst nematode is now found in all soybean producing states in the U.S. and also Argentina, Brazil, Canada, Columbia, Indonesia, and Taiwan (Niblack et al., 2004). The spread of SCN is

also attributed to wind currents, water movement, soil movement, and birds (Noel, 1992). Soon after the commercialization of genetic host resistance in soybean to SCN, physiological variation was reported among SCN populations, and a race scheme for characterization of SCN was developed (Ross, 1961; Golden et al., 1970; and Riggs and Schmitt, 1988). A separate scheme to characterize populations of SCN was developed to avoid confusion with the term race and its use on a population of organisms. This new scheme is a *Heterodera glycines* (HG) type classification system which separates the major genetic groups for host compatibility (Niblack et al., 2002). The HG type designation of a nematode population identifies what standard indicator lines it can reproduce on. Understanding the SCN HG type present in soil is important when producers select a resistant cultivar.

Since elimination of SCN populations from a field is not feasible, management of SCN by producers includes an integrated approach. The most effective methods to manage SCN are to utilize host resistance and rotate with non-host crops which both limit SCN reproduction (Niblack and Chen, 2004). However, cysts are capable of surviving several years in the absence of a host (Inagaki and Tsutsumi, 1971) and multiple legume crops and weed species can also serve as hosts (Riggs, 1992). Genetic resistance to SCN has been utilized successfully to manage SCN populations and resistant cultivars are available across MGs grown in the U.S. Cultivars with genetic resistance to populations of specific HG types have been shown to reduce the reproduction of SCN (Chen et al., 2001; De Bruin and Pedersen, 2008a; Koenning, 2000; and Wheeler et al., 1997). The effects of planting date and tillage on SCN populations have been found to be inconsistent across research studies, and management with nematicidal applications may not be a feasible option due to economic considerations, health concerns, and level of control (Niblack et al., 2004).

The first resistant soybean cultivar released in the U.S. was Pickett in 1967, which derived its resistance from Peking, a black seeded introduction from China (Brim and Ross, 1966). Additional sources of resistance to specific SCN populations were later identified and utilized in breeding programs. Two of these sources include PI 88788, which has been used frequently in the development of cultivars, and PI 437654, which at the time had resistance to all known SCN populations in the U.S. (Anand and Gallo, 1984 and Anand et al., 1988). The predominant

source of SCN resistance in northern U.S. cultivars is PI 88788. For instance, the University of Illinois Department of Crop Science Variety Testing program evaluated 336 SCN resistant entries within MG II, III, and IV in 2013, but only ten contained sources other than PI 88788 (Joos et al., 2013). The reason for the almost exclusive use of this source is due to the greater yielding resistant lines developed with PI 88788 resistance than with resistance from other sources. A likely outcome from the abundant use of resistance from PI 88788 is the selection of nematode populations that can overcome this resistance source. Surveys of soils in soybean producing regions have shown that up to 93% of nematode populations could reproduce on PI 88788 (Faghihi et al., 2010; Mitchum et al., 2007; and Niblack et al., 2008). In contrast, resistance from other sources such as PI 437654 remained effective, but development of high yielding resistant lines from other sources remains a challenge.

Resistant cultivars have repeatedly shown yield advantages compared to susceptible cultivars when SCN is present (Chen et al., 2001; De Bruin and Pedersen, 2008a; Niblack et al., 1992; Koenning, 2000; Donald et al., 2006; Wheeler et al., 1997; and Young, 1996). Furthermore, when initial egg density increases, a greater yield difference between resistant and susceptible cultivars has been found (Chen et al., 2001; Niblack et al., 1992; Koenning, 2000). Evaluation of resistant breeding lines at many locations is needed to identify high yielding resistant lines. For this reason, the Northern Regional Soybean Cyst Nematode Tests are conducted across the north-central soybean producing regions of the United States and Ontario to evaluate the yield performance of resistant germplasm from soybean breeders (Cary and Diers, 2014). Test locations range from non-infested fields to fields with high initial SCN egg counts. Data within these tests provide a tremendous resource of replicated trials growing both resistant lines and susceptible checks across MG 0 through IV.

LITERATURE CITED

Adee, E.A., C.R. Grau, and E.S. Oplinger. 1997. Population dynamics of *Phialophora gregata* in soybean residue. Plant Dis. 81:199-203.

Allington, W.B. and D.W. Chamberlain. 1948. Brown stem rot of soybean. Phytopathology 38:793-802.

Anand, S.C., K.M. Gallo. 1984. Identification of additional soybean germplasm with resistance to race 3 of the soybean cyst nematode. Plant Dis. 68:593-595.

Anand, S.C., K.M. Gallo, I.A. Baker, and E.E. Hartwig. 1988. Soybean plant introductions with resistance to races 4 or 5 of soybean cyst nematodes. Crop Sci. 28:563-564.

Bachman, M.S. and C.D. Nickell. 1997. Genetic analysis of brown stem rot resistance in soybean germplasm line LL89-605. Soybean Genet. Newsl. 24:113-115.

Bachman, M.S. and C.D. Nickell. 2000a. High frequency of brown stem rot resistance in soybean germ plasm from central and southern China. Plant Dis. 84:694-699.

Bachman, M.S. and C.D. Nickell, 2000b. Investigating the Genetic Model for Brown Stem Rot Resistance in Soybean. J. Heredity. 91:316-321.

Bachman, M.S., C.D. Nickell, P.A. Stephens, and A.D. Nickell. 1997a. Brown stem rot resistance in soybean germ plasm from central China. Plant Dis. 81:953-956.

Bachman, M.S., C.D. Nickell, P.A. Stephens, A.D. Nickell, and L.E. Gray. 1997b. The effect of *Rbs2* on yield of soybean. Crop Sci. 37:1148–1151.

Bachman, M.S., J.P. Tamulonis, C.D. Nickell, and A.F. Bent, 2001. Molecular markers linked to brown stem rot resistance genes *Rbs1* and *Rbs2* in soybean. Crop Sci. 41:527-535.

Bastien, M., H. Sonah, and F. Belzile. 2014. Genome wide association mapping of *Sclerotinia sclerotiorum* resistance in soybean with a genotype-by-sequencing approach. The Plant Genome 7(1). doi:10.3835/plantgenome2013.10.0030.

Bernard, R.L., G.R. Noel, S.C. Anand, and J.G. Shannon. 1988. Registration of Fayette soybean. Crop Sci. 28:1028-1029.

Betzelberger, A.M., C.R. Yendrek, J. Sun, C.P. Leisner, R.L. Nelson, D.R. Ort, and E.A. Ainsworth. 2012. Ozone exposure response for U.S. soybean cultivars: linear reductions in photosynthetic potential, biomass, and yield. Plant Physiol. 160:1827-1839.

Boerma, H.R. 1979. Comparison of past and recently developed soybean cultivars in maturity groups VI, VII, and VIII. Crop Sci. 19:611–613.

Brim, C.A., and J.P. Ross. 1966. Registration of Pickett soybeans. Crop Sci. 6:305-305.

Cardon L.R. and J.I. Bell. 2001. Association study designs for complex diseases. Nat. Rev. Genet. 2:91-99.

Carter, T.E., R.L. Nelson, C.H. Sneller, and Z. Cui. 2004. Genetic diversity in soybean. *In:* H.R. Boerma and J.E. Specht, editors, Soybeans: improvement, production, and uses. 3rd ed. Agron. Monogr. 16. ASA, CSSA, and SSSA, Madison, WI. p. 303-416.

Cary, T. and B. Diers (ed.) 2014. 2014 Northern regional soybean cyst nematode tests. Univ. of Illinois, Urbana.

Chamberlain, D.W. and R.L. Bernard. 1968. Resistance to brown stem rot in soybean. Crop Sci. 8:728-729.

Chen, W., C.R. Grau, E.A. Adee, and X. Meng. 2000. A molecular marker identifying subspecific populations of the soybean brown stem rot pathogen, *Phialophora gregata*. Phytopathology 90:875-883.

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

Complex Trait Consortium. 2003. The nature and identification of quantitative trait loci: a community's view. Nat. Rev. Genet. 4:911–916.

Cubbage, S. 2010. World soybean yield record shattered – again!. Missouri Soybean Farmer 15:6. http://mosoy.org/wp-content/uploads/news/WEB%20PDF-%20DEC%2010.pdf (accessed 25 Mar. 2016).

De Bruin, J.L. and P. Pedersen. 2008a. Response of old and new soybean cultivars to *Heterodera glycines* Ichinohe. Agron. J. 100:1347-1353.

De Bruin, J.L. and P. Pedersen. 2008b. Yield improvement and stability for soybean cultivars with resistance to *Heterodera glycines* Ichinohe. Agron. J. 100:1354-1359.

Donald, P.A., P.E. Pierson, S.K. St. Martin, P.R. Sellers, G.R. Noel, A.E. MacGuidwin, J. Faghihi, V.R. Ferris, C.R. Grau, D.J. Jardine, H. Melakeberhan, T.L. Niblack, W.C. Stienstra, G.L. Tylka, T.A. Wheeler, and D.S. Wysong. 2006. Assessing *Heterodera glycines*-resistant and susceptible cultivar yield response. J. Nematol. 38:76-82.

Eathington S.R., T.M. Crosbie, M.D. Edwards, R.S. Reiter, and J.K. Bull. 2007. Molecular markers in a commercial breeding program. Crop Sci. 47:S154-S163.

Eathington, S.R. and C.D. Nickell. 1994. Inheritance of brown stem rot resistance in PI 437685D. Soybean Genet. Newsl. 21:187-191.

Eathington, S.R., C.D. Nickell, and L.E. Gray. 1995. Inheritance of brown stem rot resistance in soybean cultivar BSR 101. J. Heredity 86:55-60.

Endo, B.Y. 1991. Ultrastructure of initial responses of susceptible and resistant soybean roots to infection by *Heterodera glycines*. Rev. Nematol. 14:73–94.

Faghihi, J., P. A. Donald, G. Noel, T. W. Welacky, and V. R. Ferris. 2010. Soybean resistance to field populations of *Heterodera glycines* in selected geographic areas. Plant Health Prog. doi:10.1094/PHP-2010-0426-01-RS.

Fehr, W.R., editor 1991. Principles of cultivar development: theory and technique. Vol.1 Macmillan Publishing Co., New York.

Fehr, W.R., C.E. Caviness, D.T. Burmood, and J.S. Pennington. 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. Crop Sci. 11:929–931.

Gizlice, Z., T.E. Carter, Jr., and J.W. Burton. 1994. Genetic base for North America public soybean cultivars released between 1947 and 1988. Crop Sci. 34:1143-1151.

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

Golden, A.M., J.M. Epps, R.D. Riggs, L.A. Duclos, J.A. Fox, and R.L. Bernard. 1970. Terminology and identity of infraspecific forms of the soybean cyst nematode (*Heterodera glycines*). Plant Dis. Rep. 54:544-546.

Gray, L.E. 1971. Variation in pathogenicity of *Cephalosporium gregatum* isolates. Phytopathology 61:1410–1411.

Gray, L.E. 1972. Effect of *Cephalosporium gregatum*on soybean yield. Plant Dis. Rep. 56:580–581.

Gray, L.E. 1974. Role of temperature, plant age, and fungus isolate in the development of brown stem rot in soybean. Phytopathology 64:94-96.

Gray, L.E. and C.R. Grau. 1999. Brown Stem Rot. *In*: G.L. Hartman, J.B. Sinclair, and J.C. Rupe, editors, Compendium of Soybean Diseases. 4th ed. APS Press, St. Paul, MN. p. 28-29.

Hanson, P.M., C.D. Nickell, L.E. Gray, and S.A. Sebastian. 1988. Identification of two dominant genes conditioning brown stem rot resistance in soybean. Crop Sci. 28:41-43.

Harrington, T.C. and D.L. McNew. 2003. Phlylogenetic analysis places *Thephialophora*-like anamorph genus *Cadophora* in the *Helotiales*. Mycotaxon 87:141-151.

Hassol, S.J. 2009. Global climate change impacts in the United States – 2009 Report. U.S. Global Change Research Program. Publication available on-line at: https://nca2009.globalchange.gov/agriculture/index.html (accessed 9 Mar. 2016).

Hori, S. 1915. Phytopathological notes. Sick soil of soybean caused by a nematode. J. Plant Protect. 2:927-930.

Hughes, T.J., W. Chen, and C.R. Grau. 2002. Pathogenic characterization of genotypes A and B of *Phialophora gregata* f. sp. *sojae*. Plant Dis. 86:729–735.

Hughes, T.J., N.C. Kurtzweil, B.W. Diers, and C.R. Grau. 2004. Resistance to brown stem rot in soybean germ plasm with resistance to the soybean cyst nematode. Plant Dis. 88:761-768.

Hwang, E., Q. Song, G. Jia,, J.E. Specht, D.L. Hyten, J. Costa, and P.B. Cregan. 2014. A genome-wide association study of seed protein and oil content in soybean. BMC Genomics 15:1. doi:10.1186/147-2164-15-1.

Hymowitz, T. 1990. Soybeans: The Success Story. *In:* J. Janick and J. Simon, editors, Advances in New Crops. Timber Press, Portland, OR. p. 159-163.

Hymowitz, T., and J.R. Harlan. 1983. Introduction of soybean to North America by Samual Bowen in 1765. Econ. Bot. 37:371-379.

Hyten, D.L., I. Choi, Q. Song, R.C Shoemaker, R.L. Nelson, J.M. Costa, J.E. Specht, and P.B. Cregan. 2007. Highly variable patterns of linkage disequilibrium in multiple soybean populations. Genetics 175:1937-1944.

Impullitti, A.E. and C.R. Grau. 2006. Population dynamics of *Phialophora gregata* in stem residue of a resistant and a susceptible soybean cultivar. Plant Dis. 90:759-764.

Impullitti, A.E. and D.K. Malvick. 2014. Anatomical response and infection of soybean during latent and pathologenic infection by type A and B of *Phialophora gregata*. Plos One 9(5):e98211.

Inagaki, H., and M. Tsutsumi. 1971. Survival of the soybean cyst nematode, *Heterodera glycines* Ichinohe (Tylenchida: Heteroderidae) under certain string conditions. Appl. Entomol. and Zoology, 6:156-162.

Jansen, R.C. 1993. Interval mapping of multiple quantitative trait loci. Genet. 135, 205–211.

Joos, D.K., R.W. Esgar, B.R. Henry, and E.D. Nafziger. 2013. Soybean variety test results in Illinois-2013. Crop Sci. Spec. Rep. 2013-04. Univ. of Illinois, Urbana.

Koenning, S.R. 2000. Density-dependent yield of *Heterodera glycines*-resistant and -susceptible cultivars. J. Nematol. 32:502-507.

Koenning, S.R. 2004. Population biology. *In*: D.P. Schmitt, J.A. Wrather, and R.D. Riggs editors. Biology and management of soybean cyst nematode, 2nd ed. Schmitt and Associates of Marceline. Marceline, MO. p. 73-88.

Koenning, S.R. and J.A. Wrather. 2010. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. Plant Health Prog. doi:10.1094/PHP-2010-1122-01-RS.

Lauritis, J.A., R.V. Rebois, and L.S. Graney. 1983. Development of *Heterodera glycines* Ichinohe on soybean, *Glycine max* (L.) Merr., under gnotobiotic conditions. J. Nematol. 15:272–281.

Lewers, K.S., E.H. Crane, C.R. Bronson, J.M. Schupp, P. Keim, and R.C. Shoemaker. 1999. Detection of linked QTL for soybean brown stem rot resistance in 'BSR 101' as expressed in a growth chamber environment. Mol. Breeding 5:33-42.

Lohnes, D.G. and C.D. Nickell. 1995. Inheritance of brown stem rot resistance in soybean isolines L68-0327 and L68-0469 and germplasm lines L84-5873 and L84-5932. Soybean Genet. Newsl. 22:203-207.

Long, S.P., E.A. Ainsworth, A.D.B. Leakey, J. Nosberger, and D.R. Ort. 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. Science 312:1918-1921.

MacGuidwin, A.E., C.R. Grau, and E.S. Oplinger. 1995. Impact of planting ‘Bell’, a soybean cultivar resistant to *Heterodera glycines*, in Wisconsin. J. Nematol. 27:78–85.

Mengistu, A. and C.R. Grau. 1987. Season progress of brown stem rot and its impact on soybean productivity. Phytopathology 77:1521-1529.

Mian, M.A., T. Wang, D.V. Phillips, J. Alvernaz, and H.R. Boerma. 1999. Molecular Mapping of the *Rcs3* Gene for Resistance to Frogeye Leaf Spot in Soybean. Crop Sci. 39:1687-1691.

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

Nelson, R.L., C.D. Nickell, J.H. Orf, H. Tachibana, E.T. Gritton, C.R. Grau, and B.W. Kennedy. 1989. Evaluating soybean germplasm for brown stem rot resistance. Plant Dis. 73:110-114.

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

Niblack, T.L., N.K. Baker, and D.C. Norton. 1992. Soybean yield losses due to *Heterodera glycines* in Iowa. Plant Dis. 76:943-948.

Niblack, T.L. and S.Y. Chen. 2004. Cropping systems and crop management practices. *In*: D.P. Schmitt, J.A. Wrather, and R.D. Riggs editors. Biology and management of soybean cyst nematode, 2nd ed. Schmitt and Associates of Marceline. Marceline, MO. p. 181-206.

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

Niblack, T.L., G.L. Tylka, and R.D. Riggs. 2004. Nematode Pathogens of Soybean. *In:* H.R. Boerma and J.E. Specht, editors, Soybeans: improvement, production, and uses. 3rd ed. Agron. Monogr. 16. ASA, CSSA, and SSSA, Madison, WI. p. 821-851.

Nickell, C.D. and R.L. Bernard. 1992. Registration of L84-5873 and L84-5932 soybean germplasm lines resistant to brown stem rot. Crop Sci. 32:835.

Nickell, C.D., P.M. Hanson, L.E. Gray, D.J. Thomas, and D.B. Willmot. 1990. Registration of soybean germplasm lines LN86-1595 and LN86-1947 resistant to brown stem rot. Crop Sci. 30:241.

Nickell, C.D., S.M. Lim, S. Eathington, and R. Warsaw. 1994. Registration of soybean germplasm line LL89-605, resistant to brown stem rot and bacterial blight. Crop Sci. 34:1134.

Noel, G.R. 1992. History, distribution, and economics. *In:* R.D. Riggs and J.A. Wrather editors, Biology and management of the soybean cyst nematode. APS Press, St. Paul, MN. p. 1-14.

Patzoldt, M.E., S.R. Carlson, and B.W. Diers. 2005a. Characterization of resistance to brown stem rot of soybean in five accessions from central China. Crop Sci. 45:1092-1095.

Patzoldt, M.E., W. Chen, and B.W. Diers. 2003. Evaluation of soybean plant introductions from China for resistance to brown stem rot. Plant Health Prog. doi:10.1094/PHP-2003-0702-01-RS.

Patzoldt, M.E., C.R. Grau, P.A. Stephens, N.C. Kurtzweil, S.R. Carlson, and B.W. Diers. 2005b. Localization of a quantitative trait locus providing brown stem rot resistance in the soybean cultivar Bell. Crop Sci. 45:1241-1248.

Perez, P.T., B.W. Diers, P. Lundeen, G.M. Tabor, and S.R. Cianzio. 2010. Genetic analysis of new sources of soybean resistance to brown stem rot. Crop Sci. 50:2431-2439.

Phillips, D.V. 1971. Influence of air temperature on brown stem rot of soybean. Phytopathology 61:1205-1208.

Pingali, P. 2007. Westernization of Asian diets and the transformation of food systems: Implications for research and policy. Food Policy 32: 281–298.

Poland, J.A. and T.W. Rife. 2012. Genotyping-by-sequencing for plant breeding and genetics. Plant Gen. 5:92-102.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield trends are insufficient to double crop production by 2050. PLoS ONE 8:doi.10.1371/journal.pone.0066428.

Riggs, R.D. 1992. Host range. In: R.D. Riggs and J.A. Wrather editors. Biology and management of the soybean cyst nematode. APS Press, St. Paul, MN. p.107-114.

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

Ross, J.P. 1961. Physiological strains of *Heterodera glycines*. Plant Dis. Rep. 46:766-769.

Rowntree, S.C., J.J. Suhre, N.H. Weidenbenner, E.W. Wilson, V.M. Davis, S.L. Naeve, S.N. Casteel, B.W. Diers, P.D. Esker, J.E. Specht, and S.P. Conley. 2013a. Genetic gain × management interactions in soybean: I. Planting date. Crop Sci. 53:1128:1138.

Rowntree, S., J. Suhre, E. Wilson, V. Davis, B. Diers, S. Casteel, P. Esker, and S. Conley. 2013b. Physiological and phenological responses of historical soybean cultivar releases to earlier planting. Crop Sci. 54:804-816.

Salado-Navarro, L.R., T.R. Sinclair, and K. Hinson. 1993. Change in yield and seed growth traits in soybean cultivars released in the southern USA from 1945 to 1983. Crop Sci. 33:1204–1209.

Schmutz, J., S. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D. Hyten, Q.J. Song, J.J. Thelen, J. Cheng, D. Xu, U. Hellsten, G. May, Y. Yu, T. Sakurai, T. Umezawa, M. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P.B. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, and S.A. Jackson. 2010. A genome sequence of paleopolyploid soybean (*Glycine max*). Nature 463:178–183.

Sebastian, S.A. and C.D. Nickell. 1985. Inheritance of brown stem rot resistance in soybean. J. Heredity 76:194-198.

Sebastian, S.A., C.D. Nickell, and L.E. Gray. 1985. Efficient selection for brown stem rot resistance in soybeans under greenhouse screening conditions. Crop Sci. 25:753–757.

Segura, V., B.J. Vilhjálmsson, A. Platt, A. Korte, Ü. Seren, Q. Long, and M. Nordborg. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet. 44: 825–830.

Sneller, C.H. 1994. Pedigree analysis of elite soybean lines. Crop Sci. 34:1515-1522.

Song Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS ONE 8(1): e54985. doi:10.1371/journal.pone.0054985.

Specht, J.E., B.W. Diers, R.L. Nelson, J.F. Toledo, J.A. Torrion, and P. Grassini. 2014. Soybean. *In*:, J.S.C. Smith, B. Diers, B. Carver, and J. Specht, editors, Yield gains in major U.S. field crops. CSSA Spec. Publ. 33. ASA, CSSA, and SSSA, Madison, WI. p. 311-356.

Specht, J.E., D.J. Hume, and S.V. Kumudini. 1999. Soybean yield potential–a genetic and physiological perspective. Crop Sci. 39:1560-1570.

Specht, J.E. and J.H. Williams. 1984. Contribution of genetic technology to soybean productivity—Retrospect and prospect. *In:* W.R. Fehr, editor, Genetic contribution to yield gains of five major crop plants. CSSA Special Publ. 7. CSSA and ASA, Madison, WI. p. 49-74.

Stich, B., J. Möhring, H.P. Piepho, M. Heckenberger, E.S. Buckler, and A.E. Melchinger. 2008. Comparison of mixed-model approaches for association mapping. Genetics, 178:1745-1754.

Tabor, G.M., S.R. Cianzio, G.L. Tylka, R. Roorda, and C.R. Bronson. 2006a. A new greenhouse method to assay soybean resistance to brown stem rot. Plant Dis. 90:1186-1194.

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

Tabor, G.M., G.L. Tylka, and C.R. Bronson. 2006b. Soybean stem colonization by genotypes A and B of *Cadophora gregata* increases with increasing population densities of *Heterodera glycines.* Plant Dis. 90:1297-1301.

Tachibana, H. and L.C. Card. 1972. Brown stem rot resistance and its modification by soybean mosaic virus in soybeans. Phytopathology 62:1314-1317.

Tilman, D., C. Balzer, J. Hill, and B.L. Befort. 2011. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. USA 108:20260-20264.

United Nations, Department of Economic and Social Affairs, Population Division. 2015. World population prospects: The 2015 revision, key findings and advance tables. Working Paper No. ESA/P/WP.241.

USDA. 2006. Plant Variety Protection Act and Regulations and Rules of Practice. [Online] Available at http://www.ams.usda.gov/sites/default/files/media/Plant Variety Protection Act.pdf, (accessed 9 Mar. 2016). Agriculture Marketing Service.

USDA, ARS, National Genetic Resources Program. *Germplasm Resources Information Network - (GRIN)*. [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. Available: https://npgsweb.ars-grin.gov/gringlobal/search.aspx (accessed 9 Mar. 2016).

USDA-National Agricultural Statistics Service (USDA-NASS). 2015. Data and statistics. USDA, NASS, Washington, DC. http://www.nass.usda.gov/Data_and_Statistics (accessed 9 Mar. 2016).

Ustun, A., F.L. Allen, and B.C. English. 2001. Genetic progress in soybean of the U.S. Midsouth. Crop Sci. 41:993-998.

Vaughn, J.N., R.L. Nelson, Q. Song, P.B. Cregan, and Z. Li. 2014. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. G3 4(11):223-2294.

Voldeng, H.D., E.R. Cober, D.J. Hume, C. Gillard, and M.J. Morrison. 1997. Fifty-eight years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 37:428–431.

Waller, R.S., C.D. Nickell, D.L. Drzycimski, and J.E. Miller. 1991. Genetic analysis of the inheritance of brown stem rot resistance in the soybean cultivar Asgrow A3733. J. Heredity 82:412-417.

Waller, R.S., C.D. Nickell, and L.E. Gray. 1992. Environmental effects on the development of brown stem rot in soybean. Plant Dis. 76:454-457.

Wang, J., T.L. Niblack, J.A. Tremain, W.J. Wiebold, G.L. Tylka, C.C. Marett, G.R. Noel, O. Meyers, and M.E. Schmidt. 2003. Soybean cyst nematode reduces soybean yield without causing obvious aboveground symptoms. Plant Dis. 87:623-628.

Wen, Z., R. Tan, J. Yuan, C. Bales, W. Du, S. Zhang, M.I. Chilvers, C. Schmidt, Q. Song, P.B. Cregan, and D. Wang. 2014. Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. BMC Genomics 15:809.

Wheeler, T.A., P.E. Pierson, C.E. Young, R.M. Riedel, H.R. Wilson, J.B. Eisley, A.F. Schmitthenner, and P.E. Lipps. 1997. Effect of soybean cyst nematode (*Heterodera glycines*) on yield of resistant and susceptible soybean cultivars grown in Ohio. J. Nematol. 29:703-709.

Winstead, N.N., C.B. Skotland, and J.N. Sasser. 1955. Soybean cyst nematode in North Carolina. Plant Dis. Rep. 39:9-11.

Wilcox, J.R. 2001. Sixty years of improvement in publicly developed elite soybean lines. Crop Sci. 41:1711-1716.

Wilcox, J.R., W.T. Schapaugh, Jr., R.L. Bernard, W.R. Fehr, and M.H. Niehaus. 1979. Genetic improvement of soybeans in the Midwest. Crop Sci. 19:803–805.

Wilmot, D.B. and C.D. Nickell. 1989. Genetic analysis of brown stem rot resistance in soybean. Crop Sci. 29:672-674.

Wilson, R.F. 2008. Soybean: market driven research needs. *In:* G. Stacey, editor, Genetics and Genomics of Soybean. Springer Science+Business Media, LLC, New York, NY. p. 3-15.

Workneh, F., X.B. Yang, and G.L. Tylka. 1999. Soybean brown stem rot, *Phytophthora sojae*, and *Heterodera glycines* affected by soil texture and tillage relations. Phytopathology 89:844-850.

Wrather, J.A. and S.R. Koenning. 2006. Estimates of disease effects on soybean yields in the United States 2003 to 2005. J. Nematol. 38:173-180.

Wrather, A., G. Shannon, R. Balardin, L. Carregal, R. Escobar, G.K. Gupta, Z. Ma, W. Morel, D. Ploper, and A. Tenuta. 2010. Effect of diseases on soybean yield in the top eight producing countries in 2006. Plant Health Prog. doi:10.1094/PHP-2010-0125-01-RS.

Young, L.D. 1996. Yield loss in soybean caused by *Heterodera glycines*. J. Nematol. 28:604-607.Wen, Z., R. Tan, J. Yuan, C. Bales, W. Du, S. Zhang, M.I. Chilvers, C. Schmidt, Q. song, P.B. Cregan, and D. Wang. 2014. Genome wide association mapping of quantitative resistance to sudden death syndrome in soybean. BMC Genomics 15:809. doi:10.1186/1471-2164-15-809.

Yu, J., G. Pressoir, W.H. Briggs, I. Vroh Bi, M. Yamasaki J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203–208.

Yu, N., M. Kim, Z.R. King, D.K. Harris, J.W. Buck, Z. Li, and B.W. Diers. 2015. Fine mapping of Asian soybean rust resistance gene Rpp2 from soybean PI 230970. Theor. Appl. Genet. 128:387-396.

Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. Genet. 136, 1457–1468.

Ziska, L.H. and J.A. Bunce. 2007. Predicting the impact of changing CO_2 on crop yields: some thoughts on food. New Phytol. 175:607-618.

CHAPTER TWO

Fine Mapping of Resistance Genes from Five Brown Stem Rot Resistance Sources in Soybean [1]

INTRODUCTION

Causal Organism and Host Resistance

Brown stem rot of soybean caused by the soilborne fungus *Cadophora gregata* affects soybean production in the northern United States, Canada, and Brazil and also has a minor impact in China (Wrather et al., 2010). Yield losses of up to 38% have been reported (Gray, 1972), and damage to the U.S. soybean crop was estimated to average 422,000 Mg annually from 2006 to 2009 (Koenning and Wrather, 2010). Management of this disease can be achieved with host genetic resistance (Bachman et al., 1997) and long-term crop rotation (Adee et al., 1997). Multiple screens of germplasm to identify resistance have been conducted; however, introgression of BSR resistance genes into cultivars has only been achieved with the resistance sources PI 84946-2 (Tachibana et al., 1980) and PI 88788 (Patzoldt et al., 2005b).

Three genes conferring resistance to BSR have been identified through genetic studies. *Rbs1* was identified in the germplasm line L78-4094 (Hanson et al., 1988), *Rbs2* in PI 437833 (Hanson et al., 1988), and *Rbs3* in PI 437970 (Willmot and Nickell, 1989). Additional genetic studies identified soybean accessions that contained one or more resistance genes at these loci (Bachman and Nickell, 1997; Eathington and Nickell, 1994; Eathington et al., 1995; Lohnes and Nickell, 1995). Although these studies identified previously reported *Rbs* genes, often the data were variable and definitive results were elusive. Furthermore, *Rbs1* to *Rbs3* were given unique names because the original studies showed the genes were unlinked (Hanson et al., 1988; Willmot and Nickell, 1989); however, genetic mapping studies have placed all three resistance genes onto chromosome 16 (linkage group [LG] J) near SSR markers Satt215 and Satt431 (Bachman et al., 2001; Lewers et al., 1999). Additional BSR resistance quantitative trait loci (QTL) from the cultivar Bell (Nickell et al., 1990b) and eight soybean accessions from central and south–central

[1] Previously published: Rincker K., G.L. Hartman, B.W. Diers. 2016. Fine Mapping of Resistance Genes from Five Brown Stem Rot Resistance Sources in Soybean. Plant Gen. 9(1). doi: 10.3835/plantgenome2015.08.0063. Copyright owner has provided permission to reprint.

China have been mapped with single marker analysis or interval mapping to the same region on chromosome 16 near Satt431, Satt547, or Satt244 (Patzoldt et al., 2005a,b; Perez et al., 2010). Patzoldt et al. (2005b) used near isogenic lines to confirm the BSR resistance QTL from Bell on chromosome 16 between the SSR markers Satt529 and Satt244, which is an interval of 23.1 cM (Song et al., 2004) or 10.2 Mb of cultivar Williams 82 genome sequence. Additional SSR markers (Song et al., 2010) have been developed and positioned on the soybean genome, which will aid in fine mapping the resistance QTL. The resistance allele in the cultivar Bell has been traced back to the BSR-susceptible landrace PI 88788 (Patzoldt et al., 2005b), which could contain a different resistant mechanism than other *Rbs* sources.

The major source of BSR resistance in modern cultivars is the South Korean landrace PI 84946-2, which is the source of resistance for the germplasm line L78-4094. After the initial identification of *Rbs1* in L78-4094, further studies identified PI 84946-2 as having both *Rbs1* and *Rbs3* (Eathington et al., 1995; Lohnes and Nickell, 1995). With molecular markers, Lewers et al. (1999) mapped two linked BSR resistance QTL on chromosome 16. Their mapping population included the cultivar BSR 101 (Tachibana et al., 1987) as its resistance source, which derives its resistance from PI 84946-2. This result is consistent with the mapping by Bachman et al. (2001) of a resistance gene from L78-4094. Bachman et al. (2001) also mapped to the same region a resistance gene from PI 437833. This accession had been identified as containing *Rbs2*, but it has not been used in public cultivars and only in the development of two germplasm lines (Nickell et al., 1990a).

Other germplasm identified with BSR resistance that warrant further study includes PI 86150, which was first identified as resistant by Chamberlain and Bernard (1968) and confirmed by Tachibana and Card (1972). The germplasm line L84-5873 (Nickell and Bernard, 1992) was developed as having PI 86150 as a resistance source and has not been currently used in public cultivars. The accession PI 437970 was identified as having *Rbs3* (Willmot and Nickell, 1989); however, its resistance has not been mapped or used in germplasm. Interestingly, Nelson et al. (1989) found this accession not to have levels of resistance as great as the resistant checks although variability in *C. gregata* virulence among testing locations was observed.

Before breeders attempt to pyramid BSR resistance QTL, we must determine whether there are different QTL alleles or loci on chromosome 16. Another problem presented with current QTL information is that MAS is not as efficient as it could be due to a large interval containing the resistance QTL, which, from Bell, spans 23.1 cM (Song et al., 2004) or 10.2 Mb of the Williams 82 genome sequence. Fine mapping of BSR resistance loci is needed to improve efficiency of MAS, to facilitate gene cloning, and to provide evidence whether BSR resistance is controlled by separate loci or a common locus among resistance sources. Marker-assisted selection can be used in backcrossing programs or to increase the favorable allele frequencies in breeding populations, thereby reducing resources needed to grow and evaluate plants or lines that do not carry optimal allelic combinations. In addition, MAS allows the pyramiding of resistance genes without having to inoculate plants with specific pathogens. Furthermore, the cloning of BSR resistance genes will be more successful after these genes are mapped into small intervals containing few candidate genes.

The objective of this study was to fine map the location of known BSR resistance loci on chromosome 16. The resistance sources Bell, PI 84946-2, PI 437833, PI 437970, PI 86150, and L84-5873 were used because of their inclusion in previous genetic studies and because of their parentage in germplasm and cultivar development.

MATERIALS AND METHODS

Population Development

The BSR resistant sources Bell, PI 84946-2, PI 437833, PI 437970, PI 86150, and L84-5873 were chosen for fine mapping studies. These resistance sources were crossed with the susceptible cultivars Colfax (Graef et al., 1994) or Century 84 (Walker et al., 1986) to develop mapping populations (Table 2.1). Fine mapping was conducted by first testing plants that segregated for the region on chromosome 16 where *Rbs* genes were previously mapped with genetic markers flanking the gene. Plants identified as having recombination events in this interval were selected, individually threshed, and their progeny were phenotyped for resistance. The source of segregating plants included a population of F_2 plants from the cross between Bell and Colfax, or F_3 plants from selected $F_{2:3}$ lines in all other populations (Table 2.1). For all populations except

Bell × Colfax, about 140 F_2 plants were grown in the field and threshed to develop $F_{2:3}$ lines and those lines segregating for markers across the chromosome 16 interval where *Rbs* genes map were selected. The identified recombinant F_2 and F_3 plants were screened with additional SSR markers (Cregan et al., 1999; Song et al., 2004) located on chromosome 16 to map the positions of the recombination events and to select plants with unique recombination positions within the *Rbs* interval. All selected F_3 plants and Bell × Colfax-selected F_2 plants were homozygous on one side of the recombination point and heterozygous on the other side. Selected F_3 plants were threshed individually to develop $F_{3:4}$ lines, and F_2 plants in the Bell × Colfax population were threshed individually to develop $F_{2:3}$ lines. Individual plants from these lines were inoculated with *C. gregata* and evaluated for BSR symptoms and a segregating genetic marker. Association between a molecular marker and symptom development for each recombinant line was then used to narrow the interval containing the resistance locus. Significant associations of foliar symptoms, stem symptoms, or both indicated that the resistance locus was on the heterozygous side of the recombination point. As test results narrowed the fine mapping interval, screening for recombination events took place within the narrowed region. For this reason, fewer recombinant lines were developed in populations other than Bell × Colfax (Table 2.1).

DNA Extraction

Genomic DNA from selected plants was extracted with the cetyltrimethyl ammonium bromide (CTAB) method described by Keim et al. (1988) with modifications. Specifically, leaf tissue from one or two newly expanding trifoliolates was collected into a 2.0-mL tube, freeze dried, and crushed with three (4 mm each) glass beads for 4 min on a modified paint can shaker. Six hundred microliters of CTAB extraction buffer was added to the macerated tissue and incubated for 1 h at 65°C. After cooling for 10 min, 600 µL of chloroform/isoamyl alcohol (24:1) was added to each tube, gently mixed, and then spun at 10,000 x gravity for 15 min. The aqueous layer was transferred to a new 1.7-mL tube. Five hundred microliters of isopropanol was used to precipitate the DNA, and the DNA was pelleted by spinning at 10,000 *g* for 10 min. The isopropanol was poured off and 500 µL of 70% ethanol was added to the DNA pellet for 5 min at room temperature and then spun at 10,000 x gravity. The ethanol was poured off, and the DNA pellet was allowed to dry. The DNA pellet was resuspended in 100 µL of 0.1× tris

ethylenediaminetetraacetic acid buffer and diluted 10-fold before use in polymerase chain reaction (PCR). DNA from individual plants during recombinant screening and disease testing was extracted from unifoliolates before full expansion by a quick DNA extraction method described by Bell-Johnson et al. (1998) or CTAB extraction as described earlier.

Simple-sequence repeat markers developed by Cregan et al. (1999) and Song et al. (2004, 2010) were used to genotype the samples. Markers developed by Song et al. (2010) named BARCSOYSSR are abbreviated as B hereafter. Polymerase chain reaction was performed according to Cregan and Quigley (1997) and gel electrophoresis was used to analyze PCR products in 6% (w/v) nondenaturing polyacrylamide gels (Wang et al., 2003). Amplification of SSR markers from Song et al. (2010) used the temperatures and durations of 94°C for 4 min, followed by 32 cycles of 94°C for 45 s; 58°C for 45 s; 68°C for 45 s; and a 7-min extension at 72°C.

Inoculations and Disease Evaluation

Cadophora gregata isolate *Oh*$_2$ was obtained from Dr. C. Grau, University of Wisconsin, and was originally collected from soybean growing in Ohio by Dr. L.E. Gray. *Oh*$_2$ is classified as a pathotype I (A) isolate (Gray, 1971; Hughes et al., 2002) and was chosen because of its ability to consistently produce stem symptoms in the greenhouse. Green bean agar (Chen et al., 2000) was used to culture the fungus and seed extract broth served as an inoculum, as described by Gray (1971) and Patzoldt et al. (2003). The liquid broth culture was prepared by autoclaving 100 mL of seed of the susceptible soybean cultivar Century 84 in 300 mL of distilled water. The liquid was strained and distilled water was added to one liter. Liquid cultures were maintained in the dark at 22°C, and inoculum was prepared by blending liquid cultures. The initial conidia and mycelia fragment concentration was measured with a hemocytometer and distilled water was added to adjust the concentration to 1.2×10^6 propagules mL^{-1}. Carboxylmethylcellulose was added at a rate of 7.5 g L^{-1} and blended into the inoculum. For each selected $F_{2:3}$ or $F_{3:4}$ line, up to 46 seedlings were grown in sand flats in a greenhouse until growth stage V1 (Fehr et al., 1971). Root dip inoculations (Sebastian et al., 1985; Patzoldt et al., 2003) were conducted by selecting five uniform plants, rinsing the roots in water, blotting them dry, dipping them in

50 mL of inoculum for 2 min, transplanting the plants into a 15-cm clay pot as described by Patzoldt et al. (2003), and repeating the process for the remaining plants of the line. After inoculation, the plants, in groups of five per pot, were arranged in a completely randomized design with parents and PI 88788 as check genotypes.

Plants were grown with a 14-h photoperiod and an average temperature of 18°C at night and 24°C during the day. Pots were watered from below by capillary action and each pot was fertilized weekly from above with 150 mL of water containing 0.09 g of each N, P_2O_5, K_2O; 0.5 mg chelated Fe; 0.2 mg of each chelated Cu, Mn, Zn; 0.09 mg B; and 4.0 μg Mo. When most plants were at growth stage R1 to R3, which was about 6 to 8 wk post-inoculation, BSR symptoms were measured. Foliar symptoms were measured by counting, from the base, the number of nodes on each main stem that foliar symptoms had progressed. This measurement was reported as the proportion of total nodes with leaves that had abscised or showed symptoms. Stems were then split longitudinally and the number of nodes with brown pith was counted and reported as the proportion of the total nodes on the main stem. Preliminary studies with noninoculated checks showed that they were not symptomatic and therefore were not included in subsequent evaluations. Inoculations of recombinant lines occurred in sets including a range of one to 10 lines at a time during 2011 to 2014. Recombinant lines were only tested once with the exception of two Bell × Colfax lines, which were used to validate results and narrow the interval screened for recombinants.

Associations between individual molecular markers and BSR symptoms of each recombinant line were tested with a single-factor analysis of variance in SAS v9.3 PROC MIXED and verified, if needed, with a Kruskal Wallis test in SAS v9.3 PROC NPAR1WAY (SAS Institute, 2011). Analysis of variance assumptions of homogeneous residual variance and normality of residuals were evaluated, since phenotypic data consists of a proportion of diseased plant, which potentially violates the assumption of normally distributed residuals. For this reason, the Kruskal–Wallis test was used to verify the analysis of variance results. A recombinant line with poor germination, resulting in <20 plants, was evaluated in two BSR tests, and data were analyzed in an analysis of variance model that included the test number as a variable.

RESULTS

Fine Mapping

Fine mapping of *Rbs* genes was initiated in the Bell × Colfax population by screening 200 F_2 plants with the markers Satt529 and Satt547, which flank the known *Rbs* interval. Sixty-eight recombinants were identified, grown to maturity, individually harvested to develop $F_{2:3}$ lines, and tested with additional SSR markers to map the position of the recombination point in each F_2 plant. Seven recombinant lines were selected for BSR testing because their recombination sites were located near Satt244 (Table 2.2). No associations between resistance and markers were found with lines 10262-1-19, 10262-1-32, 10262-2-2, and 10262-2-6, while significant associations were observed with foliar and stem symptoms of lines 10262-1-22 and 10262-1-30 and only foliar symptoms of line 10262-2-1 (Tests 1, 2, 4, 6; Table 2.3). This mapped the resistance QTL to between B_16_1102 and B _16_1134 (Table 2.2). Since later recombinant screening efforts were based on this narrowed interval, the lines 10262-1-22 and 10262-1-32 were tested for BSR phenotype again, and results were verified (Test 3; Table 2.3). The markers B_16_1100 and B_16_1142 were used to screen plants and identify recombinants in the second and further rounds of screening in the Bell × Colfax population. In all, 1218 F_2 plants were screened and 108 $F_{2:3}$ lines were developed (Table 2.1). Next, three more recombinant lines were selected, and no association was found in recombinant line 10262-2-7 (Test 5; Table 2.3). Significant associations were obtained with foliar and stem symptoms of recombinant lines 10262-2-8 and 10262-2-9 (Tests 11, 15; Table 2.3). This mapped the resistance QTL from Bell to between B_16_1105 to B_16_1118, which is a 0.25-Mb interval based on the Glyma 2.0 assembly (Table 2.2).

Fine mapping in the Century 84 × PI 84946-2 and PI 86150 × Century 84 populations began by screening 200 F_3 plants from each cross with the markers Satt215 (chromosome 16, 28.9 Mb) and Satt547 (Century 84 × PI 84946-2) and Satt622 (chromosome 16, 27.9 Mb) and B_16_1152 (chromosome 16, 33.8 Mb) (PI 86150 × Century 84). Further screening from these two populations and other populations was completed using the marker pairs B_16_1092 and Satt547 or B_16_1100 and B_16_1134, which were selected based on the interval identified in the

Bell × Colfax mapping population. Selected recombinant individuals were treated the same as the Bell × Colfax population and $F_{3:4}$ lines were subject to BSR assay tests (Table 2.1).

Within the Century 84 × PI 84946-2 (*Rbs1* and *Rbs3* source) population, six recombinant lines were selected for BSR testing. No association was obtained with BSR testing of the 278-1-5-4 line when analyzed across BSR Tests 4 and 5; however, significant associations were obtained with foliar and stem symptoms with lines 278-1-20-6, 278-1-18-8, 278-1-17-3, 278-1-9-4 and 278-1-5-20 (Tests 4, 5, 10, 12; Table 2.3). This mapped the resistance QTL from PI 84946-2 to between B_16_1098 and B_16_1123 (0.34 Mb) based on the Glyma 2.0 assembly (Table 2.4). Lewers et al. (1999) reported a second resistance allele on chromosome 16 in BSR101 and originating from PI 84946-2. Based on the map location in Lewers et al. (1999), this allele is expected to be located above the major QTL as depicted on Table 2.4. No significant ($P > 0.05$) marker associations with BSR resistance were found in the recombinant lines 278-1-18-6 and 278-1-18-7, which were chosen to search for this second QTL, since they were fixed for susceptible alleles in the QTL containing interval and segregating above this region (Table 2.4; Table 2.3 [Test 12]). In addition, no significant ($P > 0.05$) marker associations were found with recombinant lines 278-1-17-5 and 278-1-11-1, which segregated below the interval identified (Table 2.4; Table 2.3 [Tests 12, 15]).

Within the Century 84 × PI 437833 (*Rbs2* source) population, seven recombinant lines were selected for BSR testing (Table 2.5). Significant associations of foliar and stem symptoms were obtained with lines 263-5-17-1, 263-5-4-5, and 263-5-21-3, while only stem symptoms were significant with line 263-5-21-1 (Tests 8, 13; Table 2.3). No association was found with lines 263-5-14-7 and 263-5-13-1, and inconclusive evidence was obtained with line 263-5-16-1 (Test 8, 13; Table 2.3). The results from testing the progeny of line 263-5-16-1 was consistent with resistance segregating in the line, with plants homozygous for the marker allele from the resistant parent showing less disease than plants homozygous for the susceptible allele, but the *F*-Test was not significant ($P > 0.05$) (Table 2.3). In addition, only one progeny plant in the line showed stem symptoms, and this plant was heterozygous for the region. Unfortunately, no additional seeds of 263-5-16-1 were available to increase sample size or to test the line again. However, the results of recombinant lines 263-5-21-1 and 263-5-17-1 map the *Rbs2* gene from PI 437833 to an

interval between B_16_1105 and B_16_1115 (0.21 Mb) based on the Glyma 2.0 assembly (Table 2.5).

Eight recombinant lines within the Century 84 × PI 437970 (*Rbs3* source) population were selected for BSR testing (Table 2.6). No significant association was obtained with lines 264-6-11-2, 264-6-6-3, or 264-6-6-4 (Tests 9, 13; Table 2.3). Significant associations of foliar and stem symptoms were obtained with lines 264-6-11-6, 264-6-4-8, 264-6-5-3, 264-6-5-4, and 264-6-6-2 (Tests 9, 11, 12, 13; Table 2.3). These results place *Rbs3* from PI 437970 to the 0.21 Mb interval between B_16_1105 and B_16_1115, which is the same interval *Rbs2* maps (Table 2.6).

Six recombinant lines within the L84-5873 (resistance from PI86150) × Century 84 population were selected for BSR testing (Table 2.7). No significant association was obtained with the lines 261-1-6-10, 261-1-6-9, or 261-1-4-5, and significant associations of foliar and stem symptoms were obtained in BSR testing with lines 261-1-9-9, 261-1-11-9, and stem symptoms of 261-1-4-7 (Tests 11, 13; Table 2.3). This maps the resistance QTL allele from L84-5873 to an interval between B_16_1113 and B_16_1115 (0.05 Mb) based on the Glyma 2.0 assembly (Table 2.7).

To map resistance from PI 86150, the recombinant line 262-1-18-2 from the cross PI 86150 × Century 84 was selected for BSR testing because of a recombination site located in the lower side of the original 10.2 Mb interval identified in previous mapping studies (Table 2.8). Next, eight recombinant lines with recombination sites located near or within the fine-mapped interval from the Bell × Colfax population were selected for BSR testing. Significant associations of both foliar and stem symptoms were obtained from testing with lines 262-1-17-11, 262-1-2-11, and 262-1-17-13, while only foliar symptoms of 262-1-23-2 and the stem symptoms of 262-1-17-12 were significantly ($P > 0.05$) associated with a segregating marker (Test 14; Table 2.3). No significant association was obtained in BSR testing of lines 262-1-18-2, 262-1-22-1, 262-1-18-11, or 262-1-18-13 (Tests 6, 14; Table 2.3). These results map the resistance QTL from PI 86150 to between B_16_1114 and B_16_1115 (0.04 Mb) based on the Glyma 2.0 assembly (Table 2.8). This position is consistent with the mapping of the resistance allele from L84-5873, which is expected considering PI 86150 is the source of the resistance allele for L84-5873.

The assumptions of normality of residuals and homogeneity of residual variance were evaluated for each of the foliar and stem symptom BSR test analyses. Assumptions were violated more often in the analysis of stem than foliar data, and overall violations occurred in about half of the analyses; however, analyses using the Kruskal–Wallis tests did not change any interpretation of the analysis of variance tests.

Mapping within the PI 86150 × Century 84 population provided the most precision of all backgrounds tested, and the interval identified, B_16_1114 and B_16_1115, is inclusive of all mapping population intervals identified in this study. Because the physical position of the markers used in this fine mapping was based on the Williams 82 soybean genome version Glyma.Wm82.a2, we used the compatible genome annotation browser (http://soybase.org/gb2/gbrowse/gmax2.0/) and found four genes with predicted function in the fine-mapped *Rbs* region. Three of the genes are predicted to be nucleotide-binding site leucine-rich repeat (NBS-LRR) genes, which are the most common plant resistance genes, while the fourth is a 60S ribosomal protein L22p/L17e.

DISCUSSION

Brown stem rot resistance to the Oh_2 isolate was fine mapped within the same region on chromosome 16 (B_16_1098 and B_16_1123) from five resistance sources. The smallest interval mapped was a 0.04-Mb interval between B_16_1114 and B_16_1115 from PI 86150. These results are consistent with previous mapping of BSR resistance that showed that resistance from almost all sources map to the same region on chromosome 16 near Satt244 (Bachman et al., 2001; Patzoldt et al., 2005a,b; Perez et al., 2010). An exception to this is that of Perez et al. (2010), where they found that resistance did not map to chromosome 16 from one of the four sources tested, although this source was not fully resistant. A second exception is that of Lewers et al. (1999), who mapped two QTL on chromosome 16 from the soybean cultivar BSR101, which obtains its resistance from PI 84946-2. Lewers et al. (1999) found a minor QTL located near the restriction fragment length polymorphism (RFLP) marker B122 and a major QTL located near the RFLP marker K375. By using the Soybean Consensus Maps (Cregan et al.,

1999; Song et al., 2004), we positioned both QTLs to be above Satt244 and the minor QTL to be above the major QTL as shown in Table 2.4. We assume that we mapped the major *Rbs* resistance gene that Lewers et al. (1999) mapped from PI 84946-2 to the B_16_1098 and B_16_1123 interval, which is consistent with their mapping of BSR resistance to near K375. We found no conclusive evidence of the second minor QTL. Recombinant lines 278-1-5-4 and 278-1-17-5 did approach the significance threshold of 0.05 for stem symptoms, indicating that a minor QTL could be segregating in the populations; however, this was not supported by recombinant line 278-1-11-1 (Table 2.3, 4). The results from 263-5-16-1, within the Century 84 × PI 437833 population, potentially indicate a gene above B_16_1113 as shown in Table 2.5. However, we believe it is more likely that the results of this test are inconclusive, and unfortunately, no more seeds of that line remain for further testing. Care should be taken when inferring our results to field environments, since our greenhouse inoculation methods included a monoconidial strain of *C. gregata*, which may not be predictive of performance in all field environments. There is evidence, however, that results from greenhouse tests with Oh_2 can predict field performance in at least some environments, as Patzoldt et al. (2005b) mapped BSR resistance QTL from Bell to the same region on chromosome 16 using greenhouse tests inoculated with this isolate and field tests in two naturally infested field locations.

Before this study, the smallest interval defined as containing a BSR resistance QTL was 10.2 Mb of Williams 82 genome sequence between Satt529 and Satt244 in a Bell × Colfax population (Patzoldt et al., 2005b). They were not able to further map resistance as a result of a lack of additional useful markers. However, the availability of the soybean genome sequence and predicted candidate BARCSOYSSR markers greatly assisted the current mapping efforts (Schmutz et al., 2010; Song et al., 2010). The BARCSOYSSR markers near *Rbs* genes were used to precisely map the locations of the recombination events, which are required for fine mapping.

The BSR resistance genes *Rbs1*, *Rbs2*, and *Rbs3* were identified in classical genetic studies (Hanson et al., 1988; Willmot and Nickell, 1989). Although original research found each gene to be unlinked, molecular mapping placed all three genes to the same region on chromosome 16 (LG J) (Bachman et al., 2001; Lewers et al., 1999; Patzoldt et al., 2005b). It is possible that allelism tests of Hanson et al. (1988) and Willmot and Nickell (1989) did not have adequate

control of environmental effects, which may have led to inaccurately determining that F_2 plants or $F_{2:3}$ families were susceptible. Our research has confirmed that each of the three *Rbs* genes reside in the same region of chromosome 16, and we have further narrowed the interval containing these genes from 10.2 Mb to an overall 0.34 Mb. These results were further supported by using genome-wide association mapping and data from previous BSR screening efforts to consistently identify a region on chromosome 16 across multiple mapping panels (Rincker et al., 2016). The single-nucleotide polymorphism (ss715624573) identified in two of their association panels with the lowest *p*-value resides 0.02 Mb below the B_16_1114 and B_16_1115 interval identified in this study. It is possible that the three *Rbs* genes are in fact one gene that is located between B_16_1114 and B_16_1115, as no evidence of genes at separate loci was found in this study. Our study was not able to determine if different resistance sources used in this study have different alleles at the same locus.

The fine-mapped intervals in populations with resistance sources Bell, PI 84946-2, PI 437833, and PI 437970 contains 20 to 28 predicted genes in the Glyma 2.0 assembly. Ten of these predicted genes are NBS-LRR genes spanning 0.14 Mb and belong to the protein families PF08263 and PF00560. The fine-mapped interval in the population L84-5873 × Century 84 contains six predicted genes of which five are NBS-LRR genes. All fine-mapped intervals were inclusive of the four predicted genes located in the 0.04-Mb interval between B_16_1114 and B_16_1115 identified in the PI 86150 × Century 84 population. Three of these are NBS-LRR genes belonging to the protein families PF08263 and PF00560. These three genes, *Glyma.16 g169600*, *Glyma.16 g169700*, and *Glyma.16 g169900*, are highly homologous with the *Arabidopsis thaliana* (L.) Heynh. genes *AT1G45616.1*, *AT2G33060.1*, and *AT2G34930.1*, respectively. Since the genome sequence is based on the susceptible cultivar Williams 82, resequencing within this region from a resistance source is needed to identify the sequence of candidate genes from this source. Efforts to clone an *Rbs* gene can now focus on an interval of only 0.04 Mb.

LITERATURE CITED

Adee, E.A., C.R. Grau, and E.S. Oplinger. 1997. Population dynamics of *Phialophora gregata* in soybean residue. Plant Dis. 81:199–203. doi:10.1094/PDIS.1997.81.2.199

Bachman, M.S., and C.D. Nickell. 1997. Genetic analysis of brown stem rot resistance in soybean germplasm line LL89-605. Soybean Genet. Newsl. 24:113–115.

Bachman, M.S., C.D. Nickell, P.A. Stephens, A.D. Nickell, and L.E. Gray. 1997. The effect of *Rbs2* on yield of soybean. Crop Sci. 37:1148–1151. doi:10.2135/cropsci1997.0011183X003700040019x

Bachman, M.S., J.P. Tamulonis, C.D. Nickell, and A.F. Bent. 2001. Molecular markers linked to brown stem rot resistance genes, *Rbs1* and *Rbs2*, in soybean. Crop Sci. 41:527–535. doi:10.2135/cropsci2001.412527x

Bell-Johnson, B., G. Garvey, J. Johnson, D. Lightfoot, and K. Meksem. 1998. Biotechnology approaches to improving resistance to SCN and SDS: Methods for high throughput marker assisted selection. Soybean Genet. Newsl. 25:115–117.

Chamberlain, D.W., and R.L. Bernard. 1968. Resistance to brown stem rot in soybean. Crop Sci. 8:728–729. doi:10.2135/cropsci1968.0011183X000800060026x

Chen, W., C.R. Grau, E.A. Adee, and X. Meng. 2000. A molecular marker identifying subspecific populations of the soybean brown stem rot pathogen, *Phialophora gregata*. Phytopathology 90:875–883. doi:10.1094/PHYTO.2000.90.8.875

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, J. Chung, and J.E. Specht. 1999. An integrated genetic linkage map of the soybean. Crop Sci. 39:1464–1490. doi:10.2135/cropsci1999.3951464x

Cregan, P.B., and C.V. Quigley. 1997. Simple sequence repeat DNA marker analysis. In: G. Caetano-Anolles, et al., editors, DNA markers: Protocols, applications and overviews. John Wiley & Sons, New York. p. 173-185.

Eathington, S.R., and C.D. Nickell. 1994. Inheritance of brown stem rot resistance in PI 437685D. Soybean Genet. Newsl. 21:187–191.

Eathington, S.R., C.D. Nickell, and L.E. Gray. 1995. Inheritance of brown stem rot resistance in soybean cultivar BSR 101. J. Hered. 86:55–60.

Fehr, W.R., C.E. Caviness, D.T. Burmood, and J.S. Pennington. 1971. Stage of development descriptions for soybeans, *Glycine max* (L.). Merrill. Crop Sci. 11:929–931. doi:10.2135/cropsci1971.0011183X001100060051x

Graef, G.L., J.E. Specht, L.L. Korte, and D.M. White. 1994. Registration of ‘Colfax’ soybean. Crop Sci. 34:818. doi:10.2135/cropsci1994.0011183X003400030040x

Gray, L.E. 1971. Variation in pathogenicity of *Cephalosporium gregatum* isolates. Phytopathology 61:1410–1411. doi:10.1094/Phyto-61-1410

Gray, L.E. 1972. Effect of *Cephalosporium gregatum*on soybean yield. Plant Dis. Rep. 56:580–581.

Hanson, P.M., C.D. Nickell, L.E. Gray, and S.A. Sebastian. 1988. Identification of two dominant genes conditioning brown stem rot resistance in soybean. Crop Sci. 28:41–43. doi:10.2135/cropsci1988.0011183X002800010010x

Hughes, T.J., W. Chen, and C.R. Grau. 2002. Pathogenic characterization of genotypes A and B of *Phialophora gregata* f. sp. *sojae*. Plant Dis. 86:729–735. doi:10.1094/PDIS.2002.86.7.729

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

Koenning, S.R., and J.A. Wrather. 2010. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. Plant Health Prog. doi:10.1094/PHP-2010-1122-01-RS

Lewers, K.S., E.H. Crane, C.R. Bronson, J.M. Schupp, P. Keim, and R.C. Shoemaker. 1999. Detection of linked QTL for soybean brown stem rot resistance in 'BSR 101' as expressed in a growth chamber environment. Mol. Breed. 5:33–42. doi:10.1023/A:1009634710039

Lohnes, D.G., and C.D. Nickell. 1995. Inheritance of brown stem rot resistance in soybean isolines L68-0327 and L68-0469 and germplasm lines L84-5873 and L84-5932. Soybean Genet. Newsl. 22:203–207.

Nelson, R.L., C.D. Nickell, J.H. Orf, H. Tachibana, E.T. Gritton, C.R. Grau, and B.W. Kennedy. 1989. Evaluating soybean germplasm for brown stem rot resistance. Plant Dis. 73:110–114. doi:10.1094/PD-73-0110

Nickell, C.D., and R.L. Bernard. 1992. Registration of L84-5873 and L84-5932 soybean germplasm lines resistant to brown stem rot. Crop Sci. 32:835. doi:10.2135/cropsci1992.0011183X003200030065x

Nickell, C.D., P.M. Hanson, L.E. Gray, D.J. Thomas, and D.B. Willmot. 1990a. Registration of soybean germplasm lines LN86-1595 and LN86-1947 resistant to brown stem rot. Crop Sci. 30:241. doi:10.2135/cropsci1990.0011183X003000010070x

Nickell, C.D., G.R. Noel, D.J. Thomas, and R. Waller. 1990b. Registration of 'Bell' soybean. Crop Sci. 30:1364–1365.

Patzoldt, M.E., S.R. Carlson, and B.W. Diers. 2005a. Characterization of resistance to brown stem rot of soybean in five accessions from central China. Crop Sci. 45:1092–1095. doi:10.2135/cropsci2004.0393

Patzoldt, M.E., W. Chen, and B.W. Diers. 2003. Evaluation of soybean plant introductions from China for resistance to brown stem rot. Plant Health Prog. doi:10.1094/PHP-2003-0702-01-RS

Patzoldt, M.E., C.R. Grau, P.A. Stephens, N.C. Kurtzweil, S.R. Carlson, and B.W. Diers. 2005b. Localization of a quantitative trait locus providing brown stem rot resistance in the soybean cultivar ‘Bell’. Crop Sci. 45:1241–1248. doi:10.2135/cropsci2003.0615

Perez, P.T., B.W. Diers, P. Lundeen, G.M. Tabor, and S.R. Cianzio. 2010. Genetic analysis of new sources of soybean resistance to brown stem rot. Crop Sci. 50:2431–2439. doi:10.2135/cropsci2010.03.0159

Rincker, K., A.E. Lipka, and B.W. Diers. 2016. Genome-wide association study of brown stem rot resistance in soybean across multiple populations. Plant Gen. 9(2). doi: 10.3835/plantgenome2015.08.0064

SAS Institute. 2011. The SAS system for Microsof Windows. Release 9.3 SAS Inst. Inc., Cary, NC.

Schmutz, J., S. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D. Hyten, Q.J. Song, J.J. Thelen, J. Cheng, D. Xu, U. Hellsten, G. May, Y. Yu, T. Sakurai, T. Umezawa, M. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P.B. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, and S.A. Jackson. 2010. A genome sequence of paleopolyploid soybean (*Glycine max*). Nature 463:178–183. doi:10.1038/nature08670

Sebastian, S.A., C.D. Nickell, and L.E. Gray. 1985. Efficient selection for brown stem rot resistance in soybeans under greenhouse screening conditions. Crop Sci. 25:753–757. doi:10.2135/cropsci1985.0011183X002500050008x

Song, Q., J. Gaofeng, Z. Youlin, D. Grant, R.T. Nelson, E.Y. Hwang, D.L. Hyten, and P.B. Cregan. 2010. Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. Crop Sci. 50:1950–1960. doi:10.2135/cropsci2009.10.0607

Song, Q.J., L.F. Marek, R.C. Shoemaker, K.G. Lark, V.C. Concibido, X. Delannay, J.E. Specht, and P.B. Cregan. 2004. A new integrated genetic linkage map of the soybean. Theor. Appl. Genet. 109:122–128. doi:10.1007/s00122-004-1602-3

Tachibana, H., and L.C. Card. 1972. Brown stem rot resistance and its modification by soybean mosaic virus in soybeans. Phytopathology 62:1314–1317. doi:10.1094/Phyto-62-1314

Tachibana, H., L.C. Card, J.B. Bahrenfus, and W.R. Fehr. 1980. Registration of BSR301 soybean. Crop Sci. 20:414–415. doi:10.2135/cropsci1980.0011183X002000030042x

Tachibana, H., B.K. Voss, and W.R. Fehr. 1987. Registration of BSR101 soybean. Crop Sci. 27:612. doi:10.2135/cropsci1987.0011183X002700030045x

Walker, A.K., A.F. Schmitthenner, R.J. Fioritto, S.K. St. Martin, R.L. Cooper, and R.J. Martin. 1986. Registration of 'Century 84' soybean. Crop Sci. 26:199–200.

Wang, D., J. Shi, S.R. Carlson, P.B. Cregan, R.W. Ward, and B.W. Diers. 2003. A low-cost, high-throughput polyacrylamide gel electrophoresis system for genotyping with microsatellite DNA markers. Crop Sci. 43:1828–1832. doi:10.2135/cropsci2003.1828

Willmot, D.B., and C.D. Nickell. 1989. Genetic analysis of brown stem rot resistance in soybean. Crop Sci. 29:672–674. doi:10.2135/cropsci1989.0011183X002900030026x

Wrather, A., G. Shannon, R. Balardin, L. Carregal, R. Escobar, G.K. Gupta, Z. Ma, W. Morel, D. Ploper, and A. Tenuta. 2010. Effect of diseases on soybean yield in the top eight producing countries in 2006. Plant Health Prog. doi:10.1094/PHP-2010-0125-01-RS

TABLES AND FIGURES

Table 2.1. Summary of populations developed to fine map brown stem rot resistance of soybean.

Mapping population	Number of plants screened for recombination events within interval containing resistance gene on chr 16	Number of recombinant lines developed
Bell × Colfax	1218 F_2	108 $F_{2:3}$
Century 84 × PI 84946-2	738 F_3	26 $F_{3:4}$
Century 84 × PI 437833	699 F_3	21 $F_{3:4}$
Century 84 × PI 437970	729 F_3	15 $F_{3:4}$
PI 86150 × Century 84	654 F_3	23 $F_{3:4}$
L84-5873 × Century 84	840 F_3	13 $F_{3:4}$

Table 2.2. Fine mapping brown stem rot (BSR) resistance in the population Bell x Colfax. The genetic markers, base pair positions, names of plants selected for having recombination within interval and progeny tested, and genotypes of selected recombinant plants in the interval on chromsome16 (linkage group J).

Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants									
				LDXGE10262-1-19	LDXGE10262-1-22	LDXGE10262-2-8	LDXGE10262-2-9	LDXGE10262-2-7	LDXGE10262-1-32	LDXGE10262-1-30	LDXGE10262-2-1	LDXGE10262-2-2	LDXGE10262-2-6
BARCSOYSSR_16_703	Satt529	23,096,020	23,417,387	H†	S				S	H			
BARCSOYSSR_16_916		29,211,224	29,578,150	H	S				S	H			
BARCSOYSSR_16_952		29,872,639	30,210,186	H	S				S	H			
BARCSOYSSR_16_992		30,461,250	30,829,374	H	S				S	H			
				↓‡									
BARCSOYSSR_16_1047		31,312,198	31,680,675	S	S				S	H			
BARCSOYSSR_16_1092		32,167,801	32,657,348	S	S	R	H	R	S	H	H	R	S
BARCSOYSSR_16_1100		32,236,743	32,734,845		S	R	H	R			H	R	S
BARCSOYSSR_16_1102		32,257,365	32,755,476		S								
					↓								
BARCSOYSSR_16_1103			32,770,878		H								
BARCSOYSSR_16_1105		32,285,603	32,783,714		H	R							
						↓							
BARCSOYSSR_16_1107		32,321,995	32,820,106		H	H							
BARCSOYSSR_16_1108		32,345,742	32,843,853		H	H							
BARCSOYSSR_16_1112		32,393,343	32,895,021		H								
BARCSOYSSR_16_1115		32,496,655	32,996,795		H		H	R	S	H		R	
							↑						
BARCSOYSSR_16_1118		32,531,504	33,034,917		H	H	R	R	S	H	H	R	S
BARCSOYSSR_16_1120		32,538,098	33,041,511		H	H	R	R	S	H	H	R	S
BARCSOYSSR_16_1122		32,547,435	33,056,923		H	H	R	R	S	H	H	R	S
								↑					
BARCSOYSSR_16_1124		32,569,730	33,072,966		H			H	S	H			
BARCSOYSSR_16_1128		32,868,956	33,344,076	S	H	H	R	H	S	H	H	R	S

Table 2.2. (cont.)

Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants									
				LDXGE10262-1-19	LDXGE10262-1-22	LDXGE10262-2-8	LDXGE10262-2-9	LDXGE10262-2-7	LDXGE10262-1-32	LDXGE10262-1-30	LDXGE10262-2-1	LDXGE10262-2-2	LDXGE10262-2-6
BARCSOYSSR_16_1129		32,869,871	33,344,991		H	H	R	H	S	H	H	R	S
BARCSOYSSR_16_1132		32,918,026	33,393,145		H	H	R	H	S	H	H	R	S
									↑				
BARCSOYSSR_16_1134		32,953,687	33,428,808		H	H	R	H	H	H	H	R	S
BARCSOYSSR_16_1140		33,105,973	§		H		R				H	R	
										↑	↑	↑	↑
BARCSOYSSR_16_1142		33,209,430	33,700,198		H	H	R	H	H	S	S	H	H
BARCSOYSSR_16_1151	Satt244	33,327,246	33,818,964	S	H		R	H	H	S	S	H	H
BARCSOYSSR_16_1165	Satt547	33,538,126		S	H		R	H	H	S	S	H	H
BSR test ¶				1	1 & 3	11	15	7	2 & 3	2	7	4	4
Foliar symptoms P > F #				0.86	<0.0001††	0.01	<0.0001	0.11	0.37††	<0.0001	0.0006	0.77	0.47
Stem symptoms P > F #				0.19	0.0003††	<0.0001	<0.0001	0.82	0.86††	<0.0001	0.30	0.79	0.64

† S designates the recombinant plant that produced progeny tested was homozygous for the marker allele originating from the susceptible parent Colfax, R designates the plant was homozygous for the marker allele from the resistant parent Bell, and H designates that the plant was heterozygous.

‡ Arrow indicates the direction of resistance locus based on the result of testing progeny from this plant.

§ Marker BARCSOYSSR_16_1140 is not located in Glyma 2.0 Assembly.

¶ BSR resistance assay that the progeny of the selected plants were evaluated.

Significance level of the maker association test.

†† Only the F-test results from Test 1 are reported in line 1-22 and results from Test 2 are reported in line 1-32. F-test results from Test 3 are included in Table 2.1.

Table 2.3. Single marker F-tests of selected progeny tests

Mapping population and selected recombinant plant	Brown stem rot (BSR) test ‡	Number of plants tested §	Marker used in F-test ¶	Foliar symptoms †					Stem symptoms †				
				R	H	S	P > F #	R^2 ††	R	H	S	P > F #	R^2 ††
Bell × Colfax													
10262-1-19	1	33	B_16_1042	0.47	0.47	0.43	0.86	0.01	0.34	0.35	0.23	0.19	0.10
10262-1-22	1	39	B_16_1128	0.18	0.17	0.55	<0.0001	0.48	0.04	0.07	0.34	0.0003	0.36
	3	29	Satt547	0.00	0.12	0.42	0.0005	0.48	0.01	0.12	0.47	<0.0001	0.36
10262-2-8	11	45	B_16_1120	0.41	0.36	0.52	0.01	0.19	0.14	0.19	0.51	<0.0001	0.64
10262-2-9	15	45	B_16_1100	0.18	0.25	0.62	<0.0001	0.59	0.05	0.17	0.65	<0.0001	0.53
10262-2-7	7	44	B_16_1134	0.37	0.34	0.23	0.11	0.10	0.14	0.13	0.10	0.83	0.01
10262-1-32	2	38	Satt547	0.60	0.50	0.56	0.37	0.06	0.30	0.32	0.31	0.86	0.01
	3	40	B_16_1134	0.49	0.51	0.43	0.63	0.03	0.05	0.09	0.09	0.75	0.02
10262-1-30	2	40	B_16_1092	0.20	0.06	0.39	<0.0001	0.43	0.04	0.03	0.33	<0.0001	0.60
10262-2-1	6	45	B_16_1134	0.30	0.36	0.63	0.0006	0.30	0.13	0.19	0.21	0.30	0.01
10262-2-2	4	43	Satt547	0.32	0.27	0.24	0.77	0.01	0.12	0.08	0.07	0.79	0.01
10262-2-6	4	38	Satt547	0.55	0.60	0.50	0.47	0.04	0.49	0.45	0.42	0.64	0.03
Century 84 × PI 84946-2													
278-1-18-6	12	39	B_16_1083	0.66	0.61	0.86	0.12	0.12	0.52	0.43	0.66	0.30	0.70
278-1-18-7	12	35	B_16_1083	0.82	0.64	0.73	0.15	0.11	0.67	0.52	0.53	0.44	0.05
278-1-20-6	12	46	B_16_1100	0.40	0.41	0.82	<0.0001	0.46	0.12	0.18	0.46	0.001	0.28
278-1-18-8	12	43	B_16_1100	0.49	0.46	0.78	<0.0001	0.48	0.04	0.11	0.48	<0.0001	0.45
278-1-17-3	10	42	B_16_1100	0.50	0.25	0.74	0.0006	0.44	0.12	0.18	0.44	0.002	0.26
278-1-17-5	12	34	B_16_1145	0.69	0.79	0.81	0.45	0.05	0.43	0.57	0.66	0.09	0.14
278-1-9-4	4	37	B_16_1100	0.11	0.24	0.63	<0.0001	0.46	0.10	0.36	0.66	<0.0001	0.49
278-1-5-4	4 & 5	29	Satt547	0.74	0.76	0.98	0.50	0.19	0.21	0.65	0.57	0.07	0.38
278-1-5-20	5	41	B_16_1100	0.37	0.34	0.82	0.0003	0.35	0.11	0.13	0.38	0.005	0.23
278-1-11-1	15	45	Satt431	0.77	0.79	0.81	0.79	0.01	0.53	0.52	0.57	0.67	0.02

Table 2.3. (cont.)

Century 84 × PI 437833													
263-5-21-1	13	41	B_16_1120	0.34	0.29	0.43	0.09	0.12	0.00	0.05	0.18	0.001	0.29
263-5-16-1	13	29	B_16_1120	0.18	0.22	0.32	0.13	0.15	0.00	0.01	0.00	‡‡	
263-5-17-1	13	37	B_16_1100	0.38	0.25	0.65	<0.0001	0.41	0.07	0.05	0.58	<0.0001	0.51
263-5-14-7	8	45	B_16_1134	0.07	0.11	0.13	0.34	0.05	0.00	0.00	0.00	‡‡	
263-5-13-1	13	44	B_16_1144	0.13	0.16	0.16	0.79	0.01	0.00	0.00	0.00	‡‡	
263-5-4-5	8	45	B_16_1134	0.26	0.35	0.59	0.0003	0.32	0.02	0.04	0.17	0.009	0.20
263-5-21-3	13	35	B_16_1100	0.22	0.28	0.61	<0.0001	0.46	0.00	0.00	0.52	<0.0001	
Century 84 × PI 437970													
264-6-11-6	9	44	B_16_1100	0.38	0.37	0.79	<0.0001	0.58	0.00	0.04	0.62	<0.0001	0.86
264-6-4-8	12	30	B_16_1120	0.24	0.17	0.86	<0.0001	0.82	0.03	0.02	0.69	<0.0001	0.87
264-6-11-2	9	45	B_16_1100	0.41	0.37	0.42	0.43	0.04	0.08	0.06	0.06	0.95	0.003
264-6-5-3	12	45	B_16_1100	0.43	0.49	0.89	<0.0001	0.62	0.00	0.11	0.51	<0.0001	0.72
264-6-6-3	13	45	B_16_1120	0.26	0.24	0.28	0.54	0.03	0.00	0.00	0.00	‡‡	
264-6-6-4	13	45	B_16_1120	0.53	0.45	0.45	0.39	0.04	0.46	0.28	0.39	0.23	0.07
264-6-5-4	11	45	B_16_1100	0.22	0.26	0.60	<0.0001	0.44	0.00	0.05	0.47	<0.0001	0.55
264-6-6-2	13	43	B_16_1100	0.30	0.28	0.57	<0.0001	0.45	0.02	0.00	0.47	<0.0001	0.58
L84-5873 × Century 84													
261-1-9-9	11	45	B_16_1100	0.25	0.30	0.96	<0.0001	0.89	0.19	0.26	0.93	<0.0001	0.56
261-1-6-10	11	45	B_16_1100	0.31	0.31	0.34	0.84	0.01	0.02	0.02	0.06	0.71	0.02
261-1-6-9	11	44	B_16_1100	0.37	0.41	0.42	0.51	0.03	0.18	0.13	0.20	0.75	0.01
261-1-4-5	13	39	B_16_1118	0.21	0.25	0.27	0.59	0.03	0.00	0.00	0.00	‡‡	
261-1-11-9	15	45	B_16_1100	0.30	0.38	0.80	<0.0001	0.75	0.23	0.33	0.62	<0.0001	0.63
261-1-4-7	13	41	B_16_1118	0.36	0.31	0.46	0.06	0.14	0.00	0.01	0.30	0.0003	0.35

Table 2.3. (cont.)

PI 86150 × Century 84													
262-1-18-2	6	45	B_16_1053	0.12	0.19	0.13	0.43	0.04	0.07	0.02	0.01	0.11	0.10
262-1-23-2	14	43	B_16_1100	0.39	0.47	0.65	0.004	0.24	0.27	0.34	0.46	0.09	0.11
262-1-17-11	14	45	B_16_1115	0.40	0.42	0.68	0.0005	0.31	0.00	0.08	0.40	<0.0001	0.39
262-1-2-11	14	45	B_16_1100	0.05	0.23	0.39	0.005	0.23	0.00	0.16	0.36	0.01	0.19
262-1-22-1	14	40	B_16_1120	0.59	0.56	0.53	0.75	0.02	0.47	0.49	0.46	0.93	0.004
262-1-18-11	14	45	B_16_1120	0.26	0.28	0.28	0.73	0.02	0.09	0.05	0.04	0.57	0.03
262-1-18-13	14	44	B_16_1120	0.72	0.56	0.66	0.06	0.13	0.47	0.54	0.61	0.26	0.06
262-1-17-13	14	43	B_16_1115	0.48	0.45	0.59	0.02	0.17	0.00	0.07	0.25	0.002	0.21
262-1-17-12	14	44	B_16_1120	0.34	0.45	0.48	0.15	0.09	0.09	0.20	0.43	0.008	0.21

† Mean foliar or stem symptoms for the plants in each line that were predicted based on the genetic markers to be homozygous for the allele from Bell, PI 84946-2, PI 437833, PI 437970, L84-5873, or PI 86150, the resistant ('R') parent, heterozygous (H), or homozygous for the allele from Colfax or Century 84, the susceptible parent (S) in the interval containing the resistance QTL.

‡ BSR resistance assay that the progeny of the selected plants were evaluated in as listed in Table 2.2,2.4,2.5,2.6,2.7, and 2.8.

§ Number of progeny plants tested in the BSR resistance assay.

¶ BARCSOYSSR markers on chromosome 16 developed by Song et al. (2010) are abbreviated with the prefix B_16_.

Significance level of the marker association test.

†† R^2 value of the marker association.

‡‡ Insufficient or no stem symptoms were present in the line of this test.

Table 2.4. Fine mapping brown stem rot (BSR) resistance in the population Century 84 x PI84946-2. The genetic markers, base pair positions, names of plants selected for having recombination within the interval and progeny testing, and genotypes of lines in the interval on chromsome16 (linkage group J).

Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants									
				LDX10-278-1-18-6	LDX10-278-1-18-7	LDX10-278-1-20-6	LDX10-278-1-18-8	LDX10-278-1-17-3	LDX10-278-1-17-5	LDX10-278-1-9-4	LDX10-278-1-5-4	LDX10-278-1-5-20	LDX10-278-1-11-1
BARCSOYSSR_16_885	Satt215	28,589,359	28,944,563								S†		
BARCSOYSSR_16_1083		31,941,009	32,430,527	H	H	R	H		S	H			
BARCSOYSSR_16_1084		31,960,330	32,449,848	H									
BARCSOYSSR_16_1087		31,990,527	32,480,045	H									
				↓‡									
BARCSOYSSR_16_1089		32,142,827	32,632,374	S									
BARCSOYSSR_16_1090		32,144,613	32,634,160	S		R							S
BARCSOYSSR_16_1091		32,158,513	32,648,060	S		R							S
BARCSOYSSR_16_1093		32,207,533	32,705,634	S	H	R							S
BARCSOYSSR_16_1094		32,214,241				R		H	S	H			
BARCSOYSSR_16_1095		32,214,483	32,712,584	S	H	R							S
BARCSOYSSR_16_1096		32,219,567	32,717,668		H	R							
					↓								
BARCSOYSSR_16_1097		32,225,921	32,724,023		S	R							
BARCSOYSSR_16_1098		32,227,749	32,725,851			R							
						↓							
BARCSOYSSR_16_1100		32,236,743	32,734,845	S	S	H	H	H	S	H	S	H	S
BARCSOYSSR_16_1115		32,496,655	32,996,795	S	S	H		H		H	S	H	S
BARCSOYSSR_16_1118		32,531,504	33,034,917	S	S	H	H		S	H			
BARCSOYSSR_16_1120		32,538,098	33,041,511				H	H					
BARCSOYSSR_16_1122		32,553,509	33,056,923				H	H		H	S	H	
							↑						
BARCSOYSSR_16_1123		32,566,748	33,069,984				S	H					
								↑					
BARCSOYSSR_16_1127		32,846,820	33,321,940				S	R					
BARCSOYSSR_16_1128		32,868,956	33,344,076	S	S	H	S	R		H			

Table 2.4. (cont.)

Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants									
				LDX10-278-1-18-6	LDX10-278-1-18-7	LDX10-278-1-20-6	LDX10-278-1-18-8	LDX10-278-1-17-3	LDX10-278-1-17-5	LDX10-278-1-9-4	LDX10-278-1-5-4	LDX10-278-1-5-20	LDX10-278-1-11-1
BARCSOYSSR_16_1129		32,869,871	33,344,991				S	R		H	S	H	
BARCSOYSSR_16_1130		32,874,820	33,349,938	S	S	H	S		S	H			
BARCSOYSSR_16_1132		32,918,026	33,393,145						S				
									↑				
BARCSOYSSR_16_1135		32,965,066	33,440,187						H				
										↑			
BARCSOYSSR_16_1136		32,970,250	33,445,371					R	H	R			
BARCSOYSSR_16_1138		33,047,646	33,555,041						H				
BARCSOYSSR_16_1139		33,078,055	33,524,923	S	S	H	S	R	H	R	S	H	S
BARCSOYSSR_16_1140		33,105,973	§					R	H	R	S		
BARCSOYSSR_16_1141		33,117,153	33,596,663						H				
BARCSOYSSR_16_1142		33,209,430	33,700,198					R			S		
											↑		
BARCSOYSSR_16_1145		33,239,017	33,729,785	S	S	H	S	R	H	R	H	H	
BARCSOYSSR_16_1149		33,281,370	33,772,138					R		R			
BARCSOYSSR_16_1151	Satt244	33,327,246	33,818,964	S	S	H	S	R	H	R			
												↑	
BARCSOYSSR_16_1152		33,360,926	33,852,644					R		R	H	S	
BARCSOYSSR_16_1158		33,470,670	33,967,689					R					↑
BARCSOYSSR_16_1165	Satt547	33,538,126	34,035,215					R		R	H	S	H
BARCSOYSSR_16_1234	Satt431	35,718,476	36,221,234					R		R	H	S	H
BSR test ¶				12	12	12	12	10	12	4	4&5	5	15
Foliar symptoms P > F #				0.12	0.15	<0.0001	<0.0001	0.0006	0.45	<0.0001	0.50	0.0003	0.79
Stem symptoms P > F #				0.30	0.44	0.001	<0.0001	0.002	0.09	<0.0001	0.07	0.005	0.67

Table 2.4. (cont.)

† S designates the recombinant plant that produced progeny tested was homozygous for the marker allele originating from the susceptible parent Century 84, R designates the plant was homozygous for the marker allele from the resistant parent PI 84946-2, and H designates that the plant was heterozygous.

‡ Arrow indicates the direction of resistance locus based on the result of testing progeny from this plant.

§ Marker BARCSOYSSR_16_1140 is not located in Glyma 2.0 Assembly.

¶ BSR resistance assay that the progeny of the selected plants were evaluated.

Significance level of the maker association test.

Table 2.5. Fine mapping brown stem rot (BSR) resistance in the population Century 84 x PI437833. The genetic markers, base pair positions, names of plants selected for having recombination within the interval and progeny testing, and genotypes of lines in the interval on chromsome16 (linkage group J).

			Selected Plants						
Simple Sequence Repeat marker	Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	LDX10-263-5-21-1	LDX10-263-5-16-1	LDX10-263-5-17-1	LDX10-263-5-14-7	LDX10-263-5-13-1	LDX10-263-5-4-5	LDX10-263-5-21-3
BARCSOYSSR_16_1092	32,167,801	32,657,348	S†	R	H	R	R	H	H
BARCSOYSSR_16_1100	32,236,743	32,734,845	S	R	H	R	R	H	H
BARCSOYSSR_16_1103	32,272,767	32,770,878	S	R	H		R		
BARCSOYSSR_16_1105	32,285,603	32,783,714	S	R	H	R	R	H	
			↓‡						
BARCSOYSSR_16_1107	32,321,995	32,820,106	H	R	H		R		
BARCSOYSSR_16_1109	32,350,674	32,848,785		R	H		R		
BARCSOYSSR_16_1110	32,357,031	32,855,147		R	H				
BARCSOYSSR_16_1111	32,367,372	32,865,488		R	H				
BARCSOYSSR_16_1112	32,393,343	32,895,021		R					
				§					
BARCSOYSSR_16_1113	32,442,578	32,942,718		H	H		R		
BARCSOYSSR_16_1114	32,456,480	32,956,620			H				
					↑				
BARCSOYSSR_16_1115	32,496,655	32,996,795	H	H	S				
BARCSOYSSR_16_1118	32,531,504	33,034,917	H	H	S		R		
BARCSOYSSR_16_1120	32,538,098	33,041,511	H	H	S	R	R	H	
BARCSOYSSR_16_1122	32,553,509	33,056,923	H	H	S		R		
BARCSOYSSR_16_1124	32,569,730	33,072,966					R		
BARCSOYSSR_16_1131	32,878,276	33,353,394					R		
BARCSOYSSR_16_1132	32,918,026	33,393,145				R	R	H	
BARCSOYSSR_16_1133	32,918,653	33,393,774					R		
						↑	↑		
BARCSOYSSR_16_1134	32,953,687	33,428,808	H	H	S	H	H	H	H

Table 2.5. (cont.)

Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants						
				LDX10-263-5-21-1	LDX10-263-5-16-1	LDX10-263-5-17-1	LDX10-263-5-14-7	LDX10-263-5-13-1	LDX10-263-5-4-5	LDX10-263-5-21-3
BARCSOYSSR_16_1137		33,003,946	33,479,398				H			
BARCSOYSSR_16_1140		33,105,973	¶				H		H	
BARCSOYSSR_16_1144		33,237,426	33,728,194				H		H	
BARCSOYSSR_16_1151	Satt244	33,327,246	33,818,964						H	H
BARCSOYSSR_16_1154		33,406,683	33,898,402							H
									↑	↑
BARCSOYSSR_16_1158		33,470,670	33,967,689				H		S	R
BSR test #				13	13	13	8	13	8	13
Foliar symptoms P > F ††				0.09	0.13	<0.0001	0.34	0.79	0.0003	<0.0001
Stem symptoms P > F ††				0.001	§	<0.0001	‡‡	‡‡	0.009	<0.0001

† S designates the recombinant plant that produced progeny tested was homozygous for the marker allele originating from the susceptible parent Century 84, R designates the plant was homozygous for the marker allele from the resistant parent PI 437833, and H designates that the plant was heterozygous.

‡ Arrow indicates the direction of resistance locus based on the result of testing progeny from this plant.

§ Inconclusive test results were obtained for line 5-16-1 and insufficient stem symptoms were present to conduct an F-Test.

¶ Marker BARCSOYSSR_16_1140 is not located in Glyma 2.0 Assembly.

BSR resistance assay that the progeny of the selected plants were evaluated.

†† Significance level of the maker association test.

‡‡ No stem symptoms were present in this line.

Table 2.6. Fine mapping brown stem rot (BSR) resistance in the population Century 84 x PI437970. The genetic markers, base pair positions, names of plants selected for having recombination within the interval and progeny testing, and genotypes of lines in the interval on chromsome16 (linkage group J).

Simple Sequence Repeat marker	Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants							
			LDX10-264-6-11-6	LDX10-264-6-4-8	LDX10-264-6-11-2	LDX10-264-6-5-3	LDX10-264-6-6-3	LDX10-264-6-6-4	LDX10-264-6-5-4	LDX10-264-6-6-2
BARCSOYSSR_16_1053	31,416,354	31,784,828	S†		H					
BARCSOYSSR_16_1063	31,673,436	32,041,909	S		H					
BARCSOYSSR_16_1083	31,941,009	32,430,527		S		H			H	
BARCSOYSSR_16_1092	32,167,801	32,657,348	S	S	H	H	R	S	H	H
			↓‡							
BARCSOYSSR_16_1095	32,214,483	32,712,584	H							
BARCSOYSSR_16_1097	32,225,921	32,724,023		S						
BARCSOYSSR_16_1100	32,236,743	32,734,845	H	S	H	H	R	S	H	H
BARCSOYSSR_16_1101	32,250,359	32,748,462		S						
				↓						
BARCSOYSSR_16_1103	32,272,767	32,770,878		H		H	R	S		H
BARCSOYSSR_16_1105	32,285,603	32,783,714	H	H	H	H	R	S		H
					↓					
BARCSOYSSR_16_1107	32,321,995	32,820,106			R		R	S		H
BARCSOYSSR_16_1108	32,345,742	32,843,853			R		R	S		H
BARCSOYSSR_16_1109	32,350,674	32,848,785			R					
BARCSOYSSR_16_1110	32,357,031	32,855,147						S		
BARCSOYSSR_16_1111	32,367,372	32,865,488						S		
BARCSOYSSR_16_1112	32,393,343	32,895,021			R	H		S		
BARCSOYSSR_16_1113	32,442,578	32,942,718					R	S		
BARCSOYSSR_16_1114	32,456,480	32,956,620				H	R	S		H
						↑				
BARCSOYSSR_16_1115	32,496,655	32,996,795		H	R	S	R	S	H	H
BARCSOYSSR_16_1116	32,524,050	§					R	S		

Table 2.6. (cont.)

Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants							
				LDX10-264-6-11-6	LDX10-264-6-4-8	LDX10-264-6-11-2	LDX10-264-6-5-3	LDX10-264-6-6-3	LDX10-264-6-6-4	LDX10-264-6-5-4	LDX10-264-6-6-2
BARCSOYSSR_16_1117		32,524,918	§					R	S		
								↑	↑		
BARCSOYSSR_16_1118		32,531,504	33,034,917		H	R	S	H	H	H	H
BARCSOYSSR_16_1120		32,538,098	33,041,511	H	H	R	S	H	H		H
BARCSOYSSR_16_1122		32,553,509	33,056,923	H	H	R	S	H	H		H
BARCSOYSSR_16_1124		32,569,730	33,072,966								H
BARCSOYSSR_16_1126		32,845,988	33,321,108								H
BARCSOYSSR_16_1128		32,868,956	33,344,076		H		S			H	H
BARCSOYSSR_16_1129		32,869,871	33,344,991								H
BARCSOYSSR_16_1130		32,874,820	33,349,938		H		S			H	
BARCSOYSSR_16_1131		32,878,276	33,353,394							H	H
BARCSOYSSR_16_1132		32,918,026	33,393,145	H		R				H	
BARCSOYSSR_16_1133		32,918,653	33,393,774							H	H
										↑	↑
BARCSOYSSR_16_1134		32,953,687	33,428,808	H	H	R	S		H	R	R
BARCSOYSSR_16_1135		32,965,066	33,440,187								R
BARCSOYSSR_16_1136		32,970,250	33,445,371								R
BARCSOYSSR_16_1137		33,003,946	33,479,398			R					
BARCSOYSSR_16_1140		33,105,973	§	H	H	R	S			R	
BARCSOYSSR_16_1144		33,237,426	33,728,194	H	H	R	S	H	H	R	R
BARCSOYSSR_16_1151	Satt244	33,327,246	33,818,964	H	H	R	S			R	
BSR test ¶				9	12	9	12	13	13	11	13
Foliar symptoms P > F #				<0.0001	<0.0001	0.43	<0.0001	0.54	0.39	<0.0001	<0.0001
Stem symptoms P > F #				<0.0001	<0.0001	0.95	<0.0001	††	0.23	<0.0001	<0.0001

† S designates the recombinant plant that produced progeny tested was homozygous for the marker allele originating from the susceptible parent Century 84, R designates the plant was homozygous for the marker allele from the resistant parent PI 437970, and H designates that the plant was heterozygous.

‡ Arrow indicates the direction of resistance locus based on the result of testing progeny from this plant.

Table 2.6. (cont.)

§ Markers BARCSOYSSR_16_1116, BARCSOYSSR_16_1117, and BARCSOYSSR_16_1140 are not located in Glyma 2.0 Assembly.

¶ BSR resistance assay that the progeny of the selected plants were evaluated.

Significance level of the maker association test.

†† No stem symptoms were present in this line.

Table 2.7. Fine mapping brown stem rot (BSR) resistance in the population L84-5873 x Century 84. The genetic markers, base pair positions, names of plants selected for having recombination within the interval and progeny testing, and genotypes of lines in the interval on chromsome16 (linkage group J).

			Selected Plants					
Simple Sequence Repeat marker	Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	LDX10261-1-9-9	LDX10261-1-6-10	LDX10261-1-6-9	LDX10261-1-4-5	LDX10261-1-11-9	LDX10261-1-4-7
BARCSOYSSR_16_1083	31,941,009	32,430,527	S†	H	H	R		H
			↓‡					
BARCSOYSSR_16_1090	32,144,613	32,634,160	H					
BARCSOYSSR_16_1091	32,158,513	32,648,060	H					
BARCSOYSSR_16_1092	32,167,801	32,657,348	H	H	H		H	H
BARCSOYSSR_16_1100	32,236,743	32,734,845	H	H	H	R	H	H
BARCSOYSSR_16_1102	32,257,365	32,755,476		H	H		H	H
BARCSOYSSR_16_1103	32,272,767	32,770,878		H	H	R	H	H
BARCSOYSSR_16_1105	32,285,603	32,783,714		H	H		H	H
				↓				
BARCSOYSSR_16_1107	32,321,995	32,820,106		R	H	R	H	H
BARCSOYSSR_16_1108	32,345,742	32,843,853		R	H	R	H	H
BARCSOYSSR_16_1112	32,393,343	32,895,021				R		
BARCSOYSSR_16_1113	32,442,578	32,942,718			H	R		
					↓			
BARCSOYSSR_16_1114	32,456,480	32,956,620		R	R	R	H	
						↑	↑	
BARCSOYSSR_16_1115	32,496,655	32,996,795	H	R	R	H	R	H
BARCSOYSSR_16_1118	32,531,504	33,034,917	H	R	R	H	R	H
BARCSOYSSR_16_1119	32,536,146	33,039,559						H
BARCSOYSSR_16_1120	32,538,098	33,041,511		R	R	H	R	H
								↑
BARCSOYSSR_16_1121	32,547,435	33,050,848						R

Table 2.7. (cont.)

Simple Sequence Repeat marker	Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants					
			LDX10261-1-9-9	LDX10261-1-6-10	LDX10261-1-6-9	LDX10261-1-4-5	LDX10261-1-11-9	LDX10261-1-4-7
BARCSOYSSR_16_1122	32,553,509	33,056,923		R	R		R	R
BARCSOYSSR_16_1128	32,868,956	33,344,076	H	R	R	H		R
BARCSOYSSR_16_1130	32,874,820	33,349,938	H	R	R	H		R
BARCSOYSSR_16_1134	32,953,687	33,428,808	H	R	R	H	R	R
BARCSOYSSR_16_1144	33,237,426	33,728,194	H	R	R	H	R	R
BSR test ¶			11	11	11	13	15	13
Foliar symptoms P > F #			<0.0001	0.84	0.51	0.59	<0.0001	0.06
Stem symptoms P > F #			<0.0001	0.71	0.75	§	<0.0001	0.0003

† S designates the recombinant plant that produced progeny tested was homozygous for the marker allele originating from the susceptible parent Century 84, R designates the plant was homozygous for the marker allele from the resistant parent L84-5873, and H designates that the plant was heterozygous.

‡ Arrow indicates the direction of resistance locus based on the result of testing progeny from this plant.

¶ BSR resistance assay that the progeny of the selected plants were evaluated.

Significance level of the maker association test.

§ No stem symptoms were present in this line.

Table 2.8. Fine mapping brown stem rot (BSR) resistance in the population PI86150 x Century 84. The genetic markers, base pair positions, names of plants selected for having recombination within the interval and progeny testing, and genotypes of lines in the interval on chromsome16 (linkage group J).

				Selected Plants								
Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	LDX10-262-1-18-2	LDX10-262-1-23-2	LDX10-262-1-17-11	LDX10-262-1-2-11	LDX10-262-1-22-1	LDX10-262-1-18-11	LDX10-262-1-18-13	LDX10-262-1-17-13	LDX10-262-1-17-12
BARCSOYSSR_16_0840	Satt622	27,633,674	27,981,082	H†								
BARCSOYSSR_16_1053		31,416,354	31,784,828	H								
				↓‡								
BARCSOYSSR_16_1083		31,941,009	32,430,527	R	S			S				
BARCSOYSSR_16_1092		32,167,801	32,657,348	R	S	R	H	S	R	S	H	H
					↓							
BARCSOYSSR_16_1096		32,219,567	32,717,668		H							
BARCSOYSSR_16_1099		32,232,363	32,730,465		H							
BARCSOYSSR_16_1100		32,236,743	32,734,845	R	H	R	H	S	R	S	H	H
BARCSOYSSR_16_1102		32,257,365	32,755,476			R	H		R	S	H	
BARCSOYSSR_16_1103		32,272,767	32,770,878			R	H		R	S	H	
BARCSOYSSR_16_1105		32,285,603	32,783,714				H				H	
BARCSOYSSR_16_1107		32,321,995	32,820,106			R			R			
BARCSOYSSR_16_1108		32,345,742	32,843,853		H	R	H	S	R	S	H	
BARCSOYSSR_16_1110		32,357,031	32,855,147							S		
BARCSOYSSR_16_1111		32,367,372	32,865,488							S		
BARCSOYSSR_16_1112		32,393,343	32,895,021			R	H		R	S		
BARCSOYSSR_16_1113		32,442,578	32,942,718			R	H		R	S		
BARCSOYSSR_16_1114		32,456,480	32,956,620			R	H			S		
						↓	↑					
BARCSOYSSR_16_1115		32,496,655	32,996,795		H	H	R	S	R	S	H	H
								↑				
BARCSOYSSR_16_1116		32,524,050	§					H	R	S	H	
BARCSOYSSR_16_1117		32,524,918	§					H	R	S	H	
									↑	↑		

Table 2.8. (cont.)

Simple Sequence Repeat marker		Linkage group (LG)-J (16), position (bp) - Glyma 1.01	Linkage group (LG)-J (16), position (bp) - Glyma 2.0	Selected Plants								
				LDX10-262-1-18-2	LDX10-262-1-23-2	LDX10-262-1-17-11	LDX10-262-1-2-11	LDX10-262-1-22-1	LDX10-262-1-18-11	LDX10-262-1-18-13	LDX10-262-1-17-13	LDX10-262-1-17-12
BARCSOYSSR_16_1118		32,531,504	33,034,917					H	H	H	H	
											↑	
BARCSOYSSR_16_1119		32,536,146	33,039,559					H	H	H	S	
BARCSOYSSR_16_1120		32,538,098	33,041,511	R	H	H	R	H	H	H	S	H
BARCSOYSSR_16_1122		32,553,509	33,056,923				R					
BARCSOYSSR_16_1129		32,869,871	33,344,991		H	H	R	H	H	H	S	H
BARCSOYSSR_16_1130		32,874,820	33,349,938		H			H				H
BARCSOYSSR_16_1131		32,878,276	33,353,394									H
												↑
BARCSOYSSR_16_1132		32,918,026	33,393,145		H			H				R
BARCSOYSSR_16_1133		32,918,653	33,393,774									R
BARCSOYSSR_16_1134		32,953,687	33,428,808	R		H	R		H	H	S	R
BARCSOYSSR_16_1135		32,965,066	33,440,187									R
BARCSOYSSR_16_1136		32,970,250	33,445,371									R
BARCSOYSSR_16_1137		33,003,946	33,479,398									R
BARCSOYSSR_16_1139		33,078,055	33,524,923									
BARCSOYSSR_16_1140		33,105,973	§	R		H	R		H	H		R
BARCSOYSSR_16_1151	Satt244	33,327,246	33,818,964	R	H			H				
BSR test ¶				6	14	14	14	14	14	14	14	14
Foliar symptoms P > F #				0.43	0.004	0.0005	0.005	0.75	0.73	0.06	0.02	0.15
Stem symptoms P > F #				0.11	0.09	<0.0001	0.01	0.93	0.57	0.26	0.002	0.008

† S designates the recombinant plant that produced progeny tested was homozygous for the marker allele originating from the susceptible parent Century 84, R designates the plant was homozygous for the marker allele from the resistant parent PI 86150, and H designates that the plant was heterozygous.

‡ Arrow indicates the direction of resistance locus based on the result of testing progeny from this plant.

§ Markers BARCSOYSSR_16_1116, BARCSOYSSR_16_1117, and BARCSOYSSR_16_1140 are not located in Glyma 2.0 Assembly.

¶ BSR resistance assay that the progeny of the selected plants were evaluated.

Significance level of the maker association test.

CHAPTER THREE

Genome-Wide Association Study of Brown Stem Rot Resistance in Soybean across Multiple Populations [1]

INTRODUCTION

Brown stem rot of soybean, caused by the soilborne fungus *Cadophora gregata*, affects soybean production in the northern United States, Canada, Brazil, and also has a minor impact in China (Wrather et al., 2010). Yield losses of up to 38% have been reported (Gray, 1972), and annual damage to the U.S. soybean crop was estimated to average 422,000 Mg (15.5 million bushels) from 2006 to 2009 (Koenning and Wrather, 2010). Management of this disease is best achieved with host genetic resistance (Bachman et al., 1997b). Multiple screens of the USDA Soybean Germplasm Collection for resistance have identified almost 400 accessions with resistance similar to current resistant cultivars (Bachman et al., 1997a; Bachman and Nickell, 2000a; Chamberlain and Bernard, 1968; Hughes et al., 2004; Nelson et al., 1989; Patzoldt et al., 2003). However, introgression of disease resistance into cultivars has only been achieved with the accessions PI 84946-2 and PI 88788.

Three genes conferring resistance to BSR have been identified through genetic studies: *Rbs1*, *Rbs2*, and *Rbs3* (Hanson et al., 1988; Willmont and Nickell, 1989). These original studies showed the genes were unlinked; however, subsequent genetic mapping placed all three resistance genes onto the same region on chromosome 16 (linkage group J) near the simple-sequence repeat markers Satt215 and Satt431 (Bachman et al., 2001; Lewers et al., 1999; Patzoldt et al., 2005b). Additional BSR resistance quantitative trait loci (QTL) in experimental crosses derived from eight Chinese soybean accessions have been mapped to the same region on chromosome 16 near Satt431, Satt547, or Satt244 (Patzoldt et al., 2005a; Perez et al., 2010). The efficiency of MAS for this locus has been limited because it was mapped to a large region of nearly 10.2 Mb of the Williams 82 genome. Thus, there has been a significant need to study this region and elucidate the key locus or loci underlying this QTL. Rincker et al. (2016) recently fine

[1] Previously published: Rincker, K., A.E. Lipka, B.W. Diers. 2016. Genome-Wide Association Study of Brown Stem Rot Resistance in Soybean across Multiple Populations. Plant Gen. 9(2). doi: 10.3835/plantgenome2015.08.0064. Copyright owner has provided permission to reprint.

mapped *Rbs* genes from multiple sources and found intervals containing *Rbs1*, *Rbs2*, and *Rbs3* to be located between 32.7 and 33.1 Mb on chromosome 16. All intervals identified in this study included the same 0.04 Mb interval, suggesting the possibility that only one resistance locus is present among the sources tested.

Although additional loci have been proposed as interacting with the *Rbs* locus or acting as modifiers (Bachman and Nickell, 2000b; Sebastian et al., 1985; Waller et al., 1991), their identification using linkage mapping approaches has been difficult. Some common drawbacks of linkage mapping include the time required to develop populations, limitations in genomic resolution arising from high linkage disequilibrium (LD), and limitations in statistical power arising from small population sizes, all of which usually lead to low precision of QTL mapping (Cardon and Bell, 2001). The GWAS offers an alternative to linkage mapping for identifying QTL. Using plant accessions that capture many generations of recombination, the GWAS typically has greater resolution for locating QTL relative to linkage mapping (Cardon and Bell, 2001). Furthermore, these accessions typically include more segregating loci than what is found in a biparental cross used in a linkage mapping study.

Recent advances in genotyping and phenotyping technologies are making it possible to dissect the genetic underpinnings of important soybean traits and disease resistance with unprecedented resolution. For example, the genotyping of 19,652 accessions in the USDA Soybean Germplasm Collection with an Illumina Infinium chip, SoySNP50K, containing 52,041 SNPs (Song et al., 2013) is now complete. Phenotypic data on many of these accessions for BSR resistance and other traits are also available through the USDA Germplasm Resources Information Network (GRIN) and from previous accession screening studies (Nelson et al., 1989; Bachman et al., 1997a; Bachman and Nickell, 2000a; Patzoldt et al., 2003).

Genome-wide association studies have been conducted in soybean to identify QTL controlling a wide variety of soybean traits including disease resistance. Hwang et al. (2014) and Vaughn et al. (2014) each used the SoySNP50K array to identify QTL associated with soybean seed protein and oil concentration. In addition, Vaughn et al. (2014) relied on historical phenotypic data available in GRIN. Both studies identified previously reported QTL and refined their genomic

locations. Wen et al. (2014) also used the SoySNP50K array to study sudden death syndrome [*Fusarium virguliforme* (syn. *F. solani* f. sp. *glycines*)] resistance, which resulted in the identification of novel loci and further refinement of the genomic regions already known to contain QTL. In addition, associations with resistance to sclerotinia stem rot of soybean [*Sclerotinia sclerotiorum* (Lib.) de Bary] were identified by Bastien et al. (2014). These studies demonstrate the adaptability of GWAS to soybean traits including disease resistance, the use of GRIN phenotypic data, and the use of the SoySNP50K array as a source of genotypic data. To our knowledge, GWAS of soybean BSR resistance has not been reported.

A more precise location of previously identified resistance genes underlying the QTL on chromosome 16 (Hanson et al., 1988; Willmont and Nickell, 1989) and the elucidation of additional novel loci should increase the efficiency and effectiveness of MAS and aid in the determination of whether this QTL consists of separate genes or a common resistance gene. Genome-wide association studies with diverse soybean accessions and a dense set of SNP markers should contribute to this endeavor by refining QTL to a narrow genomic interval. Therefore, the objectives of this study were to use GWAS to (i) rigorously study the interval containing the known BSR resistance QTL on chromosome 16 and (ii) map novel resistance QTL in a diverse set of accessions previously screened for BSR resistance.

MATERIALS AND METHODS

Germplasm

Phenotypic data for BSR resistance of 4744 soybean accessions were obtained for four diversity panels designated N-1989, B-1997, B-2000, and P-2003. N-1989 included binary data (i.e., resistant vs. susceptible) based on a combination of foliar and stem observations for 2773 accessions ranging from maturity group (MG) 000 through IV and originating from primarily Asia (84%) and Europe (14%) (Nelson et al., 1989). The remaining accessions are from North America, North Africa, Australia, and unknown origins. Eighteen accessions were characterized by the USDA GRIN database as developed cultivars, including eight commercial cultivars used as susceptible checks. Because of a low number of resistant accessions, the four resistant commercial check cultivars that derived their resistance from PI84946-2 were excluded to prevent any bias from introgression of genomic content from a single source of BSR resistance.

From 1981 to 1986, susceptible accessions from N-1989 were culled from nonreplicated field tests conducted in four Midwest states (Iowa, Illinois, Minnesota, and Wisconsin) and putative resistant lines were evaluated in replicated field and greenhouse tests. Phenotypic data were obtained from GRIN (https://npgsweb.ars-grin.gov/gringlobal/search.aspx). B-1997 included quantitative data of separate foliar and stem measurements from 540 accessions originating from central China and ranging in MG II through IV (Bachman et al., 1997a). Evaluations were conducted in 1994 in a nonreplicated field test at Urbana, IL. Phenotypic data were obtained from Michael Bachman (personal communication, 2014). Foliar data were expressed as the percentage of plants exhibiting symptoms, while stem data were measured as the proportion of nodes exhibiting brown pith and averaged over four plants. B-2000 included quantitative foliar measurement data from 825 accessions originating from central and southern China and ranging in MG IV through VIII (Bachman and Nickell, 2000a). Greenhouse data from evaluations with the Oh2 isolate of *C. gregata* were included in this analysis because no selection for BSR resistance was conducted before evaluating the 825 accessions. The phenotypic data for B-2000 were expressed as the percentage of the nodes exhibiting foliar symptoms and averaged over two to five plants tested within each replication. A total of 281 accessions with symptoms failing to progress past the first trifoliate in two plants were included in a second replication. For these accessions, phenotypic data were averaged across the two replicates and obtained from Bachman (1999). P-2003 included quantitative stem data from 606 accessions originating from south–central China and ranging in MG I through IX (Patzoldt et al., 2003). Evaluation of these accessions was conducted in a greenhouse with one replication using the Oh2 *C. gregata* isolate. Phenotypic data were obtained from GRIN and expressed as a percentage of the nodes exhibiting brown pith and averaged over five plants. Methods to evaluate resistance in greenhouse tests among the N-1989, B-2000, and P-2003 panels were similar with the exception that Nelson et al. (1989) inoculated with a different isolate. Separate analyses were conducted on each of the four diversity panels. Moreover, two separate analyses were conducted for stem and foliar BSR data in the B-1997 panel. The Box-Cox (Box and Cox, 1964) procedure was conducted in SAS v9.3 (SAS Institute, 2011) to find the optimal transformations to correct for non-normality of the error terms and unequal variance of BSR resistance in the B-1997, B-2000, and P-2003 panels.

Genome-Wide Association Study

Genotypic data of soybean accessions from the SoySNP50K BeadChip (Song et al., 2013) were obtained from Perry Cregan, USDA–ARS. Single-nucleotide polymorphisms with more than 10% missing data were discarded before the GWAS and accessions were removed if there were more than 10% heterozygous genotypes. After removal of SNPs with minor allele frequency <5%, between 29,815 and 33,486 SNPs (depending on the panel) were used in the GWAS (Table 3.1). Missing SNP genotypes were imputed with the major allele. To conduct GWAS, a unified mixed linear model (MLM; Yu et al., 2006) with population parameters previously determined (Zhang et al., 2010) was implemented in the GAPIT R package (Lipka et al., 2012). To control for population structure and familial relatedness, the unified MLM included principal components (Price et al., 2006) and a kinship matrix computed by the VanRaden method (VanRaden, 2008). For each GWAS scan, the Bayesian information criterion (BIC; Schwarz 1978) was used to determine the optimal number of principal components to include in the GWAS model. The variance component estimates from this model were used to estimate narrow-sense heritability; specifically the genetic variance component estimate was divided by the sum of the genetic and residual variance component estimates. The phenotypic variation explained by the model was assessed with a likelihood-ratio-based R^2 statistic, denoted R^2_{LR} (Sun et al., 2010). The Benjamini and Hochberg (1995) procedure was implemented to the GWAS results of each panel to control for the multiple testing problem at a false discovery rate (FDR) of 10%. This rate was chosen because phenotypic data of the panels were based on limited replication and trends among panels would be identified. To further elucidate the genomic underpinnings of the peak GWAS signals, a multilocus mixed linear model (MLMM) (Segura et al., 2012) in R software was implemented using all SNPs across the genome. The final MLMM was selected using the extended BIC (Chen and Chen, 2008). All SNPs that were identified in the final MLMM were then entered as covariates in a second GWAS.

Because the unified MLM was developed to analyze quantitative traits, some of its statistical assumptions (e.g., normality of error terms) are violated when fitted to the binary BSR data in the N-1989 panel. Thus, it was imperative that the GWAS signals identified using the unified MLM were confirmed using an alternate approach specifically designed to analyze binary data, namely the logistic regression model (Agresti 2013). All statistically significant markers from

the original GWAS scan of the N-1989 panel were considered for inclusion into a logistic regression model using SAS PROC LOGISTIC (SAS Institute, 2011), where the optimal model was determined using the BIC (Schwarz 1978) in a stepwise model selection procedure. Unlike the unified MLM, these logistic regression models did not include individuals as random effects. Phenotypic variation explained by the model was assessed with a likelihood-ratio-based R^2 statistic, denoted R^2_{McF} (McFadden, 1974).

Linkage Disequilibrium Analysis

Linkage disequilibrium was assessed by computing the squared allele frequency correlations (r^2) between unimputed marker pairs using the site-by-all option in TASSEL version 5.0 (Bradbury et al., 2007). Only SNPs with a minor allele frequency >5% and <10% missing were used to estimate LD.

RESULTS

Phenotypic Data

The N-1989 panel contains 2738 susceptible and 35 resistant accessions. The remaining panels measured either foliar or stem symptoms, which were quantified as the proportion of the plant expressing symptoms. For these remaining panels, the proportion of individual plants exhibiting BSR symptoms averaged between 0.33 and 0.39 (Table 3.1). Interestingly, the B-1997 panel had, on average, threefold more plants showing no signs of foliar BSR development than stem BSR development in the B-1997 panel and BSR development in the B-2000 and P-2003 panels (Fig. 3.1). The narrow-sense heritabilities of BSR resistance in the B-1997, B-2000, and P-2003 panels ranged from 0.49 to 0.93 (Table 3.1), suggesting that genetic variability might play a substantial role in BSR resistance.

Genome-Wide Association Study and Stepwise Procedures

Consistent with previous studies, our GWAS detected significant associations between BSR resistance and marker loci on chromosome 16. Using the unified MLM, significant SNPs were

identified on chromosomes 2, 16, and 17 in the N-1989 panel (Table 3.2; Fig. 3.2). On chromosome 16, statistically significant GWAS signals were found in the vicinity of the previously reported BSR QTL (Fig. 3.3). In contrast to the N-1989 panel, no statistically significant associations were identified in either the B-1997 foliar or stem panel (Fig. 3.2). The GWAS of the B-2000 panel found significant peaks on chromosome 16, 5, and 8 (Table 3.2; Fig. 3.4; Fig. 3.2). The most significant SNP (*P*-value 4.79×10^{-35}) was located 21 kb from the interval identified by Rincker et al. (2016) as potentially containing *Rbs1*, *Rbs2*, and *Rbs3* and explained 17% of the variation for symptoms (Table 3.2; Fig. 3.4). In agreement with the N-1989 and B-2000 panel GWAS results, the peak signals from the GWAS of the P-2003 panel were in the vicinity of the putative BSR QTL on chromosome 16 (Table 3.2; Fig. 3.5; Fig. 3.2). Moreover, the marker with the strongest association with BSR resistance (*P*-value 1.59×10^{-7}) was the same as that identified in the GWAS of the B-2000 panel.

Within the N-1989 panel, the MLMM analysis identified no significant associations; however, stepwise logistic regression of significantly associated SNPs identified in the GWAS revealed that three SNPs on three different chromosomes (i.e., chromosomes 2, 16, and 17) best described the binary resistant and susceptible phenotypes (Table 3.3). When the GWAS was reconducted with these three SNPs included as covariates in the model, no statistically significant signals were detected at 10% FDR (Fig. 3.3), suggesting that these three SNPs sufficiently account for BSR resistance variability in this panel. Within the B-1997 panel, no significant associations were found using the MLMM, which is consistent with the original GWAS scan. The optimum model obtained from the MLMM analysis for the B-2000 panel included two SNPs on chromosome 16 (Table 3.3). When these two SNPs were included as covariates in the GWAS model, an additional four SNPs (three of which are proximal to the putative BSR QTL on chromosome 16 and another on chromosome 8) were significant at 10% FDR (Fig. 3.4). This could indicate that more variability exists than can be explained by the two MLMM-identified SNPs. Finally, the optimum model obtained with MLMM for the P-2003 panel identified two SNPs on chromosome 16 (Table 3.3) explaining 10% of the variation. Subsequently, the GWAS rescan that included these two SNPs as covariates in the model identified no statistically significant signals at 10% FDR (Fig. 3.5), which suggests that these two SNPs sufficiently account for BSR resistance variability in this panel.

DISCUSSION AND CONCLUSIONS

Before this study, genetic mapping of BSR resistance had been conducted primarily through the use of biparental populations. Given recent advances in genotyping and the diversity present in soybean accessions, GWAS is now a viable approach for identifying new genetic resistance loci and pinpointing the location of *Rbs* genes on chromosome 16. We conducted GWAS with a 50K SNP array in four panels of diverse soybean germplasm that include all publically available BSR resistance data. While the scope of this study is limited by the pathological screening and inoculation methods, this work constitutes one of the most extensive analyses of natural variation for BSR resistance to date.

Linkage Disequilibrium and Population Structure

The genomic region on chromosome 16 from 32 to 34 Mb that contains significant SNPs in two of the MLMM analyses is gene rich with rapid LD decay. The N-1989 panel is composed of more accessions ranging from wider geographic origins, which we would expect to have less LD. However, the greater LD present in the N-1989 panel than other panels (Fig. 3.3, 3.4, 3.5) suggests this group of accessions could have been subject to greater selection pressure, founding events, or the presence of population structure compared with the germplasm that comprised the other three association panels. Hwang et al. (2014) also used the SoySNP50K array and reported an approximate coverage of one SNP every 17 kb in euchromatic regions and 100 kb in heterochromatic regions. This coverage is sufficient for SNPs to be in LD with most QTL; however, a gap of 700 kb exists between the two significant GWAS peaks on chromosome 16 (Fig. 3.3, 3.4, 3.5) where none of the SNPs tested in GWAS were called.

For each of the four panels, the BIC was used to determine that none of the principal components needed to be in the optimum GWAS model. Given that the ancestry of Asian soybean accessions can be traced back to a domestication event that reduced diversity from a wild relative (*Glycine soja* Siebold & Zucc.) that already had low sequence diversity (Hyten et al., 2006), this is not particularly surprising. However, two explanations for this result are that the kinship matrix is

explaining both population structure and familial relatedness, or BSR resistance is not associated with population structure. To explore this possibility further, the correlation of BSR resistance and principal components was calculated, and no meaningful correlations were found (data not shown).

Genome-Wide Association Study and Stepwise Procedures

In three of the four tested panels, statistically significant GWAS signals were identified in the chromosome 16 region where three *Rbs* genes were previously mapped. Interestingly, one SNP in this region (ss715624573) had the strongest association with BSR resistance in two of the panels. Although this SNP was not identified in the N-1989 or B-1997 panels, it is only 0.22 Mb from the chromosome 16 SNP identified in N-1989 as having the peak signal. Based on the position of the significant markers in the MLMM analysis, our results have narrowed the region containing BSR resistance QTL on chromosome 16 to between 32.8 and 33.1 Mb based on the Glyma2.0 assembly (Table 3.3), which is a substantial refinement compared with the 10.2 Mb interval identified in Patzoldt et al. (2005b). These results also agree with fine mapping results of Rincker et al. (2016) that map the position of *Rbs1*, *Rbs2*, and *Rbs3* to between the positions of 32.7 and 33.1 Mb. We therefore expect these results to directly benefit MAS breeding efforts by providing a focused region on chromosome 16 from which to make selections for improved BSR resistance.

Although three *Rbs* genes have been previously mapped to the same region of chromosome 16 (Bachman et al., 2001; Lewers et al., 1999; Patzoldt et al., 2005b), the results of our study failed to confirm the presence of three distinct genes. It is possible that the three previously identified *Rbs* genes are in fact one gene and that environmental effects or errors in the original allelism tests could have resulted in the incorrect conclusion that different resistance genes were present in resistance sources. Although our study identified multiple significant SNPs located in a narrow interval, it is possible that a single QTL is present, and the different patterns of LD within the four panels result in unique sets of SNPs with significant associations (Platt et al., 2010). A second possibility is that the multiple significant SNPs identified in our study arise from multiple QTL within the region. Nucleotide binding site–leucine rich repeat (NBS-LRR) genes have been

found to be the largest class of disease resistance genes in flowering plants (Dangl and Jones, 2001), and many NBS-LRR genes have been identified in this defined region of the Williams 82 genome sequence. Furthermore, duplicate NBS-LRR genes have been found to confer disease resistance in plants (Dixon et al., 1996). In soybean, Kang et al. (2012) found that locations of disease resistance QTL were correlated with the number of NBS-LRR genes. In fact, resistance to multiple soybean pests such as Phytophthora root rot (*Rps1k*), soybean mosaic virus (*Rsv1* and *Rsv3*), and Asian soybean rust (*Rpp4*) have been mapped or cloned to NBS-LRR genes (Gao et al., 2005; Hayes et al., 2004; Meyer et al., 2009; Suh et al., 2011).

The identification of BSR resistance QTL other than at the chromosome 16 *Rbs* interval has been elusive. The variation explained in each of our panels (Table 3.3) remains low, which suggests environmental effects are large, or more loci are responsible for resistance, or both. The most promising region identified in this study is a location distal of the *Rbs* interval at the positions of 36.6 to 37.4 Mb on chromosome 16. Three of the four GWAS panels contained significant SNPs from this location. However, no significant associations were found in this region when reconducting the GWAS with SNPs identified in MLMM as covariates. Lewers et al. (1999) identified two separate QTL on chromosome 16. Based on previous linkage maps, the restriction fragment-length polymorphic markers used in their study are not located near the positions found in our study. It is possible that the markers used by Lewers et al. (1999) might have been located near the significant GWAS signals identified in the N-1989 panel at 29.6 Mb. Regardless, neither of these locations identified in GWAS on each side of the *Rbs* interval were identified in the MLMM analysis or with stepwise logistic regression. In addition to these locations on chromosome 16, other significant SNPs identified within the N-1989 (chromosomes 2 and 17) and B-2000 (chromosome 8) panels could be involved in resistance. Because of the limited replication of phenotypic data obtained for GWAS and the nature of the binary data, more evidence is needed to substantiate these QTL. Increased replication of phenotypic data within these panels may identify additional QTL, but it is possible that the frequency of BSR resistance alleles at QTL other than the chromosome 16 *Rbs* locus is too low for a traditional GWAS to detect (Cardon and Bell, 2001). In this case a biparental or family-based association mapping population would be needed. Presumably, the significant association on chromosome 5 in the B-2000 panel is a false positive because it was not included in the optimal MLMM and was not

significant when reconducting the GWAS with SNPs identified in the MLMM as covariates. The presence of another unknown QTL that interacts with the *Rbs* locus has been proposed (Bachman and Nickell, 2000b; Sebastian et al., 1985; Waller et al., 1991). If loci identified in our GWAS were substantiated, further research would be needed to determine if epistatic interactions with the *Rbs* locus exists.

The lack of significance in the B-1997 panel was unexpected because Patzoldt et al. (2005a) developed biparental populations with resistant accessions identified by Bachman et al. (1997a) and present in the B-1997 panel. Patzoldt et al. (2005a) then conducted linkage mapping with molecular markers on chromosome 16 and in all five of the resistant sources studied, QTL were mapped to the same region where *Rbs* was previously mapped. This suggests that the resistance QTL in the *Rbs* interval of chromosome 16 is present in the landrace population of the central China region where accessions included in the B-1997 GWAS were collected. It is possible that the lack of association is the result of a low frequency of resistance alleles found in the accessions of this panel, and a different design would be needed to identify an association on chromosome 16 (Cardon and Bell, 2001). Another explanation of the lack of GWAS signals for this panel is that many accessions did not have foliar symptoms and were possibly disease escapes. Yet another possibility is the inoculum for the 1994 Urbana, IL, field test where these accessions were evaluated. The *Rbs* gene may not have been effective against the population of *C. gregata* present in the field during that year. No interactions between *C. gregata* isolates and host genotypes have been reported but do exist in other soybean pathogenic fungi (Grau et al., 2004).

In this study, association mapping was used to refine the location of the known *Rbs* gene or genes and to identify new putative associations that should be tested in future research. This study demonstrates the ability to use data from the SoySNP50K array and historical BSR resistance test data to map disease resistance in soybean. The genomic positions identified will aid gene cloning efforts. Furthermore, the markers identified in this study will improve MAS and development of resistant cultivars. Regardless, the prominence of only one major QTL for BSR resistance does present a challenge in the event that an isolate is able to overcome resistance from this locus. Further research is needed to validate additional QTL or distinguish the

previously reported *Rbs* genes. In addition, research is needed to evaluate resistance to additional isolates and inoculation methods that can increase our overall resistance in germplasm.

LITERATURE CITED

Agresti, A. 2013. Categorical data analysis. 3rd ed. John Wiley & Sons. Hoboken, NJ.

Bachman, M.S. 1999. Characterization of brown stem rot resistance in soybean. Ph.D. diss. Univ. of Illinois. Urbana–Champaign.

Bachman, M.S., and C.D. Nickell. 2000a. High frequency of brown stem rot resistance in soybean germplasm from central and southern China. Plant Dis. 84:694–699. doi:10.1094/PDIS.2000.84.6.694

Bachman, M.S., and C.D. Nickell. 2000b. Investigating the genetic model for brown stem rot resistance in soybean. J. Hered. 91:316–321. doi:10.1093/jhered/91.4.316

Bachman, M.S., C.D. Nickell, P.A. Stephens, and A.D. Nickell. 1997a. Brown stem rot resistance in soybean germ plasm from central China. Plant Dis. 81:953–956. doi:10.1094/PDIS.1997.81.8.953

Bachman, M.S., C.D. Nickell, P.A. Stephens, A.D. Nickell, and L.E. Gray. 1997b. The effect of *Rbs2* on yield of soybean. Crop Sci. 37:1148–1151. doi:10.2135/cropsci1997.0011183X003700040019x

Bachman, M.S., J.P. Tamulonis, C.D. Nickell, and A.F. Bent. 2001. Molecular markers linked to brown stem rot resistance genes *Rbs1* and *Rbs2* in soybean. Crop Sci. 41:527–535. doi:10.2135/cropsci2001.412527x

Bastien, M., H. Sonah, and F. Belzile. 2014. Genome wide association mapping of *Sclerotinia sclerotiorum* resistance in soybean with a genotyping-by-sequencing approach. Plant Gen. 7:1–13. doi:10.3835/plantgenome2013.10.0030

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Royal Stat. Soc. 57:289–300.

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

Box, G.E.P., and D.R. Cox. 1964. An analysis of transformations. J. Royal Stat. Soc. B. Met. 26:211–234.

Cardon, L.R., and J.I. Bell. 2001. Association study designs for complex diseases. Nat. Rev. Genet. 2:91–99. doi:10.1038/35052543

Chamberlain, D.W., and R.L. Bernard. 1968. Resistance to brown stem rot in soybean. Crop Sci. 8:728–729. doi:10.2135/cropsci1968.0011183X000800060026x

Chen, J., and Z. Chen. 2008. Extended Bayesian information criteria for model selection with large model spaces. Biometrika 95:759–771. doi:10.1093/biomet/asn034

Dangl, J.L., and J.D. Jones. 2001. Plant pathogens and integrated defense responses to infection. Nature 411:826–833. doi:10.1038/35081161

Dixon, M.S., D.A. Jones, J.S. Keddie, C.M. Thomas, K. Harrison, and J.D.G. Jones. 1996. The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 84:451–459. doi:10.1016/S0092-8674(00)81290-8

Gao, H., N.N. Narayanan, L. Ellison, and M.K. Battacharyya. 2005. Two classes of highly similar coiled coil-nucleotide binding-leucine rich repeat genes isolated from the *Rps1-k* locus encode *Phytophthora* resistance in soybean. Mol. Plant Microbe Interact. 18:1035–1045. doi:10.1094/MPMI-18-1035

Grau, C.R., A.E. Dorrance, J. Bond, and J.S. Russin. 2004. Fungal diseases. In: H.R. Boerma and J.E. Specht, editors, Soybeans: Improvement, production, and uses. 3rd ed. Agron. Monogr. 16. ASA, CSSA, and SSSA, Madison, WI. p. 679-763.

Gray, L.E. 1972. Effect of *Cephalosporium gregatum* on soybean yield. Plant Dis. Rep. 56:580–581.

Hanson, P.M., C.D. Nickell, L.E. Gray, and S.A. Sebastian. 1988. Identification of two dominant genes conditioning brown stem rot resistance in soybean. Crop Sci. 28:41–43. doi:10.2135/cropsci1988.0011183X002800010010x

Hayes, A.J., S.C. Jeong, M.A. Gore, Y.G. Yu, G.R. Buss, S.A. Tolin, and M.A. Saghai Maroof. 2004. Recombination within a nucleotide binding-site/leucine-rich-repeat gene cluster produces new variants conditioning resistance to *Soybean mosaic virus* in soybeans. Genetics 166:493–503. doi:10.1534/genetics.166.1.493

Hughes, T.J., N.C. Kurtzweil, B.W. Diers, and C.R. Grau. 2004. Resistance to brown stem rot in soybean germ plasm with resistance to the soybean cyst nematode. Plant Dis. 88:761–768. doi:10.1094/PDIS.2004.88.7.761

Hwang, E., Q. Song, G. Jia, J.E. Specht, D.L. Hyten, J. Costa, and P.B. Cregan. 2014. A genome-wide association study of seed protein and oil content in soybean. BMC Genomics 15:1. doi:10.1186/1471-2164-15-1

Hyten, D.L., Q. Song, Y. Zhu, I.Y. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker, and P.B. Cregan. 2006. Impacts of genetic bottlenecks on soybean genome diversity. Proc. Natl. Acad. Sci. USA 103:16666–16671. doi:10.1073/pnas.0604379103

Kang, Y.J., K.H. Kim, S. Shim, M.Y. Yoon, S. Sun, M.Y. Kim, K. Van, and S. Lee. 2012. Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. BMC Plant Biol. 12:139. doi:10.1186/1471-2229-12-139

Koenning, S.R., and J.A. Wrather. 2010. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. (online) Plant Health Prog. doi:10.1094/PHP-2010-1122-01-RS

Lewers, K.S., E.H. Crane, C.R. Bronson, J.M. Schupp, P. Keim, and R.C. Shoemaker. 1999. Detection of linked QTL for soybean brown stem rot resistance in 'BSR 101' as expressed in a growth chamber environment. Mol. Breed. 5:33–42. doi:10.1023/A:1009634710039

Lipka, A.E., F. Tian, Q. Wang, J. Peiffer, M. Li, P.J. Bradbury, M.A. Gore, E.S. Buckler, and Z. Zhang. 2012. GAPIT: Genome association and prediction integrated tool. Bioinformatics 28:2397–2399. doi:10.1093/bioinformatics/bts444

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In: P. Zarembka, editor, Frontiers in Econometrics. Academic Press, New York. p. 105–142.

Meyer, J.D.F., D.C.G. Silva, C. Yang, K.F. Pedley, C. Zhang, M. Mortel, J.H. Hill, R.C. Shoemaker, R.V. Abdelnoor, S.A. Whitham, and M.A. Graham. 2009. Identification and analyses of candidate genes for *Rpp4*-mediated resistance to Asian soybean rust in soybean. Plant Physiol. 150:295–307. doi:10.1104/pp.108.134551

Nelson, R.L., C.D. Nickell, J.H. Orf, H. Tachibana, E.T. Gritton, C.R. Grau, and B.W. Kennedy. 1989. Evaluating soybean germplasm for brown stem rot resistance. Plant Dis. 73:110–114. doi:10.1094/PD-73-0110

Patzoldt, M.E., S.R. Carlson, and B.W. Diers. 2005a. Characterization of resistance to brown stem rot of soybean in five accessions from central China. Crop Sci. 45:1092–1095. doi:10.2135/cropsci2004.0393

Patzoldt, M.E., W. Chen, and B.W. Diers. 2003. Evaluation of soybean plant introductions from China for resistance to brown stem rot. (online) Plant Health Prog. doi:10.1094/PHP-2003-0702-01-RS

Patzoldt, M.E., C.R. Grau, P.A. Stephens, N.C. Kurtzweil, S.R. Carlson, and B.W. Diers. 2005b. Localization of a quantitative trait locus providing brown stem rot resistance in the soybean cultivar Bell. Crop Sci. 45:1241–1248. doi:10.2135/cropsci2003.0615

Perez, P.T., B.W. Diers, P. Lundeen, G.M. Tabor, and S.R. Cianzio. 2010. Genetic analysis of new sources of soybean resistance to brown stem rot. Crop Sci. 50:2431–2439. doi:10.2135/cropsci2010.03.0159

Platt, A., B.J. Vilhjálmsson, and M. Nordborg. 2010. Conditions under which genome-wide association studies will be positively misleading. Genetics 186:1045–1052. doi:10.1534/genetics.110.121665

Price, A.L., N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38:904–909. doi:10.1038/ng1847

Rincker, K., G.L. Hartman, and B.W. Diers. 2016. Fine mapping of resistance genes from five brown stem rot resistance sources in soybean. Plant Gen. 9(1). doi:10.3835/plantgenome2015.08.0063

SAS Institute. 2011. The SAS system for Microsoft Windows. Release 9.3. SAS Inst. Inc., Cary, NC.

Schwarz, G. 1978. Estimating the dimension of a model. Ann. Stat. 6:461–464. doi:10.1214/aos/1176344136

Sebastian, S.A., C.D. Nickell, and L.E. Gray. 1985. Efficient selection for brown stem rot resistance in soybeans under greenhouse screening conditions. Crop Sci. 25:753–757. doi:10.2135/cropsci1985.0011183X002500050008x

Segura, V., B.J. Vilhjálmsson, A. Platt, A. Korte, Ü. Seren, Q. Long, and M. Nordborg. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet. 44:825–830. doi:10.1038/ng.2314

Song, Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS ONE 8:E54985. doi:10.1371/journal.pone.0054985

Suh, S.J., B.C. Bowman, N. Jeong, K. Yang, C. Kastl, S.A. Tolin, M.A. Saghai Maroof, and S. Jeong. 2011. The *Rsv3* locus conferring resistance to *soybean mosaic virus* is associated with a cluster of coiled-coil nucleotide-binding leucine-rich repeat genes. Plant Gen. 4:55–64. doi:10.3835/plantgenome2010.11.0024

Sun, G., C. Zhu, M.H. Kramer, S.S. Yang, W. Song, H.P. Piepho, and J. Yu. 2010. Variation explained in mixed-model association mapping. Heredity 105:333–340. doi:10.1038/hdy.2010.11

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. doi:10.3168/jds.2007-0980

Vaughn, J. N., R. L. Nelson, Q. Song, P. B. Cregan, and Z. Li. 2014. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. G3: Genes, Genomes, Genet. 4:2283–2294. doi:10.1534/g3.114.013433

Waller, R.S., C.D. Nickell, D.L. Drzycimski, and J.E. Miller. 1991. Genetic analysis of the inheritance of brown stem rot resistance in the soybean cultivar Asgrow A3733. J. Hered. 82:412–417.

Wen, Z., R. Tan, J. Yuan, C. Bales, W. Du, S. Zhang, M.I. Chilvers, C. Schmidt, Q. Song, P.B. Cregan, and D. Wang. 2014. Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. BMC Genomics 15:809. doi:10.1186/1471-2164-15-809

Willmont, D.B., and C.D. Nickell. 1989. Genetic analysis of brown stem rot resistance in soybean. Crop Sci. 29:672–674. doi:10.2135/cropsci1989.0011183X002900030026x

Wrather, A., G. Shannon, R. Balardin, L. Carregal, R. Escobar, G.K. Gupta, Z. Ma, W. Morel, D. Ploper, and A. Tenuta. 2010. Effect of diseases on soybean yield in the top eight producing countries in 2006. (online) Plant Health Prog. doi:10.1094/PHP-2010-0125-01-RS

Yu, J., G. Pressoir, W.H. Briggs, I. Vroh Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203–208. doi:10.1038/ng1702

Zhang, Z., E. Ersoz, C.Q. Lai, R.J. Todhunter, H.K. Tiwari, M.A. Gore, P.J. Bradbury, J. Yu, D.K. Arnett, J.M. Ordovas, and E.S. Buckler. 2010. Adaptation of mixed linear model for genome-wide association studies. Nat. Genet. 42:355–360. doi:10.1038/ng.546

TABLES AND FIGURES

Table 3.1. Characteristics of association panels analyzed with GWAS and stepwise procedures.

						BSR Score †		
Panel	Data type	Symptoms measured	Accessions	SNPs	Box-Cox lambda	Mean	SD ‡	h^2 §
N-1989	binary	foliar and stem	2,773	33,240	na	na	na	na ¶
B-1997	proportion 0 to 1	foliar	540	33,486	log	0.09	0.15	0.49
B-1997	proportion 0 to 1	stem	540	33,486	1	0.38	0.20	0.61
B-2000	proportion 0 to 1	foliar	825	32,150	0.25	0.33	0.29	0.93
P-2003	proportion 0 to 1	stem	606	29,815	0.75	0.39	0.25	0.68

† Brown stem rot (BSR) score is measured by the percentage of plants or nodes of a plant showing foliar and stem symptoms.

‡ Standard deviation of the mean (SD).

§ Estimated within GAPIT by the use of SNP markers and genetic relatedness among individuals.

¶ Calculation of h2 by GAPIT is not appropriate for this binary trait.

Table 3.2. Significant SNP associations within each panel that were identified in a genome-wide association study of brown stem rot symptoms.

SNP		Chr.	Position GmW82.a2	P-value	Minor Allele Frequency	Sample Size	R^2_{LR} from Model without SNP	R^2_{LR} from Model with SNP	FDR-Adjusted P-values	Effect Size
N-1989										
Gm16_32298597_T_G	ss715624549	16	32,796,708	2.23×10^{-7}	0.31	2773	0.01	0.02	4.30×10^{-3}	0.03
Gm16_32265231_C_T	ss715624543	16	32,763,342	2.59×10^{-7}	0.21	2773	0.01	0.02	4.30×10^{-3}	-0.03
Gm17_37574384_T_C	ss715627222	17	37,284,864	5.18×10^{-7}	0.45	2773	0.01	0.02	5.74×10^{-3}	0.04
Gm16_32940363_G_A	ss715624611	16	33,415,484	1.35×10^{-6}	0.15	2773	0.01	0.02	1.12×10^{-2}	0.02
Gm17_37742364_A_G	ss715627239	17	37,452,896	5.25×10^{-6}	0.09	2773	0.01	0.02	3.49×10^{-2}	-0.02
Gm02_4208733_C_T	ss715582351	2	4,260,493	7.09×10^{-6}	0.31	2773	0.01	0.02	3.93×10^{-2}	0.03
Gm16_32349812_C_T	ss715624559	16	32,847,923	8.72×10^{-6}	0.13	2773	0.01	0.02	4.14×10^{-2}	-0.02
Gm02_4312213_C_T	ss715582534	2	4,363,973	1.31×10^{-5}	0.38	2773	0.01	0.02	5.08×10^{-2}	0.03
Gm16_36650773_A_G	ss715624945	16	37,153,578	1.37×10^{-5}	0.10	2773	0.01	0.02	5.08×10^{-2}	-0.02
Gm16_29272856_G_A	ss715624168	16	†	1.99×10^{-5}	0.09	2773	0.01	0.02	5.54×10^{-2}	-0.02
Gm16_33172651_T_C	ss715624616	16	33,663,403	2.01×10^{-5}	0.47	2773	0.01	0.02	5.54×10^{-2}	0.05
Gm02_4198619_A_G	ss715582341	2	4,250,379	2.14×10^{-5}	0.38	2773	0.01	0.02	5.54×10^{-2}	0.01
Gm16_29253523_G_A	ss715624163	16	29,621,120	2.17×10^{-5}	0.10	2773	0.01	0.02	5.54×10^{-2}	-0.02
Gm02_4215608_A_G	ss715582353	2	4,267,368	3.27×10^{-5}	0.41	2773	0.01	0.02	7.69×10^{-2}	0.01
Gm16_29211869_G_A	ss715624149	16	29,578,794	3.47×10^{-5}	0.05	2773	0.01	0.02	7.69×10^{-2}	-0.02
B-2000										
Gm16_32517934_C_T	ss715624573	16	33,018,083	4.79×10^{-35}	0.41	825	0.18	0.34	1.54×10^{-30}	-0.09
Gm16_32342381_C_T	ss715624558	16	32,840,492	1.84×10^{-25}	0.18	825	0.18	0.29	2.95×10^{-21}	0.10
Gm16_32340079_G_A	ss715624557	16	32,838,190	1.03×10^{-21}	0.47	825	0.18	0.27	1.10×10^{-17}	-0.07
Gm16_32843154_G_T	ss715624593	16	33,318,274	1.12×10^{-19}	0.22	825	0.18	0.26	8.99×10^{-16}	-0.08
Gm16_32526792_T_C	ss715624574	16	†	1.43×10^{-18}	0.13	825	0.18	0.26	9.17×10^{-15}	-0.10
Gm16_32681330_C_T	ss715624587	16	33,182,780	3.17×10^{-18}	0.38	825	0.18	0.25	1.70×10^{-14}	-0.06
Gm16_32148034_A_G	ss715624527	16	32,637,581	3.28×10^{-17}	0.37	825	0.18	0.25	1.50×10^{-13}	0.07

Table 3.2. (cont.)

SNP		Chr.	Position GmW82.a2	P-value	Minor Allele Frequency	Sample Size	R^2_{LR} from Model without SNP	R^2_{LR} from Model with SNP	FDR-Adjusted P-values	Effect Size
Gm16_32665742_T_C	ss715624585	16	33,167,192	2.12×10^{-16}	0.18	825	0.18	0.25	7.11×10^{-13}	0.08
Gm16_32266024_T_C	ss715624544	16	32,764,135	2.21×10^{-16}	0.17	825	0.18	0.25	7.11×10^{-13}	0.09
Gm16_32876100_A_G	ss715624596	16	33,351,218	1.44×10^{-15}	0.29	825	0.18	0.24	4.20×10^{-12}	-0.06
Gm16_32887637_C_A	ss715624602	16	33,362,755	8.01×10^{-14}	0.33	825	0.18	0.23	2.15×10^{-10}	-0.06
Gm16_32236491_G_A	ss715624540	16	32,734,593	6.30×10^{-13}	0.16	825	0.18	0.23	1.56×10^{-9}	0.07
Gm16_32227195_A_G	ss715624538	16	32,725,297	8.66×10^{-13}	0.30	825	0.18	0.23	1.99×10^{-9}	0.06
Gm16_32099047_A_C	ss715624520	16	32,588,581	5.88×10^{-12}	0.47	825	0.18	0.22	1.26×10^{-8}	0.05
Gm16_32139895_C_T	ss715624526	16	32,629,442	5.56×10^{-11}	0.47	825	0.18	0.22	1.12×10^{-7}	0.05
Gm16_32113900_A_G	ss715624524	16	32,603,434	1.39×10^{-9}	0.49	825	0.18	0.21	2.62×10^{-6}	-0.04
Gm16_32534697_A_G	ss715624576	16	33,038,110	9.18×10^{-9}	0.42	825	0.18	0.21	1.64×10^{-5}	0.04
Gm16_32200441_A_C	ss715624535	16	32,698,542	1.17×10^{-7}	0.34	825	0.18	0.20	1.98×10^{-4}	0.04
Gm16_33224286_G_T	ss715624623	16	33,715,054	1.30×10^{-7}	0.24	825	0.18	0.20	2.09×10^{-4}	-0.04
Gm16_32161757_T_C	ss715624529	16	32,651,304	1.43×10^{-7}	0.43	825	0.18	0.20	2.18×10^{-4}	-0.04
Gm16_32279783_C_T	ss715624546	16	32,777,894	1.90×10^{-7}	0.23	825	0.18	0.20	2.78×10^{-4}	-0.05
Gm16_32154906_T_C	ss715624528	16	32,644,453	8.38×10^{-7}	0.08	825	0.18	0.20	1.17×10^{-3}	-0.06
Gm05_39374746_C_T	ss715591790	5	40,974,254	1.03×10^{-6}	0.28	825	0.18	0.20	1.38×10^{-3}	0.04
Gm05_39535948_C_T	ss715591815	5	40,813,158	6.78×10^{-6}	0.06	825	0.18	0.20	8.72×10^{-3}	0.07
Gm16_36654985_C_A	ss715624946	16	37,157,790	2.03×10^{-5}	0.35	825	0.18	0.19	2.51×10^{-2}	0.03
Gm05_39415018_A_G	ss715591796	5	40,934,088	2.56×10^{-5}	0.06	825	0.18	0.19	3.04×10^{-2}	0.07
Gm08_7571195_T_C	ss715602692	8	7,577,565	2.91×10^{-5}	0.06	825	0.18	0.19	3.34×10^{-2}	0.06
Gm16_33188623_A_C	ss715624618	16	33,679,375	3.39×10^{-5}	0.24	825	0.18	0.19	3.75×10^{-2}	-0.04
Gm05_39513085_C_A	ss715591812	5	40,836,021	5.58×10^{-5}	0.06	825	0.18	0.19	5.98×10^{-2}	-0.07
Gm16_31470069_T_C	ss715624456	16	31,838,543	6.08×10^{-5}	0.48	825	0.18	0.19	6.31×10^{-2}	0.03
Gm16_36049380_G_A	ss715624865	16	36,552,189	6.89×10^{-5}	0.34	825	0.18	0.19	6.93×10^{-2}	-0.03

Table 3.2. (cont.)

SNP		Chr.	Position GmW82.a2	P-value	Minor Allele Frequency	Sample Size	R^2_{LR} from Model without SNP	R^2_{LR} from Model with SNP	FDR-Adjusted P-values	Effect Size
Gm16_32298597_T_G	ss715624549	16	32,796,708	7.27×10^{-5}	0.31	825	0.18	0.19	7.09×10^{-2}	0.03
Gm08_7589397_A_G	ss715602693	8	7,595,767	7.71×10^{-5}	0.07	825	0.18	0.19	7.21×10^{-2}	-0.06
Gm08_7589824_T_C	ss715602695	8	7,596,194	7.85×10^{-5}	0.07	825	0.18	0.19	7.21×10^{-2}	0.06
Gm16_31956105_C_A	ss715624501	16	32,445,623	1.01×10^{-4}	0.44	825	0.18	0.19	9.05×10^{-2}	-0.03
P-2003										
Gm16_32517934_C_T	ss715624573	16	33,018,083	1.59×10^{-7}	0.49	606	0.08	0.12	3.72×10^{-3}	-0.06
Gm16_32617666_T_C	ss715624583	16	33,119,116	3.52×10^{-7}	0.21	606	0.08	0.12	3.72×10^{-3}	-0.07
Gm16_32534697_A_G	ss715624576	16	33,038,110	3.74×10^{-7}	0.49	606	0.08	0.12	3.72×10^{-3}	0.06
Gm16_32227195_A_G	ss715624538	16	32,725,297	1.87×10^{-6}	0.30	606	0.08	0.11	1.39×10^{-2}	0.06
Gm16_32342381_C_T	ss715624558	16	32,840,492	5.00×10^{-6}	0.19	606	0.08	0.11	2.98×10^{-2}	0.06
Gm16_32340079_G_A	ss715624557	16	32,838,190	6.43×10^{-6}	0.41	606	0.08	0.11	3.20×10^{-2}	-0.05
Gm16_32161757_T_C	ss715624529	16	32,651,304	1.28×10^{-5}	0.40	606	0.08	0.11	5.08×10^{-2}	-0.05
Gm16_32681330_C_T	ss715624587	16	33,182,780	1.36×10^{-5}	0.39	606	0.08	0.11	5.08×10^{-2}	-0.05
Gm16_32139895_C_T	ss715624526	16	32,629,442	2.15×10^{-5}	0.45	606	0.08	0.10	7.14×10^{-2}	0.05
Gm16_36888387_C_T	ss715624973	16	37,381,270	2.85×10^{-5}	0.46	606	0.08	0.10	8.49×10^{-2}	-0.05

† SNP marker is not located in Glyma 2.0 assembly

Table 3.3. Significant SNPs identified in stepwise procedures for the foliar or stem symptoms of brown stem rot.

Stepwise Logistic Regression

Panel	Symptoms	SNP added to Model		Chr.	Position, Glyma2.0 assembly	P-value	Bayesian information criterion (BIC)	R^2_{McF} †	R^2_{McF} with all SNPs identified in stepwise logistic regression ‡
N-1989	foliar and stem	Gm17_37574384_T_C	ss715627222	17	37,284,864	5.09E-11	368.181	0.07	0.21
N-1989	foliar and stem	Gm16_32298597_T_G	ss715624549	16	32,796,708	2.09E-08	339.544	0.10	
N-1989	foliar and stem	Gm02_4208733_C_T	ss715582351	2	4,260,493	4.09E-05	330.345	0.06	

Muli-locus mixed model

Panel	Symptoms	SNP added to Model		Chr.	Position, Glyma2.0 assembly	P-value	Extended BIC §	R^2_{LR} †	R^2_{LR} with all SNPs identified in MLMM ¶
B-2000	foliar	Gm16_32340079_G_A	ss715624557	16	32,838,190	1.04E-44	-767.80526	0.10	0.22
B-2000	foliar	Gm16_32517934_C_T	ss715624573	16	33,018,083	2.93E-15	-804.055686	0.17	
P-2003	stem	Gm16_32617666_T_C	ss715624583	16	33,119,116	2.09E-06	-19.7654342	0.04	0.10
P-2003	stem	Gm16_32342381_C_T	ss715624558	16	32,840,492	1.89E-02	-17.4209656	0.03	

† Defined as the likelihood-ratio based R^2 statistic that measures the increase in phenotypic variation explained by adding the indicated SNP into the model.

‡ Defined as the likelihood-ratio based R2 statistic that measures the increase in phenotypic variation explained by adding the three SNPs identified in the "SNP added to model" column compared to the intercept-only model.

§ Chen and Chen, 2008.

¶ Defined as the likelihood-ratio based R2 statistic that compares the variation explained by the final multilocus mixed models (MLMMs) fitted to (i) the foliar symptoms in the B-2000 panel and (ii) the stem symptoms in the P-2003 panel relative to the intercept-only model. For each of these panels, the respective SNPs included in the final MLMM are indicated in the "SNP added to model" column.

Figure 3.1. Histograms of phenotypic data in genome-wide association studies are shown in the following panels: (A) foliar symptoms from the B-1997 study, (B) stem symptoms from the B-1997 study, (C) foliar symptoms from the B-2000 study, and (D) stem symptoms from the P-2003 study.

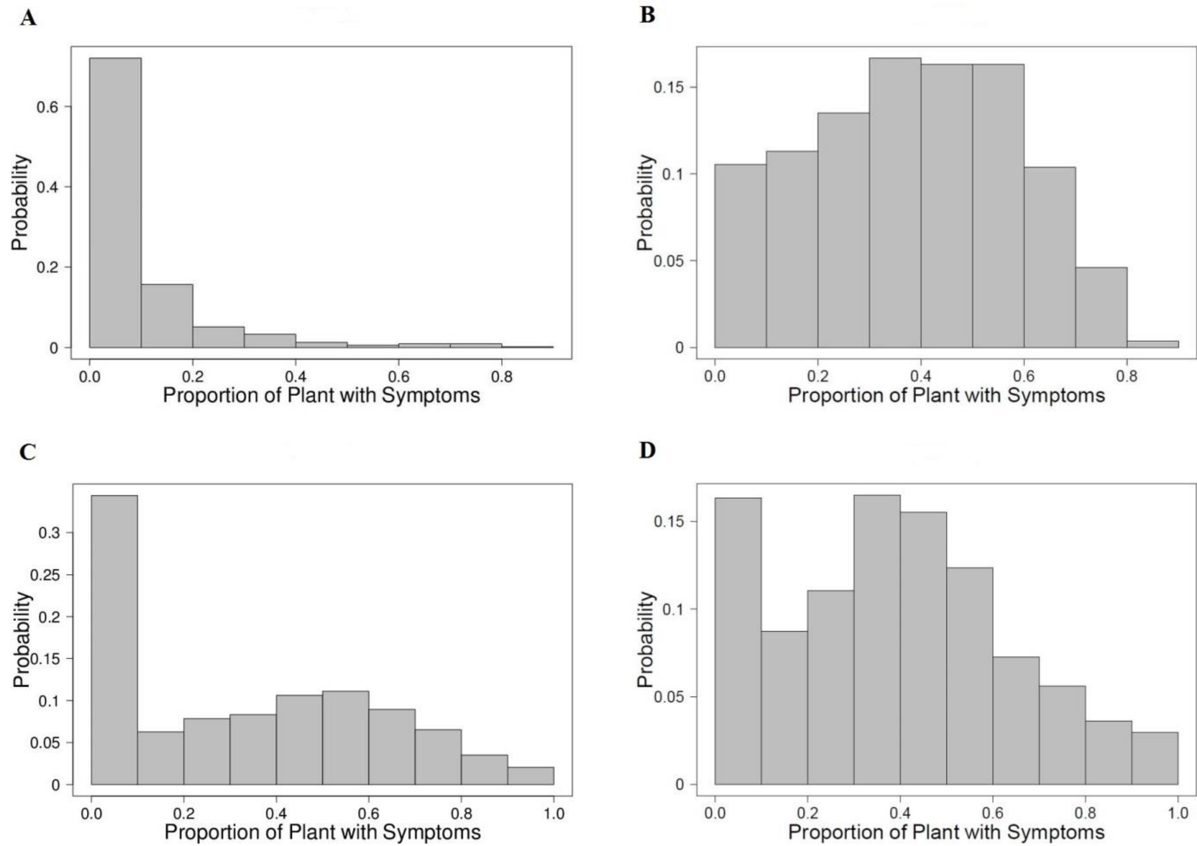


Figure 3.2. Genome-wide association study of foliar and stem symptoms of brown stem rot. Manhattan plots of association results from a unified mixed model analysis. Negative $\log_{10}$–transformed *P*-values (*y*-axis) from a GWAS are plotted against physical position (Glyma.W82.a2) on each of 20 chromosomes. Chromosomes are alternatingly colored and a horizontal line indicates the least significant SNP at 10% FDR, or if no SNPs are significant, it is placed at a Bonferroni corrected significance threshold.

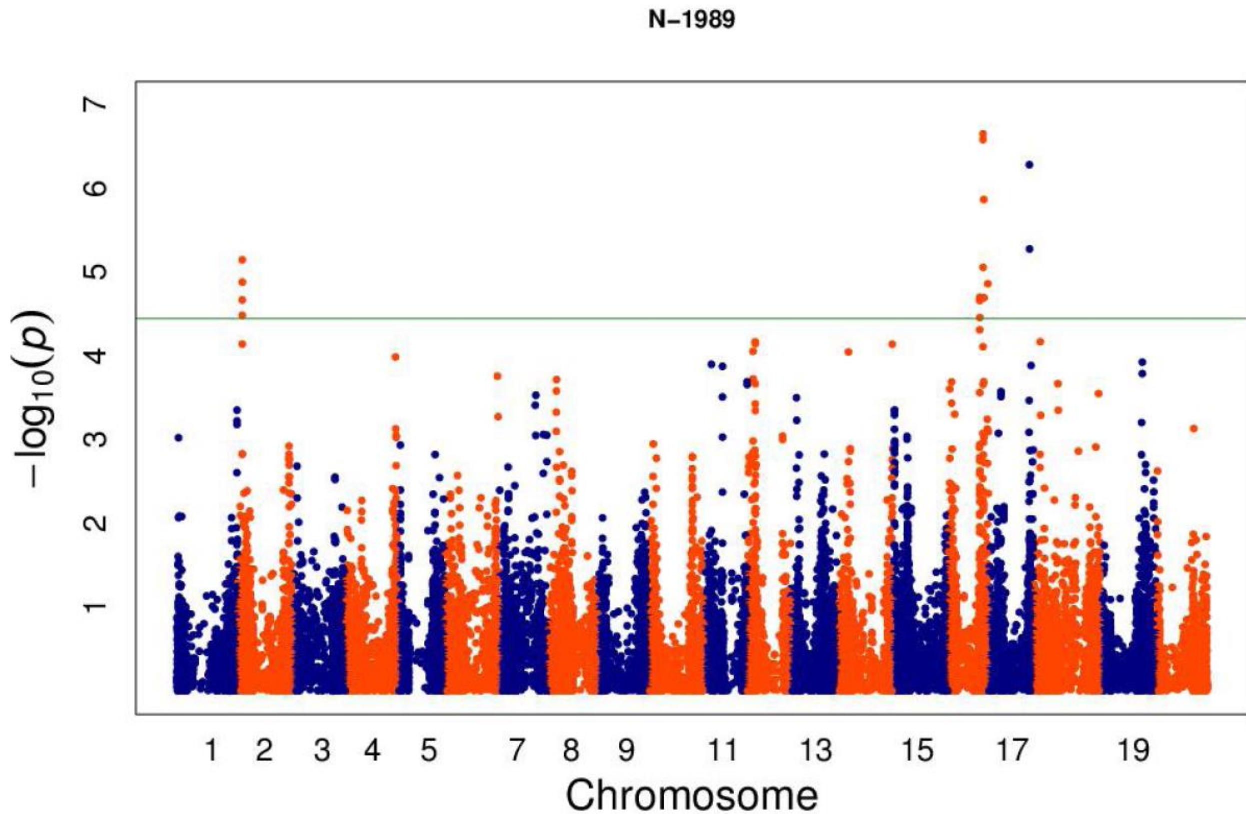


Figure 3.2. (cont.)

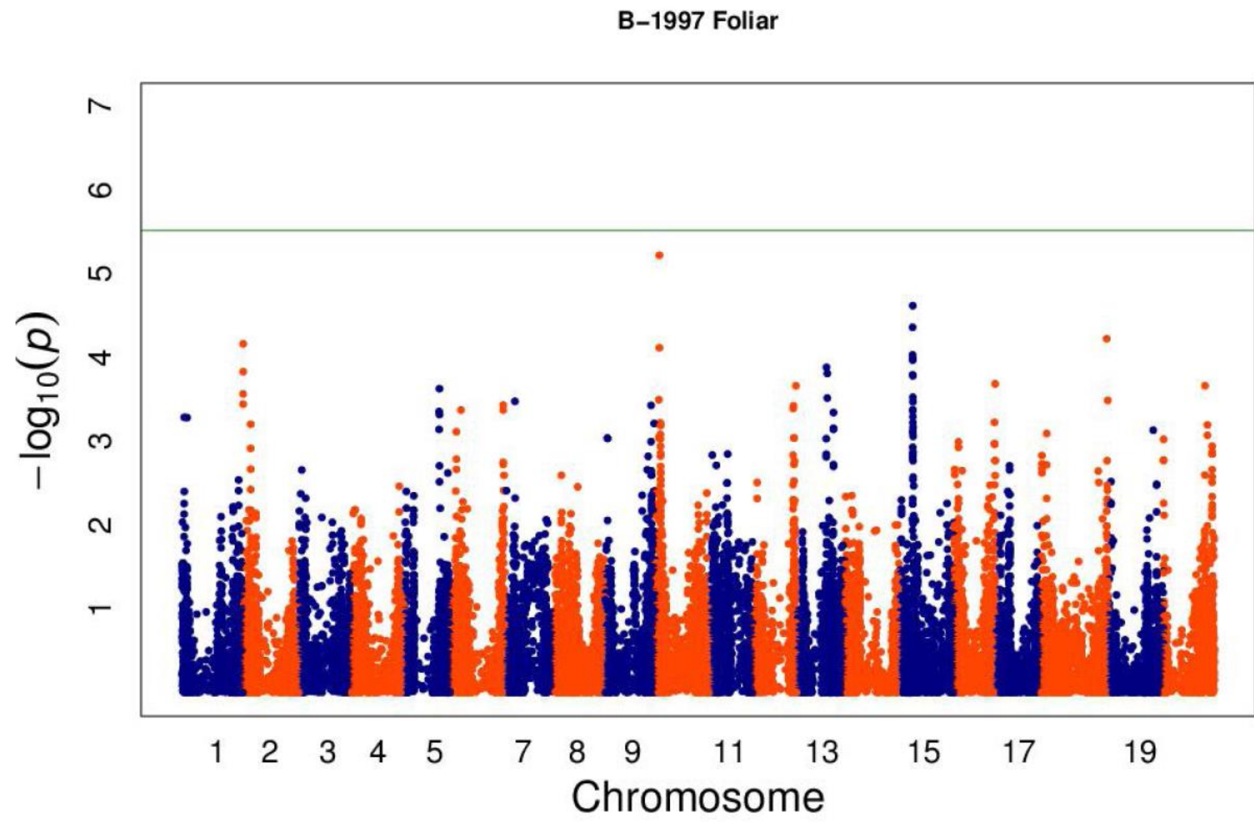


Figure 3.2. (cont.)

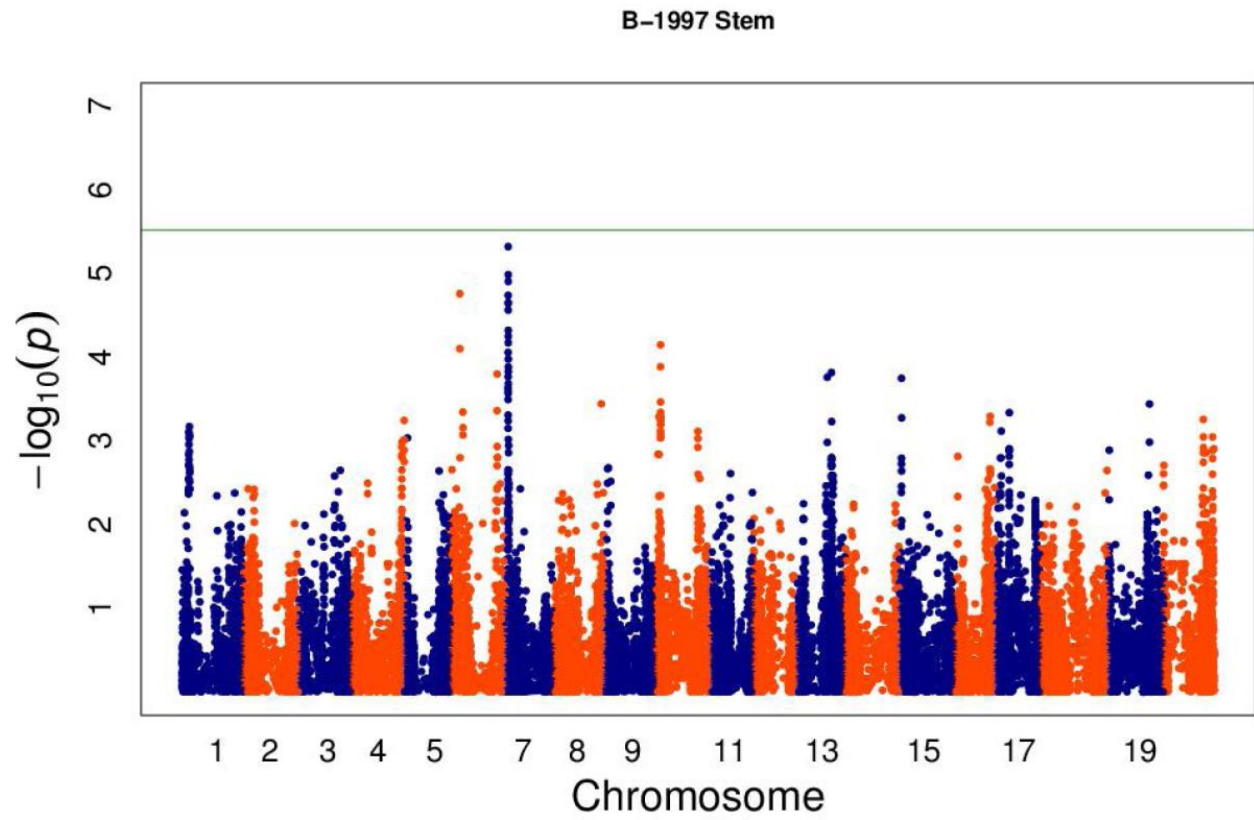


Figure 3.2. (cont.)

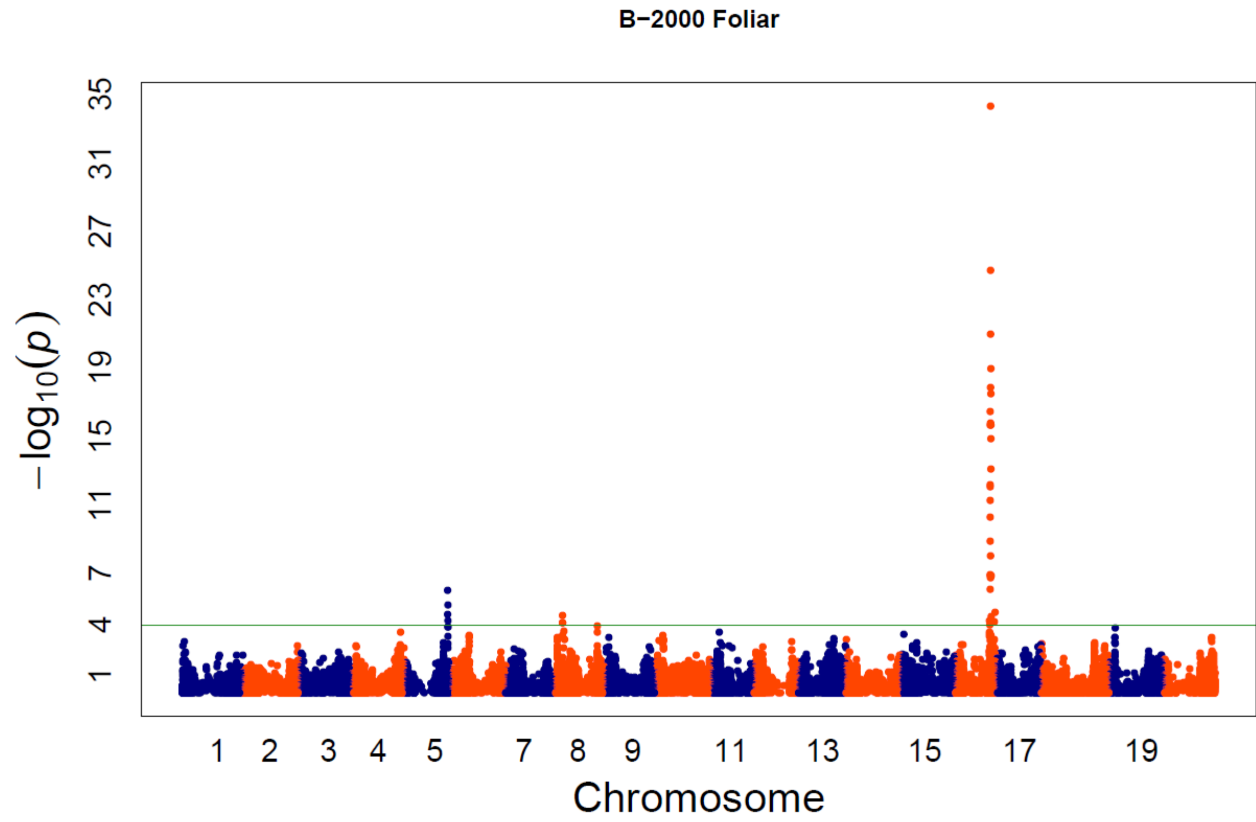


Figure 3.2. (cont.)

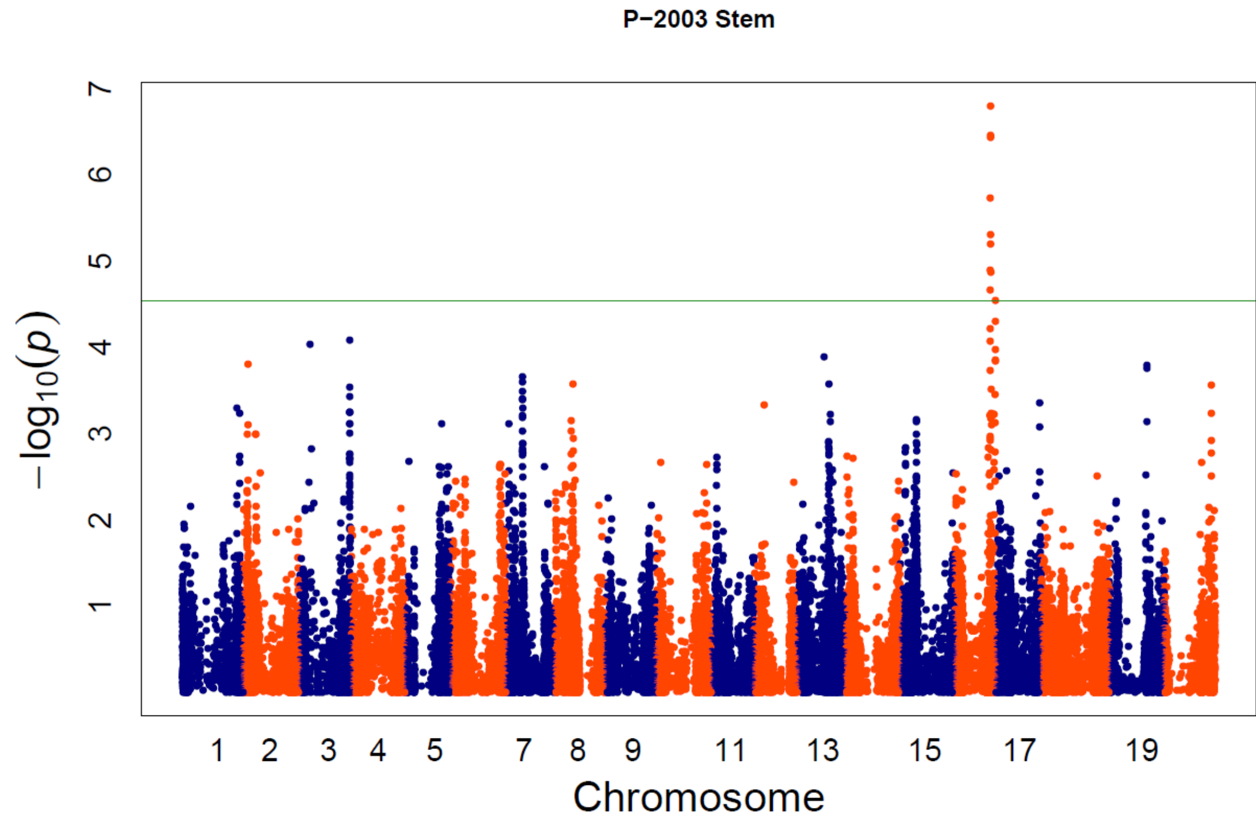


Figure 3.3. Genome-wide association study (GWAS) results for brown stem rot (BSR) resistance of soybean in the N-1989 panel. (A) Association results from a unified mixed model analysis of BSR and linkage disequilibrium (LD) estimates (r^2) across a chromosome 16 region. Negative $\log_{10}$–transformed *P*-values (left, *y*-axis) from a GWAS for BSR and r^2 values (right, *y*-axis) are plotted against physical position (Glyma.W82.a2) for a 8-Mb region on chromosome 16. The blue vertical lines are $-\log_{10}$ *P*-values for single-nucleotide polymorphisms (SNPs) that are statistically significant for BSR at 10% false discovery rate (FDR), whereas the gray vertical lines are $-\log_{10}$ *P*-values for SNPs that are nonsignificant at 10% FDR. Triangles are the r^2 values of each SNP relative to the peak SNP (indicated in red) at 32,796,708 bp of chromosome 16. The black horizontal dashed line indicates the $-\log_{10}$ *P*-value of the least statistically significant SNP at 10% FDR. The black vertical dashed lines indicate the positions of BARCSOYSSR_16_1114 and BARCSOYSSR_16_1115 identified by Rincker et al. (2016), which fine map an interval containing an *Rbs* gene. (B) Association results from a conditional unified mixed model analysis of BSR and LD estimates (r^2) across a chromosome 16 region, as in (A). The three SNPs (ss715627222, ss715624549, and ss715582351) from the stepwise logistic regression model were included as covariates in the unified mixed model to control for the *Rbs* effect.

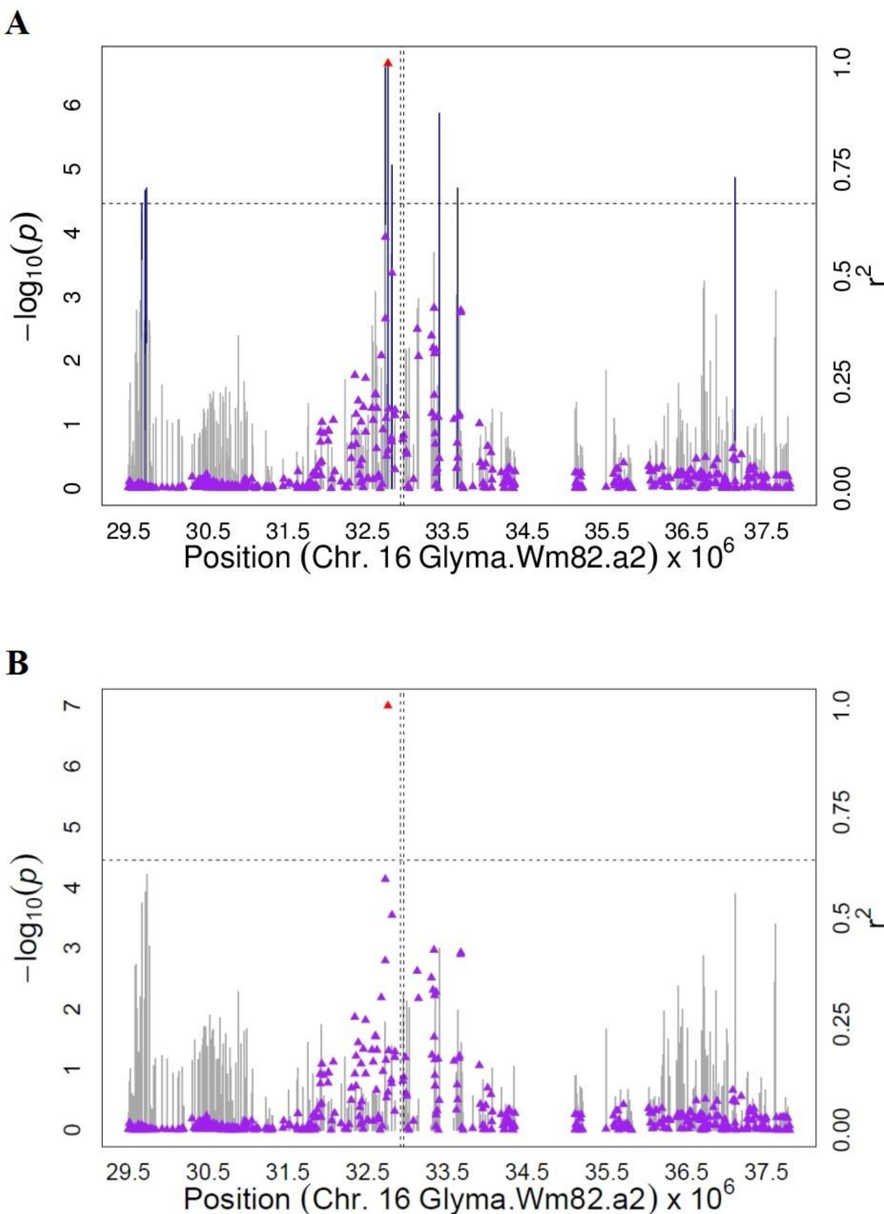


Figure 3.4. Genome-wide association study (GWAS) results for brown stem rot (BSR) resistance of soybean in the B-2000 panel. (A) Association results from a unified mixed model analysis of BSR and linkage disequilibrium (LD) estimates (r^2) across a chromosome 16 region. Negative $\log_{10}$–transformed *P*-values (left, *y*-axis) from a GWAS for BSR and r^2 values (right, *y*-axis) are plotted against physical position (Glyma.W82.a2) for a 6-Mb region on chromosome 16. The blue vertical lines are $-\log_{10}$ *P*-values for single-nucleotide polymorphisms (SNPs) that are statistically significant for BSR at 10% false discovery rate (FDR), whereas the gray vertical lines are $-\log_{10}$ *P*-values for SNPs that are nonsignificant at 10% FDR. Triangles are the r^2 values of each SNP relative to the peak SNP (indicated in red) at 33,018,083 bp of chromosome 16. The black horizontal dashed line indicates the–$\log_{10}$ *P*-value of the least statistically significant SNP at 10% FDR. The black vertical dashed lines indicate the positions of BARCSOYSSR_16_1114 and BARCSOYSSR_16_1115 identified by Rincker et al. (2016), which fine map an interval containing an *Rbs* gene. (B) Association results from a conditional unified mixed model analysis of BSR and LD estimates (r^2) across a chromosome 16 region, as in (A). The two SNPs (ss715624557 and ss715624573) from the optimal multilocus mixed model model were included as covariates in the unified mixed model to control for the *Rbs* effect.

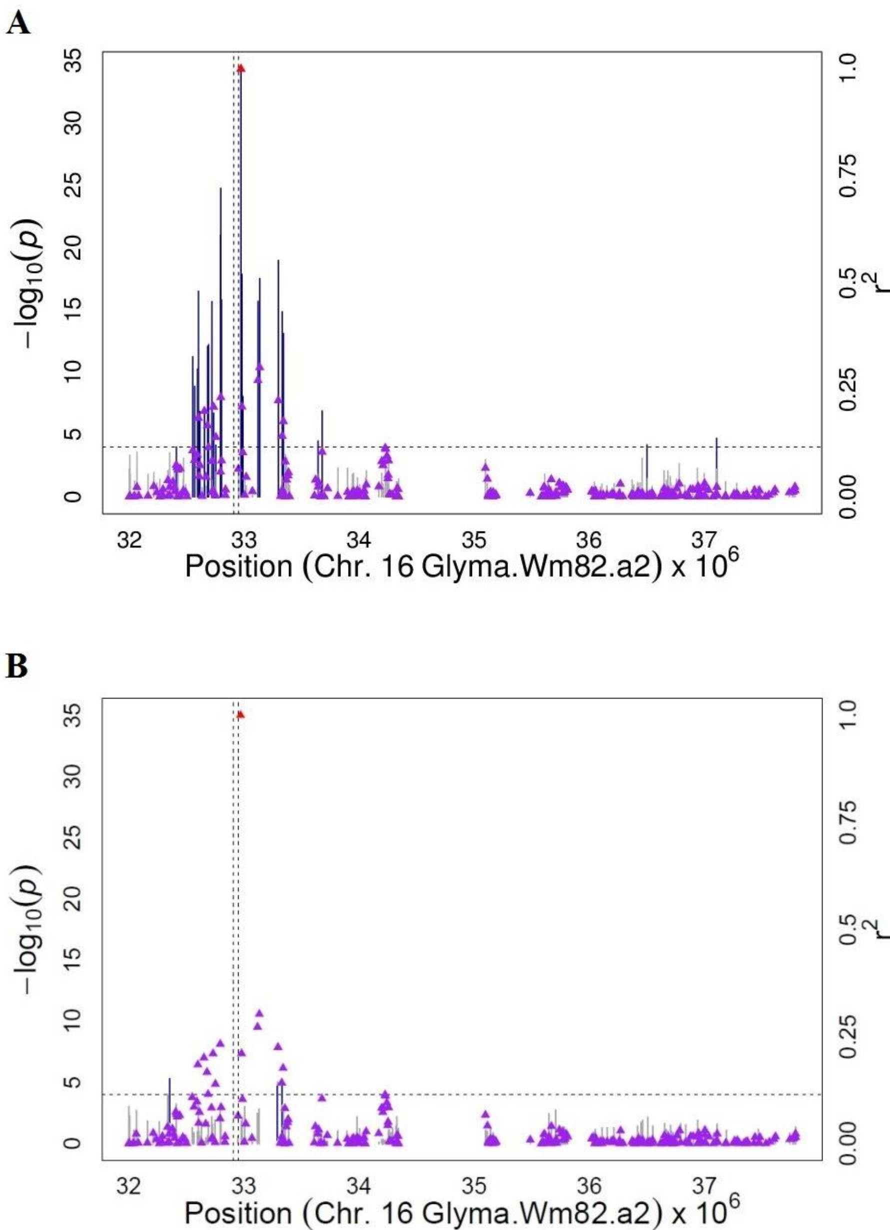
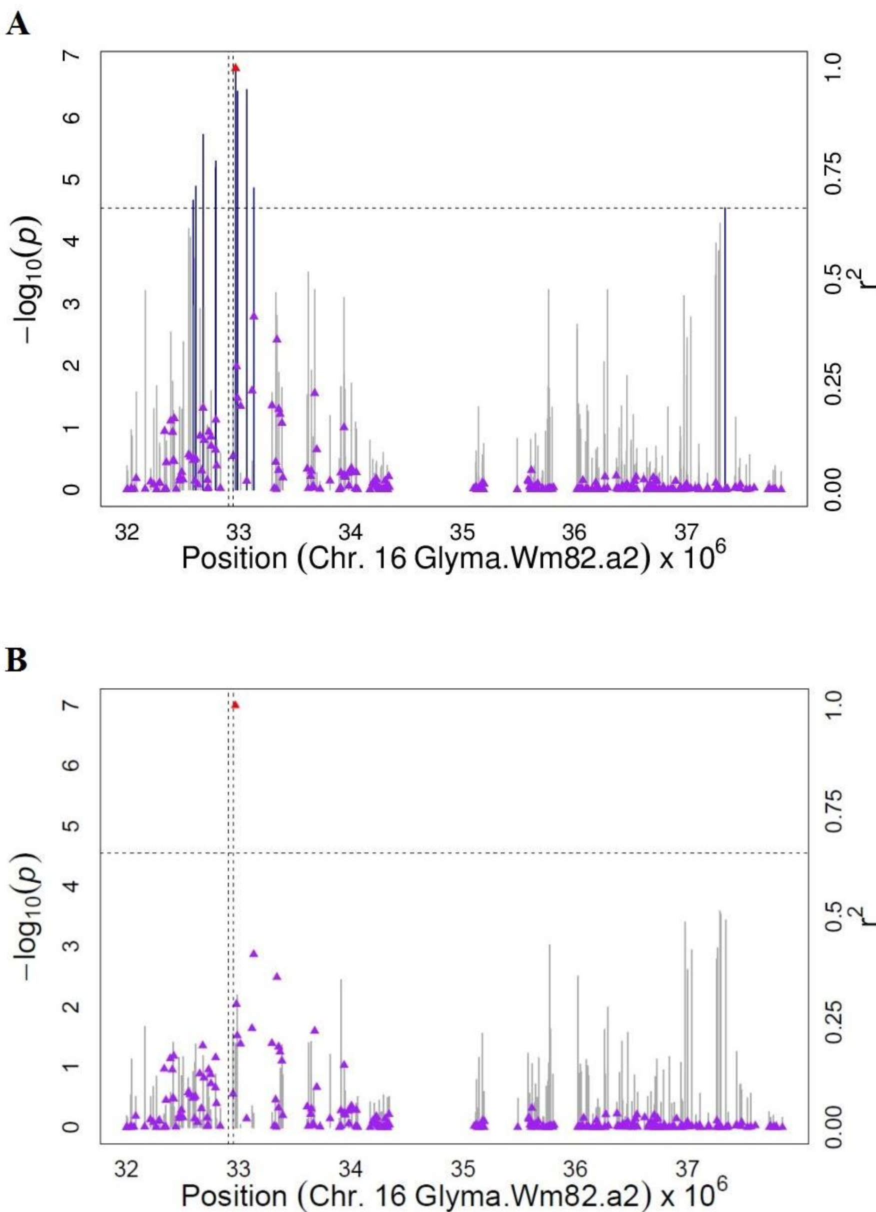


Figure 3.5. Genome-wide association study (GWAS) results for brown stem rot (BSR) resistance of soybean in the P-2003 panel. (A) Association results from a unified mixed model analysis of BSR and linkage disequilibrium (LD) estimates (r^2) across a chromosome 16 region. Negative $\log_{10}$–transformed P-values (left, *y* axis) from a GWAS for BSR and r^2 values (right, *y* axis) are plotted against physical position (Glyma.W82.a2) for a 6-Mb region on chromosome 16. The blue vertical lines are–$\log_{10}$ *P*-values for single-nucleotide polymorphisms (SNPs) that are statistically significant for BSR at 10% false discovery rate (FDR), whereas the gray vertical lines are–$\log_{10}$ *P*-values for SNPs that are nonsignificant at 10% FDR. Triangles are the r^2 values of each SNP relative to the peak SNP (indicated in red) at 33,018,083 bp of chromosome 16. The black horizontal dashed line indicates the–$\log_{10}$ *P*-value of the least statistically significant SNP at 10% FDR. The black vertical dashed lines indicate the positions of BARCSOYSSR_16_1114 and BARCSOYSSR_16_1115 identified by Rincker et al. (2016) which fine map an interval containing an *Rbs* gene. (B) Association results from a conditional unified mixed model analysis of BSR and LD estimates (r^2) across a chromosome 16 region, as in (A). The two SNPs (ss715624583 and ss715624558) from the optimal multilocus mixed model were included as covariates in the unified mixed model to control for the *Rbs* effect.



CHAPTER FOUR

Genetic Improvement of U.S. Soybean in Maturity Groups II, III, IV [1]

INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] production in the United States has evolved from a minor forage crop in the early 1900's to a crop with a harvest of 82.1 million metric tons in 2012 (USDA-NASS, 2013). Soybean is currently a leading source of protein and oil for human food, animal feed, and industrial products (Wilson, 2008). The global rate of yield increase in soybean will need to almost double to keep up with growing demand for the crop that is predicted for 2050 (Ray et al., 2013).

On-farm yield gains arise from the combined impact of grower adoption of new cultivars, improved cultural practices, interactions between new cultivars and improved cultural practices, and environmental factors such as increased atmospheric CO_2 levels (Long et al., 2006; Rowntree et al., 2013a, 2013b; Specht et al., 2014; Ziska and Bunce, 2007). Average on-farm soybean yields in the USA have increased from 738 kg ha^{-1} in 1924 to 2658 kg ha^{-1} in 2012 (USDA-NASS, 2013). Specht et al. (2014) recently reviewed the USA national estimated yields compiled by NASS from 1924 to 2012 and calculated a 23 kg ha^{-1} yr^{-1} on-farm yield gain rate with a simple linear regression model. The authors also noted that a two-segment linear model had a better fit to the data, which showed an increase of 22 kg ha^{-1} yr^{-1} from 1924 to 1983 and 29 kg ha^{-1} yr^{-1} from 1983 to 2012. Wilcox (2001) estimated yield gains across a 60 yr period by reviewing Uniform Soybean Test data for maturity group (MG) 00 to IV entries and found yield increases with maturity groups ranging from 22 kg ha^{-1} yr^{-1} for MG 00 to 31 kg ha^{-1} yr^{-1} for MG III.

The development of soybean cultivars available to farmers has changed significantly over the past 80 years. In the early 1900s, farmers grew either plant introductions (PIs) from East Asia or

[1] Previously published: Rincker K., R. Nelson, J. Specht, D. Sleper, T. Cary, S.R. Cianzio, S. Casteel, S. Conley, P. Chen, V. Davis, C. Fox, G. Graef, C. Godsey, D. Holshouser, G. Jiang, S.K. Kantartzi, W. Kenworthy, C. Lee, R. Mian, L. McHale, S. Naeve, J. Orf, V. Poysa, W. Schapaugh, G. Shannon, R. Uniatowski, D. Wang, B. Diers. 2014. Genetic Improvement of U.S. Soybean in Maturity Groups II, III, IV. Crop Sci. 54:1419-1432.
Copyright owner has provided permission to reprint.

selections from these PIs. The first cultivars developed from breeding programs in North America were released in the 1940s. These cultivars originated from the hybridizations of PIs, which resulted in the development of cultivars better adapted to North American production systems (Hartwig, 1973). As soybean production increased, public and proprietary breeding efforts expanded and began to include selection for pest resistance along with selection for yield (Carter et al., 2004). Proprietary breeding programs have continued to expand in number and in size and these programs currently provide most of the soybean seed sold to farmers in the USA (Specht et al., 2014).

Genetic yield gain was estimated in a number of studies by growing soybean cultivars with different release years in a common environment. Luedders (1977) reported results of testing twenty-one MG I, II, III, and IV soybean cultivars grown in Missouri. He reported an overall increase of 708 kg ha^{-1} for cultivars released between 1964 and 1971 compared to PIs introduced in the 1920s and 30s. When the cultivar yield means reported by Luedders were regressed on release dates, an annual increase of 16 kg ha^{-1} was obtained. Wilcox et al. (1979) tested five MG II and five MG III cultivars in the Midwestern USA and reported 700 kg ha^{-1} in MG II and 625 kg ha^{-1} in MG III yield increase in cultivars released in the 1970's compared to early PIs. When regression analysis was done with these cultivar means, rates of 11 kg ha^{-1} yr^{-1} for MG II and 13 kg ha^{-1} yr^{-1} for MG III were obtained. In Georgia, Boerma (1979) calculated a genetic gain in yield of 14 kg ha^{-1} yr^{-1} for eighteen MG VI, VII, and VIII cultivars released from 1942 to 1973. Specht and Williams (1984) evaluated 240 MG 00 to IV cultivars in Nebraska that were introduced or released from 1902 to 1977 and estimated an average genetic gain of 19 kg ha^{-1} yr^{-1}. Similarly, Salado-Navarro et al. (1993) observed genetic yield gains of 16 to 19 kg ha^{-1} yr^{-1} in MG VI to VIII cultivars tested in Florida. Voldeng et al. (1997) evaluated forty-one cultivars in Quebec and Ontario from MG 000 to 0 that were released from 1934 to 1992 and found an overall yield increase of 11 kg ha^{-1} yr^{-1}. To account for differences in maturity among the tested cultivars, Voldeng et al. (1997) created a maturity-corrected yield index and the regression of this index on year of release resulted in a better fit to a quadratic model, which suggested an accelerating rate of genetic gain over years. In Tennessee, Ustun et al. (2001) reported that soybean yield had increased 14 kg ha^{-1} yr^{-1} in eight MG V to VII cultivars grown in ten Tennessee environments. Recently, De Bruin and Pedersen (2008)

observed an overall genetic gain of 25 kg ha^{-1} yr^{-1} in evaluations of twenty-three cultivars ranging from MG I to III grown in six Iowa environments.

A number of factors can influence yield gains over time. Selection of parents from only elite cultivars and breeding lines has resulted in a narrowing of the genetic base for cultivars released in North America (Gizlice et al., 1994; Sneller, 1994), and this limited genetic diversity could eventually reduce rates of genetic gain. The negative impact on yield of abiotic stresses, such as increased ozone concentration, may continue to increase (Betzelberger et al., 2012). In addition, pests, such as soybean cyst nematode (*Heterodera glycines* Ichinohe), soybean aphid (*Aphis glycines* Matsumura), and Phytophthora root rot (*Phytophthora sojae* Kauffmann and Gerdemann) have added pressure to breeding programs to not only select for yield but also allocate limited breeding resources to concurrent selection for pest resistance, which can slow genetic gains for yield potential. Incorporation of transgenes for herbicide tolerance could also slow gains for yield by diverting resources from yield-focused breeding efforts. In contrast, increasing atmospheric CO_2 levels has and will enhance photosynthesis and water use efficiency in the C3 photosynthetic soybean which theoretically could result in a yield benefit (Specht et al., 1999; Sinclair et al., 1984). Climate change also will have negative impacts on soybean production because of higher summer temperatures and more extreme weather events (Hassol, 2009), however, warmer springs and falls could have a positive impact by lengthening the growing season which could be exploited by earlier planting and the use of cultivars with later maturities (Specht et al., 2014).

Another factor that likely impacted genetic gain was the passage of the Plant Variety Protection (PVP) Act of 1970. This spurred proprietary investment in soybean breeding programs because it gave plant breeders exclusive control over newly developed cultivars, thereby resulting in a greater return on investment (Fehr, 1991; USDA, 2006). Since the passage of the act, Carter et al. (2004) documented an increase in the number of North American soybean breeders, especially in the private sector. In addition, improvements in plot planting, harvesting equipment, and computing capacity enabled increases in field plot testing and improved precision in selection, resulting in greater genetic gains (Eathington et al., 2007). New technologies and methodologies may arise that could further accelerate future yield gains. For

example, genomic selection has been proposed to shorten the breeding cycle (Meuwissen et al., 2001; Bernardo, 2010b; Hefner et al., 2009), and DNA and RNA sequencing will increase our understanding of gene function, which can be utilized in the development of future cultivars (Poland and Rife, 2012; Schmutz et al., 2010).

In addition to gains in genetic yield potential, genotype by environment interactions are constant challenges for plant breeders developing new cultivars. By increasing yield potential in a favorable environment, an improved cultivar must still perform well in high stress environments. Finlay and Wilkinson (1963) developed a method to assess yield stability via regression of a cultivar mean yield with the mean yield of all cultivars in a specific environment. Cultivar regression coefficients greater than one denote less yield stability (i.e., a steeper yield response to increases in environmental productivity), while coefficients less than one denote more yield stability (i.e. a shallower yield response). By using this method, De Bruin and Pedersen (2008), Voldeng et al. (1997), and Wilcox et al. (1979) all determined that soybean yields have increased over generations of breeding without detectable reductions in yield stability.

Although the magnitude of yield improvement over time is of primary interest when developing new soybean cultivars, other traits also may have changed due to direct or indirect selection. For instance, resistance to lodging is an important trait that breeders must consider when releasing a cultivar. Previous research has shown that lodging resistance has improved over generations of breeding. Luedders (1977) found that lodging scores (scaled in units from 1-erect, 5-prostrate) decreased by 0.9 units when cultivars released between 1964 and 1971 were compared to early PIs. Specht and Williams (1984) reported a decrease of 1.0 unit between cultivars released before 1977 and original PIs. Voldeng et al. (1997) also detected a decrease in lodging score from old to new cultivars in the early MGs, and Wilcox et al. (1979) found a decrease in lodging among MG III cultivars, but not among MG II cultivars.

Although seed protein and oil concentration, mature plant height, seed size, and seed quality are important traits, unless their magnitudes deviate substantially from the norm, yield, maturity, and lodging are the primary considerations in most cultivar release decisions. Over generations of breeding, Specht and Williams (1984) noted little change in these secondary traits in MGs IV or

earlier cultivars, but did detect a decrease in mature plant height in MG III and an increase in MG 00. Similarly, Wilcox (2001) noted an increase of plant height in MG I but a decrease in II, III, and IV across generations. Ustun et al. (2001) reported an overall decrease in plant height when original PIs were compared to cultivars from the 1970's grown in the mid-southern USA.

Seed protein and oil concentration were found to be related to year of cultivar release by Voldeng et al. (1997), Wilcox (2001), and Ustun et al. (2001). Voldeng et al. (1997) reported that seed protein decreased 4 g kg^{-1} yr^{-1} whereas seed oil increased 4 g kg^{-1} yr^{-1} over generations. Across years of the MG 00-IV Uniform Soybean Test, Wilcox (2001) noted that seed protein concentration decreased significantly by 0.29 g kg^{-1} yr^{-1} in MG I and 0.27 g kg^{-1} yr^{-1} in MG II. In addition, a significant increase in seed oil concentration of 0.19 g kg^{-1} yr^{-1} in MG 00 and a decrease of 0.11 g kg^{-1} yr^{-1} in MG III were observed. Ustun et al. (2001) reported an overall decrease in seed protein and an increase in seed oil across generations within their historic set of eight cultivars adapted to the mid-south USA.

The objective of this study was to more comprehensively estimate annual genetic gain for seed yield by using a greater number of cultivars and a longer time frame of release than those used in past studies, with the purpose of updating and improving our understanding of genetic changes resulting from the past 80 years of North American soybean breeding in MG II, III, and IV. This objective was fulfilled by directly comparing the performance of soybean cultivars released from 1923 to 2008 in field tests in 17 USA states and one Canadian province.

MATERIALS AND METHODS

The experiments included PIs and publicly and proprietarily developed cultivars that were released or introduced between 1923 and 2008. Public cultivars and PIs were selected for inclusion in the MG II, III, or IV tests if they were widely grown by producers after their release. Testing all released cultivars was not possible, so for each MG-specific entry list, cultivars were first sorted by decade of release year so that the authors could construct a final entry list of cultivars whose release years were uniformly distributed (to the extent possible) within and across the eight decades of breeding. Seed of public cultivars were obtained from the USDA Soybean Germplasm Collection. Soybean cultivars from the proprietary sector (i.e., Monsanto,

Pioneer, and Syngenta) were nominated by the respective company breeders for inclusion in these tests, and seed of the cultivars was provided to the authors by the companies. Seed increases of all test entries were grown in 2009 to provide seed for 2010 trials and seed was produced in 2010 for the 2011 trials. Seed increase sites in both years were Lincoln, NE (MG II); Urbana, IL (MG III); and Columbia, MO (MG IV).

Cultivars in each MG-specific set were yield tested during 2010 and 2011 at 13 to 15 locations representative of typical production environments where cultivars of the given MG are grown. Table 4.1 lists the cultivar names and release years for the 60 MG II, 59 MG III, and 49 MG IV cultivars. The three MG-specific trials were separate experiments, and the experimental design at each site was a randomized complete block with between 2 and 12 replications (Table 4.2). Years and sites were not a balanced factorial, so the two were combined to be treated as a site-year environmental factor in the data analysis. The number of site-year environments in which data were collected for a given trait is shown in Table 3.

Cultivars were planted in yield plots four rows wide with 0.76 m row spacing and at a rate of approximately 30 seeds m^{-1} of row. Plot length varied depending on the planting system of the author-collaborator. Seed yield was estimated by harvesting the inner two rows of four row plots with a plot combine and adjusted to 130 g kg^{-1} moisture. Date of maturity was recorded as the day when at least 95% of the pods in a plot had attained a mature color (R8; Fehr et al., 1971). Plant height was recorded as the average distance from the soil surface to the apex of the main stem after R8. Lodging of the plants within a plot was rated after R8 using a scoring scale of 1 (all plants erect) to 5 (all plants prostrate). Seed mass was measured as the weight of a sample of 100 or 200 seeds, but expressed as a 100-seed weight value. Seed quality was visually examined and rated using a scoring scale of 1 (very good quality) to 5 (very poor quality), taking into account greenishness and the number of seeds with defective seed coats and moldy or rotten seeds, but not mechanical harvest damage. Seed protein and oil concentration was measured with near infrared reflectance and expressed on a 130 g kg^{-1} moisture basis.

Entries in the 2010 trials conducted at DeKalb, Perry, Monmouth, Urbana, Brownstown, and Dixon Springs, IL were a subset of 35 to 45 cultivars of the historic sets, but eight replications

were used in these trials (Fox et al., 2013). In addition, some cultivars were replicated more frequently than others in the trials conducted at Waseca2, MN in 2010, West Lafayette, IN in 2011, and Arlington, WI in both years, for the purpose of allowing the inclusion of more cultivars in studies aimed at examining how contrasts in agronomic practices (i.e., early vs. late planting) influence the magnitude of the genetic yield gain rate (Rowntree et al., 2013a).

To deal with the unbalanced nature of the data set, Proc Mixed of SAS was used, with covariance parameters estimated by the restricted maximum likelihood method, and the fixed effect denominator degrees of freedom estimated with the Kenward Rogers option (Kenward and Rogers 1997) (SAS Institute, 2011). Best linear unbiased predictors (BLUPs) of entries were calculated using estimate statements in a model, with environment, replications nested in environments, cultivar, and cultivar × environment interaction treated as random effects. Regression of cultivar trait values (BLUPs) on the respective year of cultivar release within each MG provided an annualized estimate of genetic change. Cultivar release dates were obtained from published plant cultivar registration articles, published pedigree lists (Bernard et al., 1988), or from the company providing seed. An overall analysis across MGs for yield was completed to test for significant differences in slopes or intercepts among each MG using Proc Mixed of SAS (SAS Institute, 2011). However, to account for maturity date variation within maturity groups, yield BLUPs were also calculated using maturity date as a covariate in the model. This overall analysis of yield data included the random factors of environment and replications nested in environments and also the fixed effects of MG, year of release, and MG × year of release interaction. In addition to linear model estimates of cultivar yield improvement, a two-segment linear regression model was also fit to the data and the two model fits were compared using the Akaike Information Criterion (AIC). The AIC rewards models with a better coefficient of determination (R^2) fit but penalize models that require more parameters to achieve the better fit (Sakamoto et al., 1986; Posada and Buckley, 2004). The parameters in the segmented linear regression model, as shown below,

Y1 = intercept1 + slope1*X

YatX0 = slope1*X0 + intercept1

Y2 = YatX0 + slope2*(X-X0)

Y = IF(X<X0, Y1, Y2)

were estimated using nonlinear regression curve-fitting with iteration to converge on parameter estimates that provided a best-fit to the data, with results then plotted using GraphPad Prism (GraphPad Software, 2013)

Yield stability (Finlay and Wilkinson, 1963) of cultivars was evaluated by obtaining environment and cultivar × environment BLUPs from the previously described random effects model without the maturity covariate. BLUPs of cultivar × environment combinations were regressed on the environment BLUPs to obtain slopes that describe the stability of each cultivar. Stability slope trends were described by the regression of each cultivar regression value with its year of release. To compare yield stability of old vs. new cultivars, the six oldest and six newest cultivars in MG II and MG III sets, and five such of each type in MG IV were chosen. Average BLUP yields for the old and new cultivars were regressed with the environment BLUPs. Calculation of BLUPs and regression coefficients were conducted in SAS Proc Mixed (SAS Institute, 2011).

RESULTS AND DISCUSSION

The weather was hot and dry in 2010 at the southernmost test sites of Perkins, OK; Lexington, KY; Dixon Springs, IL; Stuttgart, AR; and Suffolk, VA. In 2011, growing season conditions were generally better, although environments at Manhattan, KS, Stuttgart, AR, and Brownstown, IL, experienced hot and dry conditions that suppressed yields. During 2011, there was an early frost at Beresford, SD, and the Ontario locations of Harrow and Woodslee were planted late, but these events did not greatly affect yield. Overall, mean yields at test sites were typical for the regions where the sites were located. Average yields for environments ranged from 2123 to 4584 kg ha^{-1} in MG II, 1799 kg ha^{-1} to 4577 kg ha^{-1} in MG III, and 953 kg ha^{-1} to 4443 kg ha^{-1} in MG IV. The upper range value is approximately the same in all three MG trial sets, but the lower range values differ, particularly for MG IV. There were no yield trials with a mean yield less than 2000 kg ha^{-1} in MG II, but there was one in MG III, and five in MG IV.

Yield Improvement

Results from field tests revealed that seed yields consistently increased over the past 80 yr due to breeding efforts. Across environments, the estimated linear rate of genetic yield gain was

23 kg ha^{-1} yr^{-1} in both MG II and MG III, and 20 kg ha^{-1} yr^{-1} in MG IV (Fig. 4.1). A consideration in the analysis is that within each MG, the more recently released cultivars matured an average of 8 d later than older cultivars. Within yield tests, later-maturing cultivars tend to yield greater than earlier-maturing cultivars because a later maturity date allows plants to assimilate more carbon through an extended duration of photosynthesis, thereby resulting in greater seed yield than early maturing cultivars. After adjusting for maturity with the covariate analysis, the linear rates for yield gain were reduced to 20 kg ha^{-1} yr^{-1} in both MG II and MG III, and to 18 kg ha^{-1} yr^{-1} in MG IV. These estimates of annual genetic yield gain are greater than estimates provided by Luedders (1977) and Wilcox et al. (1979), but comparable with Specht and Williams' (1984) estimates of 29 kg ha^{-1} yr^{-1} in MG II, 17 kg ha^{-1} yr^{-1} in MG III, and 22 kg ha^{-1} yr^{-1} in MG IV. While the genetic yield gain estimates obtained in the present study do not reflect the entire diversity of soybean production areas of the United States, it is worth noting that these regions account for about 75% of the total U.S. soybean cropping area and production (USDA-NASS, 2013). It nonetheless appeared that genetic improvement (arising from the continual release of ever-higher yielding cultivars that are quickly adopted by producers) has been a key driver of on-farm yield improvement. Rates of genetic yield gain estimated in this study are concordant with the 23 kg ha^{-1} yr^{-1} rate of on-farm soybean yield improvement that has occurred during the same 80-yr period in the United States (Specht et al., 2014).

Additional data analyses indicated that the regression coefficients and y-intercepts for MG II and MG III were not significantly different, but that those for MG IV were significantly lower. Hypothesized reasons for the lower MG IV regression parameters were that (i) the five trials with yields < 2000 kg ha^{-1} among the 27 MG IV trials may have diminished the parameter estimates, or (ii) the MG IV cultivar releases may have intrinsically less yield potential than the MG II and III cultivars. To evaluate these hypotheses, regression parameters for the MG III and MG IV cultivar sets were reestimated using just the site-year trial yield data collected at 15 locations where both MG III and MG IV trials were grown. One such location, the nonirrigated Manhattan, KS 2011 site, was the lowest-yielding MG III site and was also a low-yielding MG IV site. Remaining locations with both MG III and MG IV tests comprised moderate- to high-yielding environments. No significant difference was detected between MG III and MG IV for rates of yield gain ($P = 0.35$) or intercept ($P = 0.34$). This suggests that the lower MG IV

intercept value in the overall analysis was not due to a lower yield potential of MG IV cultivars, but was instead a result of the less productive soils, drier conditions, and potentially greater biotic stresses encountered at the MG IV testing sites than at the MG II and MG III sites.

The fit of the yield data to either a simple linear or a two-segment linear model was compared to determine if the rate of yield improvement was constant or discontinuous over 80 yr of breeding. The segmented model was statistically more probable than the simple linear model based on the AIC values computed for each MG (Fig. 4.1). The bestfit breakpoint year was 1968 for MG II, 1964 for MG III, and 1971 for MG IV. The postbreakpoint regression coefficient across MGs averaged 29 kg ha^{-1} yr^{-1} and was 2.6 times greater than the average prebreakpoint coefficient of 11 kg ha^{-1} yr^{-1}. Specht et al. (2014) reported that a two-segment linear model also provided the best fit to on-farm U.S. yield data from 1924 to 2012, and noted that yield gain rate increased 50% after the best-fit breakpoint year of 1983.

These data clearly demonstrate that annual genetic gain for seed yield is higher now than in the past, but the large confidence intervals for the breakpoints make it difficult to be specific about when the change occurred or the reasons for the change. There were fewer entries released between 1920 and 1950 than between 1980 and 2010, so the estimate of yield for those early years is not as precise as for later years. In both MGs II and III, there were entries from the 1920s that yielded as much as the entries from the 1960s and these few entries affected the prebreakpoint rate of gain. The improved rate of gain that is sustained post 1970 is likely to be a result of increase in investments by commercial companies in soybean breeding that were stimulated by the passage of the 1970 PVP Act, and also by the increase in the number of public sector soybean breeders. No changes in rate of genetic gain from major scientific and technical advances of the past 15 yr were observed; however, it may be too soon to detect a change if it has occurred. In addition, these results provide no evidence of a yield plateau in any of the three MGs.

Yield Stability

A stability analysis was conducted to evaluate the yield response of historic sets of cultivars when grown in environments of varying productivity to determine whether yield stability of

cultivars has changed over generations of breeding. Stability regression coefficients of cultivars ranged from 0.69 to 1.21 in MG II, 0.80 to 1.26 in MG III, and 0.74 to 1.23 in MG IV (Fig. 4.2). A cultivar stability regression coefficient (b) of 1.0 indicates that the response of a cultivar to different environments is the same as the mean response of all other cultivars in the experiment (Bernardo, 2010a). A cultivar classified as having a high b value is considered less stable than a cultivar with a low b value. The high b value could indicate poor performance in stressful environments, or conversely, it could also indicate that a cultivar has a better ability to take advantage of favorable environments compared with a cultivar with a lower coefficient. When stability coefficients of cultivars were regressed on year of release (Fig. 4.2), significant positive regression coefficients were detected, which were 0.005 yr^{-1} ($P < 0.0001$) for MG II and MG IV and 0.002 yr^{-1} ($P = 0.0003$) for MG III. This finding indicates that new cultivars are less stable (but alternatively, more responsive) than old cultivars, relative to the yield mean of all cultivars in each test environment. Voldeng et al. (1997), Wilcox et al. (1979), and De Bruin and Pedersen (2008) reported that yield stability did not change over generations of breeding, but statistical power in those studies was far less than that in the present study in terms of cultivar numbers and test sites.

A group of new and old cultivars from each MG were compared to determine whether increases in stability coefficients (reductions in stability) over time are the result of newer cultivars having reduced performance in stressful environments or the result of newer cultivars being better able take advantage of favorable environments than older cultivars. We found that new cultivars had greater yields in both low and high yielding environments compared with old cultivars (Fig. 4.3; $P < 0.0001$). In addition, yield stability coefficients of the new cultivar group were found to be significantly greater than the old cultivar group in all three MGs ($P < 0.0001$; Fig. 4.3). These results show that, despite the reduced stability of modern cultivars, the modern cultivars yielded better on average than old cultivars across all environments tested, and that modern cultivars can take advantage of high-yielding environments better than old cultivars. Araus et al. (2002) had a similar finding in wheat (Triticum aestivum L.) and documented that new wheat cultivars were more yield responsive than were old cultivars to more productive test environments.

The relationship between genetic gain rates and location mean yields is shown in Fig. 4.4. The graphs clearly indicate that measured rate of genetic yield improvement is functionally dependent on productivity of the test environment in which it is measured. Because breeding leads to improvement in genetic yield potential over time, and because high-yielding environments allow more of the genetic yield potential to be expressed, it was not surprising that there is an increase in the rate of genetic gain for yield as the environmental yield potential increases. This finding is indicative of an interaction of genetic improvement with environmental productivity improvement, wherein new cultivars synergistically act with modern agronomic practices to improve yield more than would be expected based on the additive effects of the genetic (old vs. new cultivar yield potential) factor and the agronomic (low vs. high productivity) factor. In that regard, it would be of interest in the future to use these historic cultivar sets to evaluate the rate of genetic yield gain in environments whose productivity exceeds 5000 kg ha^{-1}.

Agronomic Traits

Within each MG tested, the date when cultivars reached maturity increased linearly at a rate of 0.09 to 0.10 d yr^{-1} (Table 4.4, Fig. 4.5). Although this increase is small from year to year, it amounts to approximately 1 d per decade, and over the course of the release years in this study, maturities of new cultivars are about 1 wk later than those of cultivars from the 1950s. This change occurred despite the inclusion of maturity checks used to bracket maturities within each MG of the Uniform Soybean Tests, in which public sector experimental lines are yield-tested before they are released (Crochet and Hughes, 2012). Over time, however, these maturity checks are replaced as new cultivars are released, resulting in later maturity dates for cultivars within a MG. The gradual change towards later maturity is not surprising, because there is a tendency for later-maturing genotypes to have a yield advantage over earlier-maturing lines. Later maturity of recently released cultivars compared with old cultivars also occurred for cultivars developed by proprietary sector breeders, most likely for the same reasons. It can be argued that later-maturing, recently released cultivars within a given MG do offer a better adaptive fit to growers using earlier planting dates (Rowntree et al., 2013), particularly because climate change has resulted in a longer growing season, as is evident in the shifting of USDA plant zones (Kaplan, 2012).

Lodging scores have decreased in newly released cultivars compared with older cultivars or PIs. A decrease of 0.012 to 0.018 units yr^{-1} was found, which translates to a total decrease of about 1 to 1.5 units of lodging over the 80 yr of cultivar releases (Table 4.4). Previous research also has shown a reduction in lodging (Luedders, 1977; Wilcox et al., 1979; Specht and Williams, 1984; Voldeng et al., 1997; and Wilcox, 2001). Lodging is an important trait in selection, as severe lodging can result in harvest losses and growers have a strong desire for cultivars with reduced lodging.

Mature plant height significantly decreased across generations of breeding, with decreases by MG ranging from 0.13 to 0.21 cm yr^{-1} (Table 4.4). Tall cultivars tend to lodge more than short cultivars, and therefore as breeders have selected for reduced lodging, there has been a trend for cultivars to become shorter over time. Specht and Williams (1984) and Wilcox (2001) also reported a decrease of mature plant height over generations of breeding. Even though these decreases are significant, height data are much more variable than seed yield data. Short and tall cultivars existed early in soybean breeding and still do today. Furthermore, this variability results from tall cultivars released in the 1920s or short ones released in the 2000s having a large influence on the overall regression line. For example, the semideterminate cultivar NE3001 (University of Nebraska, Lincoln), which was released in 2004, had a mean height across environments of 61 cm, compared with the mean height of 97 cm averaged across the other MG III cultivars that were released after 2000. When NE3001 was excluded from the MG III data set, the rate of decrease lessened from a rate of 0.21 cm yr^{-1} to 0.07 cm yr^{-1}. Plant height is not necessarily a trait of primary importance in selection aimed at improving yield and lodging, and both tall and short cultivars meeting the yield and lodging selection criteria are often released.

Genetic changes in seed protein and oil concentrations are important to the soybean processing industry in terms of the extractable percentage of meal and oil obtainable from a given volume of soybean seed. Seed protein concentration decreased at a rate of 0.22 g kg^{-1} yr^{-1} in the MG II and MG III sets, and 0.16 g kg^{-1} yr^{-1} in the MG IV set (Table 4.4, Fig. 4.6). Concordantly but inversely, oil concentration increased at a rate of 0.14 g kg^{-1} yr^{-1} in MG II, 0.10 g kg^{-1} yr^{-1} in MG III, and 0.05 g $kg^{-1} yr^{-1}$ in MG IV (Table 4.4, Fig. 4.6). These rates were significant at the

0.05 significance level, with the exception of the low rate of oil increase in MG IV (Table 4.4). Similar results were reported by Wilcox et al. (1979) and Voldeng et al. (1997). On the basis of annual Uniform Soybean Test data, Wilcox (2001) also noted similar findings in MG II, but reported a decrease in seed oil concentrations in MG III. Wilson (2004) described the difficulty of breeding for higher protein, oil, and yield in soybean germplasm due to the negative correlation between protein and yield, protein and oil, and the positive correlation between oil and yield. Because soybean growers in the United States are compensated on the basis of seed weight sold and not seed composition, increasing yield remains the primary goal of breeding programs. Reduction of protein concentration is, therefore, the consequence of this focus on yield and the negative correlation frequently observed between yield and protein (Burton, 1987). Unless the market for soybean changes to include compensation to growers for seed composition, it is likely that the trend of reduced seed protein and greater oil concentrations will continue. Still, increases in seed yield result in an overall increased total mass of protein produced per hectare (Table 4.4), which in this study averaged between 6.6 and 7.4 kg ha^{-1} yr^{-1} in the three MG sets (Table 4.4). Concordantly, the total mass of oil produced per hectare also rose on the scale of 4.4 to 4.7 kg ha^{-1} yr^{-1} in the three MGs. In essence, there is a genetic gain not only for seed yield, but also for seed protein and oil yields per se over the last 80 yr of breeding.

Data collected for 100-seed weight and seed quality were the most variable of the traits measured. No significant change over time in 100-seed weight was detected in any MG (Table 4.4). This suggests that the significant yield increase observed over time in this study must result from new cultivars producing more seed per unit area compared with old cultivars. With regard to seed quality (Table 4.4), breeding efforts in the MG II and IV cultivar sets have resulted in slight, but significant, decreases in seed quality scores, indicating improved seed quality. Unfortunately, insufficient data from the MG III set prevented evaluation of the significance of estimated genetic change in seed quality for this MG. Although a small reduction in seed quality score suggests that new MG II and IV cultivars produce seed that may be slightly more appealing, that finding may be of questionable significance, given the difficulty of inferring a biological basis for such an effect. Previous research by Specht and Williams (1984) found an increase in seed size and a decrease in seed quality over generations for breeding soybean. Small or nonsignificant changes in these traits were expected because neither of these traits is a

selection target in the development of commodity type cultivars with yield potential as the primary focus.

Estimates of genetic gains in other North American crops are comparable with the estimates in soybean observed in the current study. Battenfield et al. (2013) estimated the rate of winter wheat yield gain grown in the Great Plains region of the United States to be 15 kg ha^{-1} yr^{-1} when compared with a tall cultivar released in 1919, and 11 kg ha^{-1} yr^{-1} when compared with a semi-dwarf cultivar released in 1971. In the eastern United States, Green et al. (2012) estimated yield gains from 1919 to 2009 of soft red winter wheat to be 37 kg ha^{-1} yr^{-1}. Dry bean (Phaseolus vulgaris L.) breeding in the western United States has faced many challenges, with broad breeding objectives such as disease resistance and growth types. However, yield improvements made in the great northern, pink, pinto, and red bean market classes were estimated to range from 3% over a 16-yr period to 35% over a 43-yr period (Singh et al., 2007). In contrast, maize (Zea mays L.) has experienced tremendous improvements to yield potential. Duvick et al. (2004) estimated improvements of hybrids grown from 1930 to 2001 to be 77 kg ha^{-1} yr^{-1} when grown at their optimum density. However, Specht et al. (2014) noted that when on-farm irrigated corn and soybean yield improvement is examined synchronously in terms of an annual corn–soybean yield ratio, that ratio has been a 3:1 constant for the past 42 yr. Though the physiological differences between the two crop species account for soybean yielding about three-fold less than corn in an absolute sense, the relative pace of on-farm corn and soybean yield improvement has been effectively and remarkably equivalent.

CONCLUSIONS

Plant breeders are continually making selections to improve cultivars that producers grow. Over the past 80 yr, soybean seed yield has increased considerably, and this experiment produced estimates that the genetic improvement of seed yield has totaled 1450 to 1578 kg ha^{-1} for the MGs studied, after adjustments for the date of maturity. In an analysis that included comparisons of on-farm yield gains with genetic gains in the northern United States, Specht et al. (2014), estimated that two-thirds of the on farm yield gains were the result of genetic improvements and that one-third was the result of agronomic improvements. However, there also is evidence of synergistic interactions between genetic and agronomic improvements (Rowntree et al., 2013).

Efforts of soybean breeders in both public and proprietary programs have developed soybean into a leading North American crop for domestic and export markets and increasing the rate of yield gain is important for meeting ever-greater worldwide demand for food, feed, and fuel. Further improvements will continue to face challenges of limited genetic diversity and biotic and abiotic stresses (Gizlice et al., 1994; Sneller, 1994; Betzelberger et al., 2012). However, breeders will have new tools available such as technology to collect phenotypic data, DNA and RNA sequences, molecular markers, mapping studies, and methodologies to aid in yield improvement (Bernardo, 2010b; Furbank and Tester, 2011; Poland and Rife, 2012; Schmutz et al., 2010).

LITERATURE CITED

Bernard, R.L., G.A. Juvik, E.E. Hartwig, and C.J. Edwards, Jr. 1988. Origins and pedigrees of public soybean varieties in the United States and Canada. USDA Technical Bulletin No. 1746. 68 p.

Bernardo, R. 2010. Genomewide selection with minimal crossing in self-pollinated crops. Crop Sci. 50:614-627.

Betzelberger, A.M., C.R. Yendrek, J. Sun, C.P. Leisner, R.L. Nelson, D.R. Ort, and E.A. Ainsworth. 2012. Ozone exposure response for U.S. soybean cultivars: linear reductions in photosynthetic potential, biomass, and yield. Plant Physiol. 160:1827-1839.

Boerma, H.R. 1979. Comparison of past and recently developed soybean cultivars in maturity groups VI, VII, and VIII. Crop Sci. 19:611–613.

Carter, T.E., R.L. Nelson, C. H. Sneller, and Z. Cui. 2004. Genetic diversity in soybean. *In:* H. R. Boerma and J. E. Specht, editors, Soybeans: improvement, production, and uses. 3rd ed. Agron. Monogr. 16. ASA, CSSA, and SSSA, Madison, WI. p. 303-416.

De Bruin, J. L. and P. Pedersen. 2008. Yield improvement and stability for soybean cultivars with resistance to *Heterodera glycines* Ichinohe. Agron. J. 100:1354-1359.

Eathington S.R., T.M. Crosbie, M.D. Edwards, R.S. Reiter, and J.K. Bull. 2007. Molecular markers in a commercial breeding program. Crop Sci. 47:S154-S163.

Fehr, W.R., editor 1991. Principles of cultivar development: theory and technique. Vol.1 Macmillan Publishing Co., New York.

Fehr, W., C. Caviness, D. Burmood, and J.S. Pennington. 1971. Stage of development descriptions for soybeans, *Glycine max* (L). Merrill. Crop Sci. 11:929–931.

Finlay, K.W., and G.N. Wilkinson. 1963. The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. 14:742-754.

Fox, C. M., T.R. Cary, A.L. Colgrove, E.D. Nafziger, J.S. Haudenshield, G.L. Hartman, J.E. Specht, and B.W. Diers. 2013. Estimating soybean genetic gain for yield in the Northern United States – influence of cropping history. Crop Sci. Accepted for publication.

Gizlice, Z., T.E. Carter, Jr., and J.W. Burton. 1994. Genetic base for North America public soybean cultivars released between 1947 and 1988. Crop Sci. 34:1143-1151.

GraphPad Software, 2013. GraphPad Prism for Windows. Vers. 6.02. GraphPad Software, La Jolla, CA.

Hartwig, E.E. 1973. Varietal development. *In* B. E. Caldwell (ed.) Soybeans: Improvement, production, and uses. 1st ed. Agron. Monogr. 16. ASA, CSSA, and SSSA, Madison, WI. p. 187-210.

Hassol, S.J. 2009. Global climate change impacts in the United States – 2009 Report. U.S. Global Change Research Program. Publication available on-line at: https://nca2009.globalchange.gov/agriculture/index.html (accessed 9 Mar. 2016)

Hefner, E.L., M.E. Sorrels, and J.L. Jannink. 2009. Genomic selection for crop improvement. Crop Sci. 49:1-12.

Kenward, M. and J. Roger. 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983-997.

Long, S.P., E.A. Ainsworth, A.D.B. Leakey, J. Nosberger, and D.R. Ort. 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. Science 312:1918-1921.

Luedders,V.D. 1977. Genetic improvement in yield of soybeans. Crop Sci. 17:971–972.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Poland, J.A. and T.W. Rife. 2012. Genotyping-by-sequencing for plant breeding and genetics. Plant Gen. 5:92-102.

Posada, D. and T.R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53:793-808.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield trends are insufficient to double crop production by 2050. PLoS ONE 8:doi.10.1371/journal.pone.0066428.

Rowntree, S.C., J.J. Suhre, N.H. Weidenbenner, E.W. Wilson, V.M. Davis, S.L. Naeve, S.N. Casteel, B.W. Diers, P.D. Esker, J.E. Specht, and S.P. Conley. 2013a. Genetic gain × management interactions in soybean: I. Planting date. Crop Sci. 53:1128:1138.

Rowntree, S., J. Suhre, E. Wilson, V. Davis, B. Diers, S. Casteel, P. Esker, and S. Conley. 2013b. Physiological and phenological responses of historical soybean cultivar releases to earlier planting. Crop Sci. Accepted for publication.

Sakamoto, Y., M. Ishiguro, and G. Kitagawa. 1986. Akaike information criterion statistics. KTK Scientific Publishers, Tokyo, Japan.

Salado-Navarro, L.R., T.R. Sinclair, and K. Hinson. 1993. Change in yield and seed growth traits in soybean cultivars released in the southern USA from 1945 to 1983. Crop Sci. 33:1204–1209.

SAS Institute. 2011. The SAS system for Microsoft Windows. Release 9.3. SAS Inst., Cary, NC.

Schmutz, J., S. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D. Hyten, Q.J. Song, J.J. Thelen, J. Cheng, D. Xu, U. Hellsten, G. May, Y. Yu, T. Sakurai, T. Umezawa, M. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P.B. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, and S.A. Jackson. 2010. A genome sequence of paleopolyploid soybean (*Glycine max*). Nature 463:178–183.

Sinclair, T.R., C.P. Tanner, and J.M. Bennet. 1984. Water use efficiency in crop production. BioScience 34:36-40.

Sneller, C.H. 1994. Pedigree analysis of elite soybean lines. Crop Sci. 34:1515-1522.

Specht, J.E., B.W. Diers, R.L. Nelson, J.F. Toledo, J.A. Torrion, and P. Grassini. 2014. Soybean. *In*:, J.S.C. Smith, B. Diers, B. Carver, and J. Specht, editors, Yield gains in major U.S. field crops. CSSA Spec. Publ. 33. ASA, CSSA, and SSSA, Madison, WI. p. 311-356.

Specht, J.E., D.J. Hume, and S.V. Kumudini. 1999. Soybean yield potential–a genetic and physiological perspective. Crop Sci. 39:1560-1570.

Specht, J.E. and J.H. Williams. 1984. Contribution of genetic technology to soybean productivity—Retrospect and prospect. *In:* W.R. Fehr, editor, Genetic contribution to yield gains of five major crop plants. CSSA Special Publ. 7. CSSA and ASA, Madison, WI. p. 49-74.

USDA. 2006. Plant Variety Protection Act and Regulations and Rules of Practice. [Online] Available at http://www.ams.usda.gov/sites/default/files/media/Plant Variety Protection Act.pdf, (accessed 9 Mar. 2016). Agriculture Marketing Service.

USDA-National Agricultural Statistics Service (USDA-NASS). 2015. Data and statistics. USDA, NASS, Washington, DC. http://www.nass.usda.gov/Data_and_Statistics (accessed 9 Mar. 2016).

Ustun, A., F.L. Allen, and B.C. English. 2001. Genetic progress in soybean of the U.S. Midsouth. Crop Sci. 41:993-998.

Voldeng, H.D., E.R. Cober, D.J. Hume, C. Gillard, and M.J. Morrison. 1997. Fifty-eight years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 37:428–431.

Wilcox, J.R. 2001. Sixty years of improvement in publicly developed elite soybean lines. Crop Sci. 41:1711-1716.

Wilcox, J.R., W.T. Schapaugh, Jr., R.L. Bernard, W.R. Fehr, and M.H. Niehaus. 1979. Genetic improvement of soybeans in the Midwest. Crop Sci. 19:803–805.

Wilson, R.F. 2008. Soybean: market driven research needs. *In:* G. Stacey, editor, Genetics and Genomics of Soybean. Springer Science+Business Media, LLC, New York, NY. p. 3-15.

Ziska, L.H. and J.A. Bunce. 2007. Predicting the impact of changing CO_2 on crop yields: some thoughts on food. New Phytol. 175:607-618.

TABLES AND FIGURES

Table 4.1. Cultivar entries and year of release from each maturity group (MG). Proprietary cultivars are not identified by name (as per an MTA agreement between the authors and company officials that allowed use of those cultivars in this 2-yr experiment).

MG II		MG III		MG IV	
Cultivar	Year of release	Cultivar	Year of release	Cultivar	Year of release
Korean	1928	Dunfield	1923	Macoupin	1930
Mukden	1932	Illini	1927	Scioto	1933
Richland	1938	A.K. (Harrow)	1928	Boone	1935
Hawkeye	1947	Mandell	1934	Chief	1940
Harosoy	1951	Mingo	1940	Patoka	1940
Lindarin	1958	Lincoln	1943	Gibson	1942
Harosoy 63	1963	Adams	1948	Wabash	1948
Hawkeye 63	1963	Shelby	1958	Perry	1952
Amsoy	1965	Ford	1958	Clark	1953
Corsoy	1967	Ross	1960	Clark 63	1963
Beeson	1968	Wayne	1964	Cutler	1968
Amsoy 71	1970	Adelphia	1964	Bonus	1971
Wells	1972	Calland	1968	Franklin	1977
Harcor	1975	Williams	1971	Union	1977
Wells II	1978	Woodworth	1974	Douglas	1980
Vickery	1978	Cumberland	1978	Lawrence	1981
Corsoy 79	1979	Oakland	1978	Sparks	1981
Beeson 80	1979	Pella	1979	Morgan	1986
Century	1979	Williams 82	1981	Spencer	1988
Amcor	1979	Zane	1984	Flyer	1988
Century 84	1984	Harper	1984	Corsica	1991
Elgin	1984	Chamberlain	1986	KS4694	1993
Preston	1985	Resnik	1987	Stressland	1994
Burlison	1988	Pella 86	1987	Cisne	1995
Elgin 87	1988	Dunbar	1992	Mustang	1995
Conrad	1988	Thorne	1992	Omaha	1996
Jack	1989	Macon	1995	LS93-0375	2001
Kenwood	1989	IA 3004	1995	LN97-15076	2003
RCAT Angora	1991	Maverick	1996	LD00-3309	2005
IA 2021	1995	Pana	1997	P-ID-4- 1	1985
IA 2022	1995	IA 3010	1998	P-ID-4- 2	1989
Savoy	1996	U98-311442	2001	P-ID-4- 3	1992
Dwight	1997	IA 3014	2001	P-ID-4- 4	2001

Table 4.1. (cont.)

MG II		MG III		MG IV	
Cultivar	Year of release	Cultivar	Year of release	Cultivar	Year of release
IA 2038	1998	IA 3023	2003	P-ID-4- 5	2004
IA 2050	2000	NE3001	2004	P-ID-4- 6	1980
IA 2052	2000	IA 3024	2004	P-ID-4- 7	1980
Loda	2001	P-ID-3- 1	1978	P-ID-4- 8	1990
IA 2068	2003	P-ID-3- 2	1986	P-ID-4- 9	1990
IA 2065	2005	P-ID-3- 3	1991	P-ID-4-10	2000
IA 2094	2006	P-ID-3- 4	1996	P-ID-4-11	2000
P-ID-2- 1	1989	P-ID-3- 5	1997	P-ID-4-12	1973
P-ID-2- 2	1990	P-ID-3- 6	1998	P-ID-4-13	1984
P-ID-2- 3	2004	P-ID-3- 7	1999	P-ID-4-14	1992
P-ID-2- 4	2001	P-ID-3- 8	2002	P-ID-4-15	1993
P-ID-2- 5	1993	P-ID-3- 9	1989	P-ID-4-16	1994
P-ID-2- 6	1991	P-ID-3-10	1990	P-ID-4-17	1996
P-ID-2- 7	1977	P-ID-3-11	1996	P-ID-4-18	1997
P-ID-2- 8	1977	P-ID-3-12	1997	P-ID-4-19	2006
P-ID-2- 9	1988	P-ID-3-13	2004	P-ID-4-20	2008
P-ID-2-10	1994	P-ID-3-14	2007		
P-ID-2-11	1982	P-ID-3-15	1983		
P-ID-2-12	1996	P-ID-3-16	1991		
P-ID-2-13	2008	P-ID-3-17	1992		
P-ID-2-14	2008	P-ID-3-18	1993		
P-ID-2-15	1985	P-ID-3-19	1994		
P-ID-2-16	1994	P-ID-3-20	2000		
P-ID-2-17	2001	P-ID-3-21	2001		
P-ID-2-18	1997	P-ID-3-22	2006		
P-ID-2-19	2005	P-ID-3-23	2006		
P-ID-2-20	2005				

Table 4.2. Number of replications in the maturity group (MG)-specific field trials conducted at each listed location and year.

	MG II		MG III		MG IV	
Location	2010	2011	2010	2011	2010	2011
Lamberton, MN	4	4				
Waseca, MN	4	4				
Waseca2, MN	6 †					
Beresford, SD		3				
Arlington, WI	12 †	12 †				
Ingram, MI		4				
Harrow, ON	3	3				
Woodslee, ON	3	3				
Hoytville, OH	4	4				
Nevada, IA	3					
Finch, IA	3					
Story City, IA		3				
Boone, IA		3				
Dekalb, IL	8 ‡	3				
Monmouth, IL	8 ‡					
Arthur, IL				3		
Perry, IL			8 ‡			
Urbana, IL		3	8 ‡	3		3
Wooster, OH			3	3		
West Lafayette, IN			4	2 †		
Muscatine, IA §			3	3		
Ames, IA			3			
Crawfordsville, IA				3		
Mead, NE §	4		4		4	
Mead, NE	4		4		4	
Clay Center, NE §		4		4		4
Havelock, NE		4		4		4
Manhatton, KS §			4	4	4	4
Manhatton, KS			4	4	4	4
Novelty, MO			2	2	3	2
Columbia, MO			3	3	2	3
Albany, MO				2		2
Harrisburg, IL				2		2
Belleville, IL						2
Brownstown, IL					8 ‡	3
Dixon Springs, IL					8 ‡	

Table 4.2. (cont.)

Location	MG II 2010	MG II 2011	MG III 2010	MG III 2011	MG IV 2010	MG IV 2011
Stuttgart, AR					3	3
Suffolk, VA					4	
Lexington, KY					4	3
Queenstown, MD					3	3
Perkins, OK					4	

† Cultivars not equally replicated.

‡ Subsets of 35 to 45 cultivars of the MG II and III historic sets were grown with four reps following 11 years continuous corn and four reps following a corn soybean rotation (Fox et al., 2014).

§ Irrigated location.

Table 4.3. Number of site-year environments that data were collected for yield, other agronomic traits, and seed composition for maturity group (MG) II, III and IV historic cultivar sets.

Trait	MG II	MG III	MG IV
Yield	27	26	27
Date of maturity	22	22	20
Plant height	24	20	21
Lodging	27	22	20
100 seed weight	15	11	9
Seed quality	9	5 †	3
Protein	14	8	6
Oil	14	8	6

†All reported data is from a single replication of the tests.

Table 4.4. Estimated rates of genetic change for the listed trait based on simple linear regression of cultivar trait BLUP value on cultivar release year within maturity groups (MG) II, III, and IV.

MG	n †	*b*	± SE	R^2	n	*b*	± SE	R^2	n	*b*	± SE	R^2
	Yield				Yield adjusted for maturity covariate				Date of maturity			
	------kg ha^{-1} yr^{-1}------				------kg ha^{-1} yr^{-1}------				-------days yr^{-1}--------			
II	6001	23.1 ***	1.295	0.845	4644	19.6 ***	1.307	0.794	4668	0.102 ***	0.017	0.369
III	4991	22.8 ***	1.280	0.847	4289	19.7 ***	1.258	0.812	4324	0.090 ***	0.018	0.303
IV	4451	19.5 ***	1.233	0.845	3474	18.1 ***	1.301	0.805	3577	0.090 ***	0.020	0.294
	Lodging				Plant height				Seed protein			
	-----score yr^{-1} ‡ ------				---------cm yr^{-1}---------				-----g kg^{-1} yr^{-1} § -----			
II	5904	-0.012 ***	0.002	0.413	5006	-0.133 *	0.052	0.102	2633	-0.222 ***	0.061	0.189
III	4495	-0.018 ***	0.002	0.667	4195	-0.205 ***	0.053	0.207	1416	-0.221 ***	0.051	0.251
IV	3582	-0.014 ***	0.001	0.666	3631	-0.129 **	0.042	0.167	1072	-0.159 *	0.068	0.105
	Seed oil				Protein produced				Oil produced			
	-----g kg^{-1} yr^{-1} § ------				------kg ha^{-1} yr^{-1} § ------				----kg ha^{-1} yr^{-1} § -----			
II	2633	0.136 ***	0.032	0.235	2633	6.594 ***	0.482	0.763	2633	4.373 ***	0.271	0.818
III	1416	0.103 **	0.030	0.170	1413	6.856 ***	0.469	0.789	1413	4.720 ***	0.337	0.774
IV	1072	0.051 NS	0.031	0.055	1071	7.378 ***	0.528	0.806	1071	4.369 ***	0.347	0.771
	100-seed weight				Seed quality							
	---------g yr^{-1}---------				------score yr^{-1} ¶ ------							
II	2622	-0.019 NS	0.010	0.063	1308	-0.0006 **	0.0002	0.1626				
III	1595	0.010 NS	0.008	0.029	295	#						
IV	1415	0.002 NS	0.008	0.002	490	-0.0008 *	0.0003	0.1232				

*, **, *** Significantly different from zero at 0.05, 0.01, and 0.001 probability level.

† The total number of observations (i.e., replicates, sites, years, cultivars) included in BLUP analyses.

‡ Lodging is scored visually from 1 (all plans erect) to 5 (all plants prostrate).

§ Protein and oil units are expressed on a 130 g kg^{-1} moisture basis.

¶ Seed quality is scored visually from 1 (very good) to 5 (very poor).

Insufficient data collected to calculate a trend.

Figure 4.1. Scatter plot of cultivar yield vs. cultivar release year for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model and a two-segment linear (Seg Lin) model were fit to the data, and the best-fit trend line statistics for each model are presented in the text box. The dotted lines are the 95% confidence intervals (CIs) for the linear fit. The vertical line denotes the estimated breakpoint year generated with the two-segment linear fit. Comparison of the Akaike information criterion (AIC) values computed for each model was used to determine the most probable model for the given data. Each cultivar data point is the mean of 27 (MGs II and IV) or 26 (MG III) site-years.

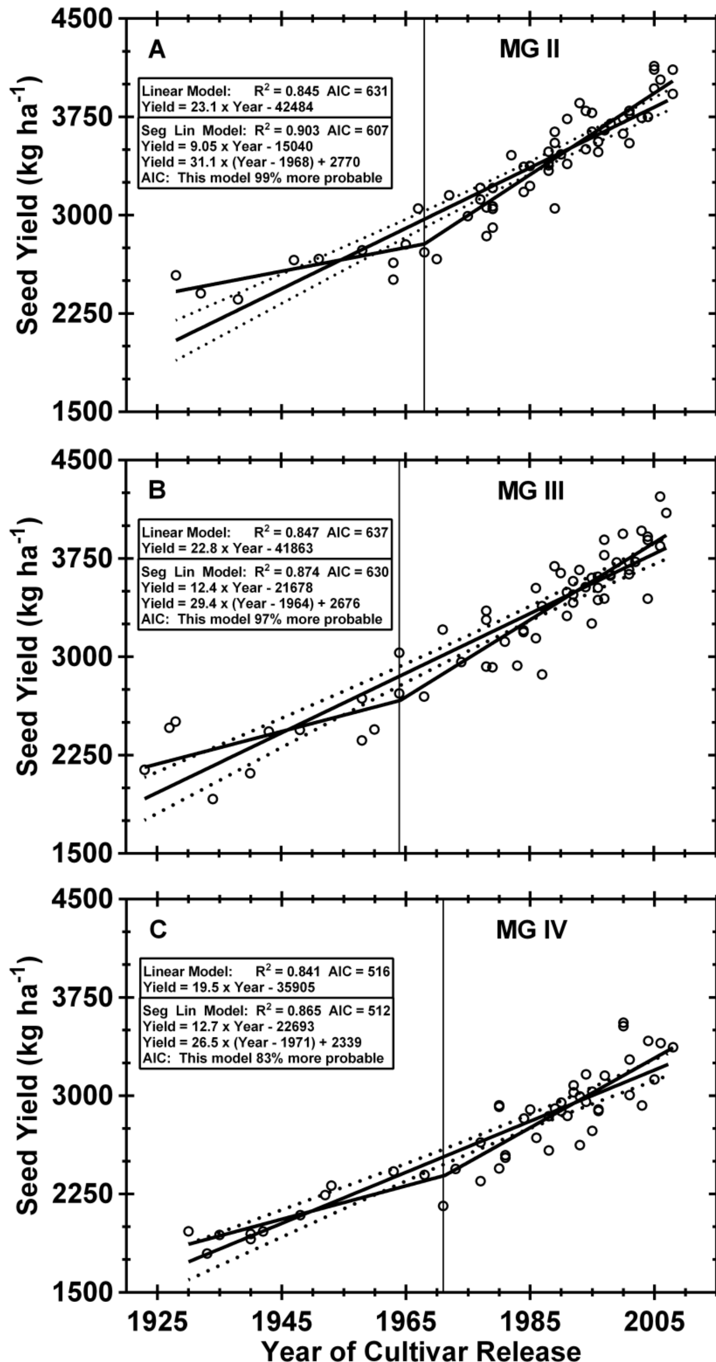


Figure 4.2. Scatter plot of cultivar yield stability coefficients vs. cultivar release year for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. Dotted lines are the 95% confidence intervals (CIs) for the linear fit. Each cultivar data point is the mean of 27 (MGs II and IV) or 26 (MG III) site-years.

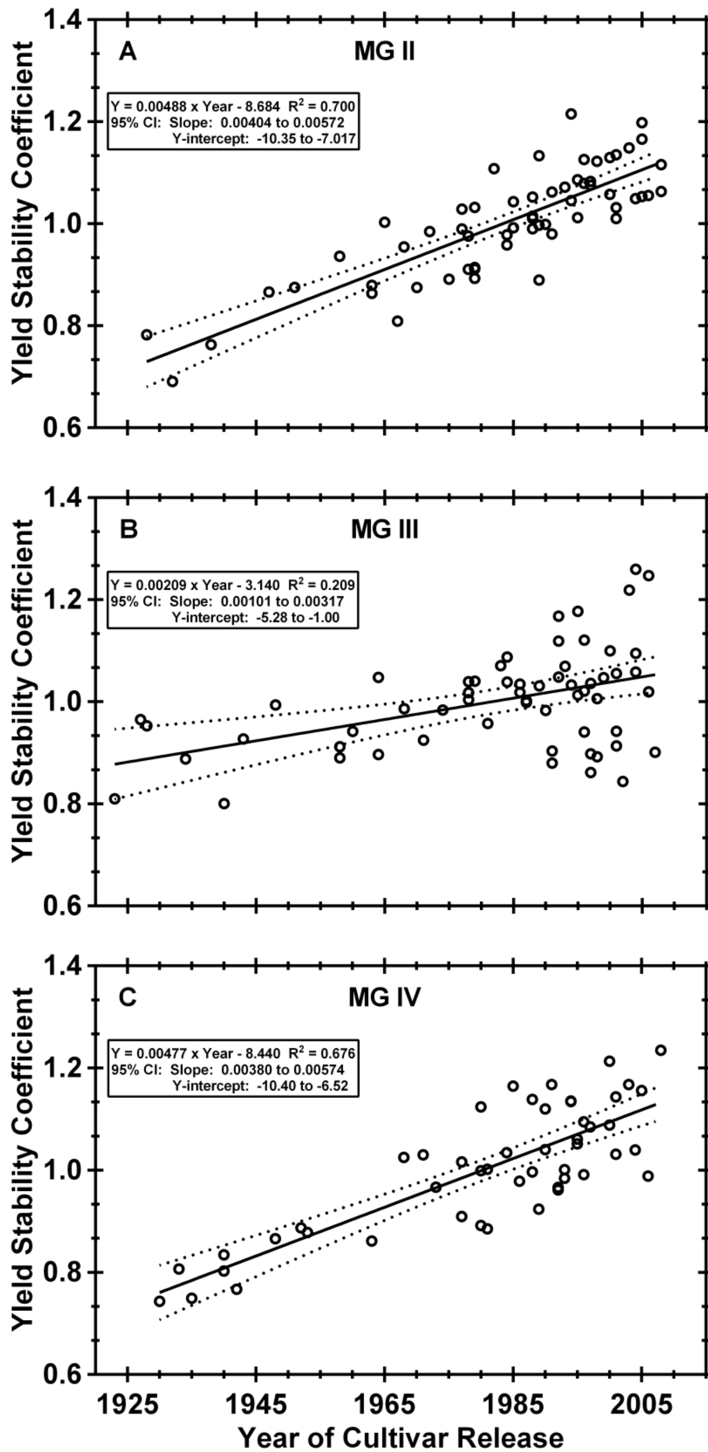


Figure 4.3. Scatter plot of cultivar yield vs. the respective mean yield of site-year trial used to derive cultivar yield estimates for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. The two sets of cultivars shown here represent cultivars selected from the decile extremes (6 or 5 oldest releases vs. 6 or 5 newest releases) in the time-span distribution of all cultivars (60, 59, 49) in respective MG sets (II, III, IV). Dotted lines are the 95% confidence intervals (CIs) for the linear fits in the oldest or newest groups. Each cultivar data point is the mean of 2 to 12 replications at that site-year.

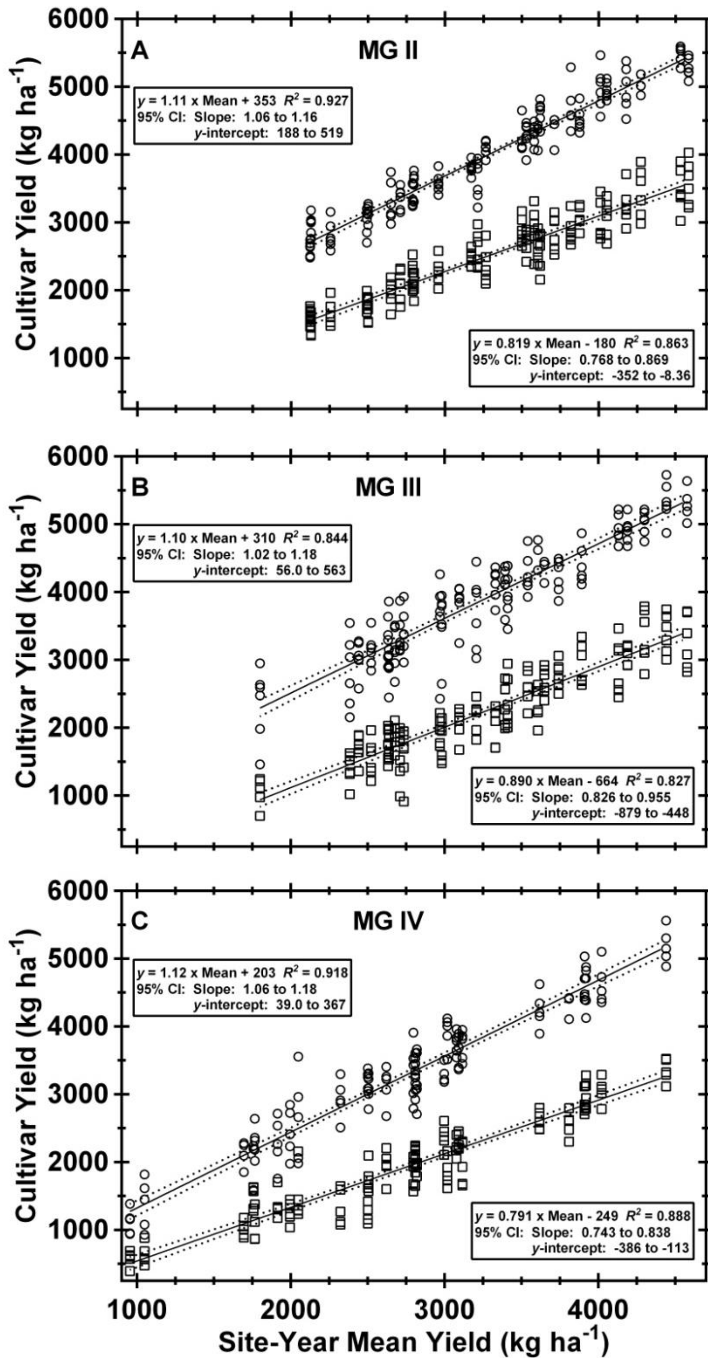


Figure 4.4. Scatter plot of genetic yield gain estimates vs. the respective means of site-years used to derive genetic yield gain estimates for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. Dotted lines are the 95% confidence intervals (CIs) for the linear fit. Each genetic gain estimate is taken across all cultivars at that site-year.

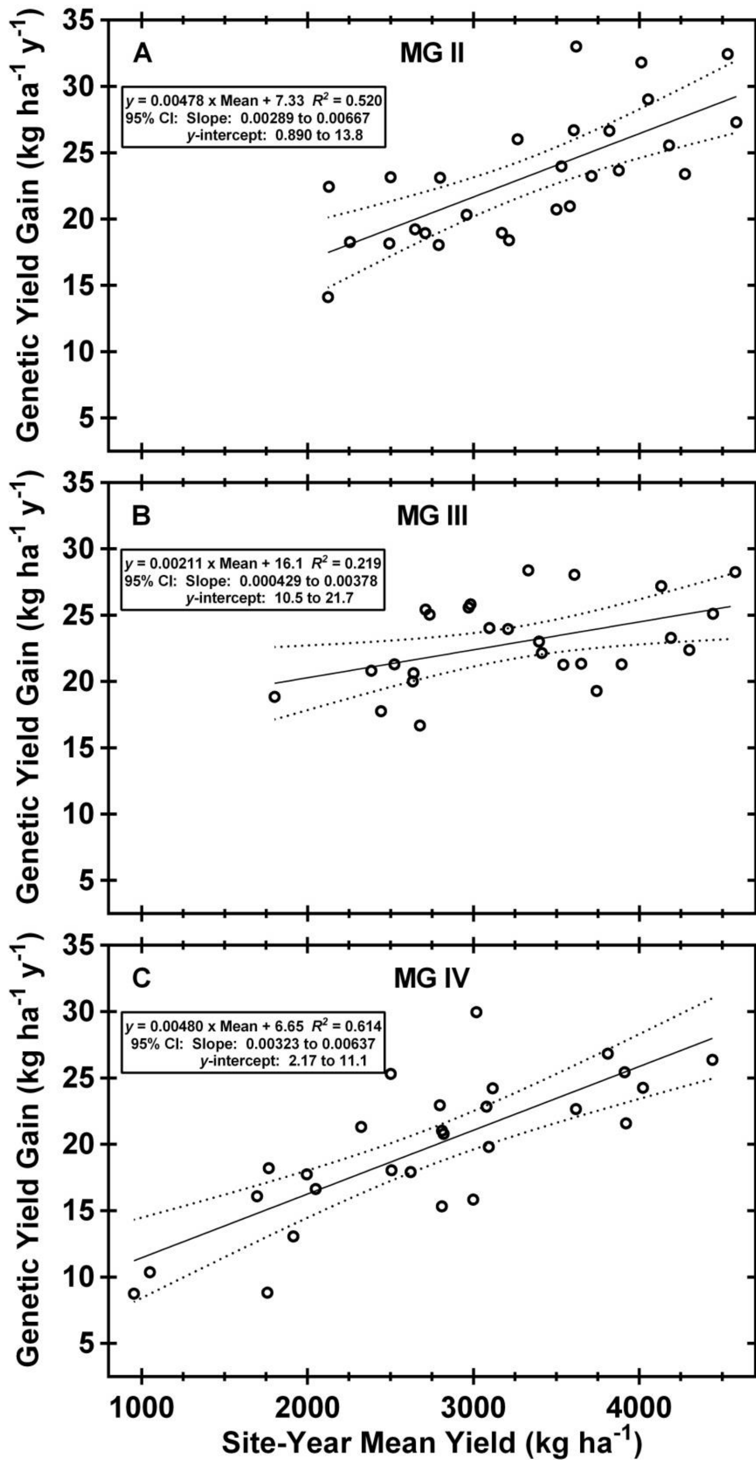


Figure 4.5. Scatter plot of cultivar R8 maturity date (901 corresponds to September 1, 930 is September 30 and 931 is October 1) vs. cultivar release year for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. Dotted lines are the 95% confidence intervals (CIs) for the linear fit. Each cultivar data point is the mean of 27 (MGs II and IV) or 26 (MG III) site-years.

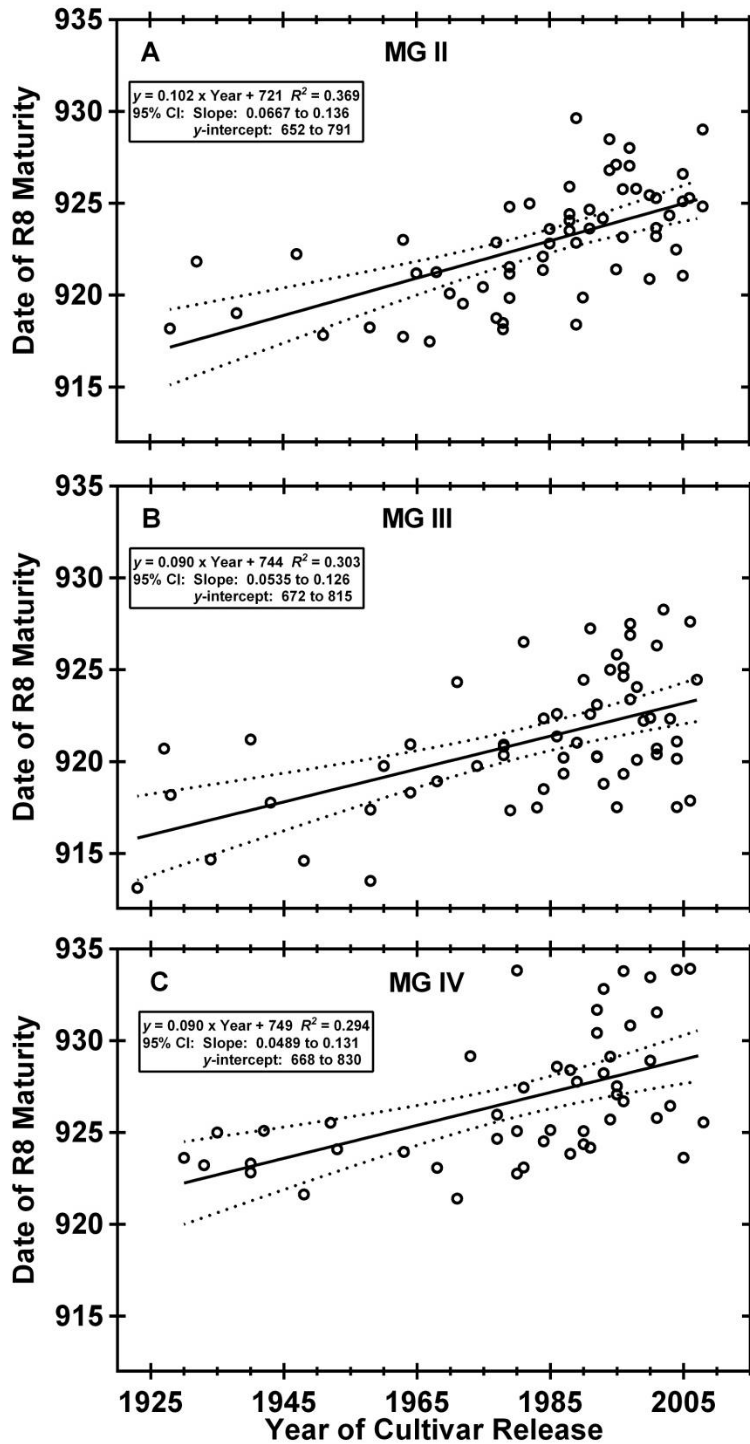
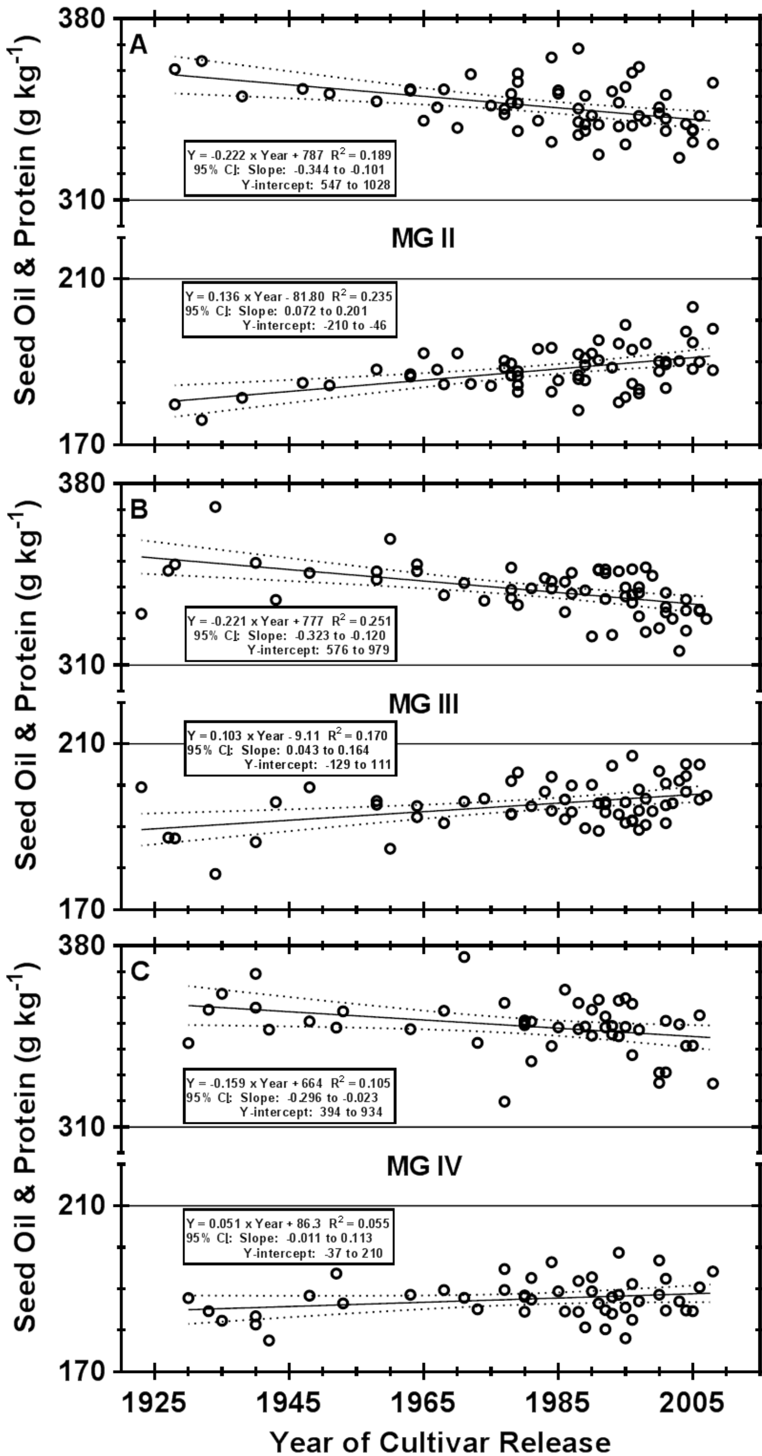


Figure 4.6. Scatter plot of cultivar seed oil and protein content (lower and upper chart) vs. cultivar release year for (A) Maturity Group (MG) II, (B) MG III, and (C) MG IV. A simple linear model was fit to the data, and the best-fit trend line statistics are presented in the text box. Dotted lines are the 95% confidence intervals (CIs) for the linear fit. Each cultivar data point is the mean of 27 (MGs II and IV) or 26 (MG III) site-years.



CHAPTER FIVE

Impact of Soybean Cyst Nematode Resistance on Soybean Yield

INTRODUCTION

Soybean cyst nematode (SCN) (*Heterodera glycines* Ichinohe) is the pest estimated to cause the most damage to soybean [*Glycine max* (L.) Merr.] production in the United States due to its widespread distribution and ability to reduce seed yield. Losses in the United States were estimated to average 2,771,493 Mg annually during 2003 – 2005 (Wrather and Koenning, 2006) and 3,468,684 Mg annually during 2006 – 2009 (Koenning and Wrather, 2010). Growers are often slow to respond to SCN infestations in fields because yield losses often occur when aboveground symptoms, such as plant stunting, are not present (Wang et al. 2003 and Young, 1996). Damage to soybean plants occurs when juvenile nematodes penetrate roots and feed from vascular tissue (Koenning, 2004). As the lifecycle of SCN continues, eggs are produced in cysts that offer protection to the eggs for several years until optimum conditions are present for hatching (Koenning, 2004).

Populations of SCN are described by a *Heterodera glycines* (HG) type classification system which separates the major genetic groups based on host compatibility (Niblack, 2002). The HG type designation of a nematode population identifies which standard indicator lines it can reproduce on. The indicator lines represent SCN resistant sources that have been utilized in breeding resistant cultivars. Therefore, an SCN population that can reproduce on a standard indicator line would then be expected to reproduce on cultivars with resistance derived from this indicator line. Understanding the SCN HG type present in the soil is important when producers select a resistant cultivar.

The most effective methods to manage SCN are to utilize host resistance and to rotate with non-host crops (Niblack and Chen, 2004). Resistance to at least one SCN HG type has been found in 158 soybean accessions (https://npgsweb.ars-grin.gov/gringlobal/search.aspx). Furthermore, resistance has been bred into cultivars; however, PI 88788 is the predominate source of SCN resistance for cultivars available to growers in the northern USA. For instance, the University of

Illinois Department of Crop Science Variety Testing program evaluated 336 SCN resistant entries within maturity groups (MG) II, III, and IV in 2013, but only ten contained sources other than PI 88788 (Joos et al., 2013). As a likely consequence of the abundant use of a single source of resistance, Niblack et al. (2008) found that 70% of the SCN infested soil samples collected in Illinois had populations that could reproduce on PI 88788. Similar results were found in other soybean producing regions as well (Faghihi et al., 2010 and Mitchum et al., 2007). In contrast, resistance from other sources such as PI 437654 and Peking remain effective, but development of high yielding resistant lines from sources other than PI 88788 continues to be a challenge.

Resistant cultivars have repeatedly shown yield advantages compared to susceptible cultivars when SCN is present (Brucker et al., 2005; Chen et al., 2001; DeBruin and Pedersen, 2008; Delheimer et al., 2010; Donald et al., 2006; Koenning, 2000; Niblack et al., 1992; Wheeler et al., 1997; Young, 1996). Furthermore, as initial egg densities increases, a greater yield difference between resistant and susceptible cultivars has been found (Chen et al., 2001; Koenning, 2000; Niblack et al., 1992). With the end goal of developing high yielding resistant cultivars, evaluation of breeding lines at many well characterized locations is needed. For this reason, the Northern Regional Soybean Cyst Nematode Tests are conducted across the north-central soybean producing region of the United States and southern Ontario to evaluate the yield performance of SCN resistant germplasm from public soybean breeders. Yield test locations range from non-infested fields to fields with high initial SCN egg counts. Additionally, most locations are further characterized by HG type testing of the SCN population in the field. Data from these tests are a tremendous resource of replicated yield trial results from resistant and susceptible lines grown in MG 00 through IV locations.

The spatial variability of SCN in field environments has been documented (Avendaño et al., 2003; Donald et al., 1999; and Francl, 1989) and controlling this variability is challenging to researchers. To address this spatial variability, researchers could either increase soil sampling or increase the number of environments sampled. The benefit provided by utilizing the Northern Regional SCN Tests is the power gained by utilizing a large sample of environments to accurately calculate the yield impact that resistance to SCN provides.

Our objectives of this study are three-fold, i) to determine the impact of SCN resistance on yield by comparing SCN resistant and susceptible breeding lines in fields that vary for SCN pressure as measured by egg counts at planting, ii) explore maturity group as a predictor that impacts the yield of resistant breeding lines compared to susceptible lines, and iii) gain insights into the relationship and interactions of an SCN population's egg count at planting and ability to reproduce on PI 88788. The Northern Regional Soybean Cyst Nematode Tests provide a rich resource to evaluate yield relationships over many locations and breeding lines.

MATERIALS AND METHODS

Composition of the dataset

The Northern Regional Soybean Cyst Nematode Tests (Cary and Diers, 2014) are conducted in the north-central United States and Canada across MG 00 – IV through support provided by the United Soybean Board. Each year public soybean breeders nominate experimental lines that they believe are SCN resistant and these lines are organized, together with check cultivars, into separate MG tests. Seed of these experimental lines are shipped to the University of Illinois, repackaged, and distributed to the public breeder collaborators who grow the tests at typically 6 to 20 locations annually for each MG (Table 5.1). After experimental lines are separated into tests by MG, the MG tests are also separated into conventional (non-transgenic) tests or transgenic tests and between preliminary tests, which include entries that have not been previously tested in the regional test, or uniform tests that include fewer entries which have been previously evaluated in preliminary tests. Whether there were separate preliminary and uniform tests and conventional and Roundup Ready tests was determined based on need and all tests were not grown for each combination of year and MG. Field plots were replicated two to four times at each location and grown in multiple rows with row spacing ranging from 19 to 76 cm and the center rows were harvested for yield.

Soil samples from each test location were collected at planting by taking cores 15-20 cm deep spread over the test area. The samples were shipped to the University of Illinois and analyzed to determine the initial SCN egg count which is reported as eggs per 100 cm^{-3} soil (Pi) (Faghihi and Ferris, 2000). The HG type of the SCN sample from each location was determined according to

Niblack et al. (2002). Briefly, the nematodes were first allowed to reproduce on the susceptible cultivar Essex (Smith and Camper, 1973) to increase the population, if needed. One seedling from an indicator line was then placed in a tube filled with sterilized sandy soil and inoculated with 1,000 eggs from the soil sample being tested. Each test was grown in the greenhouse with 27° C soil temperature, 16 h of light for 30 d, and each indicator line replicated six times (Cary and Diers, 2014). A female index (FI) was then calculated for each indicator line by dividing the mean number of cysts on the indicator line with the mean number of cysts on the susceptible cultivar Lee 74 (Caviness et al., 1975) and multiplying by 100. The FI on indicator line PI 88788 is utilized in our analysis and is designated as FI_{88788}.

Data from the Northern Regional Soybean Cyst Nematode Tests in the years 2004 to 2014 were maintained at the University of Illinois as the means of experimental lines and checks for each location. Yield data were combined with two more sets of data, 1) a characterization of the location's SCN population (Pi and HG type, as mentioned above) along with 2) the FI of each experimental line and check from greenhouse SCN testing at the University of Illinois (2005-2014) and at the University of Missouri (2004). The test entries were each tested separately with a SCN HG type 0 and a 2.5.7 population. A FI was calculated for each entry in the same manner as HG type testing and designated as FI_{entry}. Entries were rated separately for each HG type 0 and 2.5.7 population based on the FI_{entry} number as highly resistant (HR, $FI_{entry} < 10$), resistant (R, FI_{entry} of 10 to 24), moderately resistant (MR, FI_{entry} of 25 to 39), low resistance (LR, FI_{entry} 40 to 59) or no effective resistance (NR, $FI_{entry} > 60$) (Niblack, 2005). To avoid confusing ratings from the two HG type populations, the ratings were suffixed with the number of the population HG type, either 0 or 257. For entries tested in multiple years, an average of FI_{entry} value was calculated across years and included in the overall dataset. *Heterodera glycines* Type 0 ratings from entries in the year 2008 were not included due to low cyst counts.

This screening and rating system provided two variables to characterize an entry's resistance. However, neither rating variable alone is fully informative of an entry's resistance to SCN populations commonly found in production fields. For example, highly resistant and resistant entries to HG type 0 may or may not have resistance to HG type 2.5.7. Also, entries with no effective resistance to HG type 2.5.7 may or may not contain some level of resistance to the HG

type 0 screening population. Therefore, a new rating was developed to better classify resistance to both screening populations. First, if entries were found to have high resistance (HR) or resistance (R) to the HG type 2.5.7 population, they were classified as HR257 or R257. Next, those entries not placed in the first categories and were found to have high resistance or resistance to the HG type 0 population, were classified as HR0 or R0. Then, entries with no resistance (NR) to both screening populations were rated as S. All other entries were rated as having a mid-level resistance and included entries with moderate or low resistance to HG type 0 and a moderate, low, or no resistance to HG type 2.5.7. This created a rating system that when summarized for resistance from greatest to least is ordered as follows: HR257, R257, HR0, R0, mid-level, and S.

Steps were taken to account for yield data quality. First, entries and environments with average yield below 673 kg ha^{-1} were removed since these do not represent a typical production agriculture environment and were likely overly influenced by environmental factors such as moisture that were outside the scope of this research. Next, locations with high coefficient of variation (>20%), as listed in the Northern Regional Soybean Cyst Nematode Tests Report were removed. These high coefficients of variation values indicate that a high level of variability exists among entries and across replications, caused by such factors as localized pests, unequal moisture within blocks or plot errors. Overall, 6% of the data were removed based on data quality parameters, primarily high coefficient of variation. To evaluate the yield impact of resistant entries compared to susceptible entries, a new variable was calculated from the yield data generated. First, all susceptible entries (S, $FI_{entry} > 60$), within a test were identified by having no effective resistance to both HG type 0 and 2.5.7 screening population. The average yield of all susceptible entries was then calculated within each environment of that test. Next, the yield of all other entries were divided by the susceptible average within that environment and multiplied by 100 to create a new variable, yield as a percentage of susceptible.

Statistical analysis

The initial egg counts, Pi, was transformed, $\log_{10}(Pi+1)$, due to a non-normal distribution of raw Pi values. Our model to analyze the yield impact of resistance compared to susceptible entries included the independent variables of resistance rating from the experimental lines and checks,

the log transformed egg counts, from the field samples, the squared term of log transformed egg counts, and the interactions of resistance rating with each the log transformed egg counts and the squared term. Means of the yields within each resistance ratings were also calculated at non-infested environments. To further explore the additional predictor variable of MG, a model was analyzed with this variable included. Since the spatially variable nature of SCN populations combined with our data structure requires large datasets to identify trends, exploring all levels of MG and possible interacting factors was not feasible. To circumvent this, we further analyzed MG by grouping MGs into early (MG 00, 0, I, and II) and late (III and IV) data subsets and analyzed the results from a model described above on each subset. Data were analyzed with Proc Mixed of SAS 9.3 (SAS Institute, 2011). Predicted mean yields at levels of specified Pi were calculated with an LSMEANS statement and p-values were adjusted to control for Type I statistical errors by the Dunnett procedure.

Data of SCN population virulence on PI 88788, FI_{88788}, from HG type tests were also non-normally distributed and transformed by $\log_{10}(FI_{88788}+1)$. To gain insights into the effects of the SCN population virulence on PI 88788 in combination with Pi, we developed a separate model which included $\log_{10}(Pi+1)$, $\log_{10}(FI_{88788}+1)$, and a resistance rating composed of either susceptible entries or entries with resistance from PI 88788 (HR0 and R0). Data from non-infested sites were excluded from the analysis to gain better predictions of regression slopes at varying SCN infestations. Significance of factors and interactions were calculated in SAS Proc Mixed of SAS 9.3 (SAS Institute, 2011). Predictions were then graphed between Pi values of 100 and 10,000 where the majority of our SCN infested dataset is located.

RESULTS

After data quality steps were taken, the overall dataset from 11 years of tests included six maturity groups, 1,682 soybean entries, 408 environments, and over 25,000 observations. When considering multiple tests conducted in a single environment, 1,247 different test-environment combinations were represented in the dataset (Table 5.1). When filtering to only locations with HG type 2 ($FI_{88788} > 10$), 183 environments and 595 test-environments were represented (Table 5.1). The variable of Pi ranged from non-infested environments to over 30,000 eggs 100 cm^{-3} of soil. The distribution of the test-environment's log transformed Pi is shown in

Fig. 5.1, Panel A. Yields ranged from 673 to 6893 kg ha^{-1} and yield response ranged from 23-256%. The number of susceptible entries in a test varied from 1 to 12 entries with an average of 3.6 entries and a prevalence of fewer susceptible entries in MG IV tests. The number of entries within our rating system is as follows: HR257, 57; R257, 42; HR0, 679; R0, 333; mid-level, 316; S, 193 while the remaining entries were found to have incomplete screening test results.

Resistance rating and initial egg count as predictors of yield

Initial egg count is a metric to quantify the number of eggs within a volume of soil and potentially predict the amount of infection and subsequent yield loss. As expected and similar to previous studies, we found all resistant breeding lines to have a yield advantage compared to the susceptible (S) at high Pi (Fig. 5.2). The analysis of the whole data set with the single resistance rating scale show that the resistance rating variable, $(\log_{10}(Pi+1))^2$, and the interaction of both resistance rating × $\log_{10}(Pi+1)$ along with resistance rating × $(\log_{10}(Pi+1))^2$, were significant factors explaining yield ($P < 0.0001$) (Table 5.2). Interpretation of this polynomial model and effect estimates indicate an overall upward trend of resistant ratings for yield and differences among the slopes exist (Fig. 5.2). Additionally, all resistance classes except HR257 show an increased rate of yield gain compared to susceptible as Pi increases. As a result of this increase of yield that resistance provides, we find that breeding lines with any level of resistance show significantly higher yield than the susceptible at Pi = 1,000 and 10,000 ($\log_{10}(Pi+1) = 3$ and 4) (Dunnett Adj $P < 0.0001$). The resistance rating with the highest predicted yield over susceptible at Pi = 10,000 ($\log_{10}(Pi+1)$ of 4) was found to be the R0 rating with 121% and at Pi = 1,000 ($\log_{10}(Pi+1) = 3$) the HR0 and R257 with 109%. The ratings of R257 and HR0 even show a significantly higher yield than the susceptible at Pi = 100 ($\log_{10}(Pi+1) = 2$) (Dunnett Adj $P < 0.05$).

In non-infested environments (Pi = 0) significant differences were found among resistance ratings ($P < 0.0001$). Contrasts with the susceptible entries (S) show that the highly resistant classifications (HR0 and HR257) and mid-level resistance yielded less than the susceptible entries (Dunnett Adj $P < 0.001$, Table 5.3). The mean percent of susceptible values for these groups were found to be 2.9 and 3.9% less than the susceptible entries. However, resistant entries (R257 and R0) were found to not be significantly different than the susceptible entries.

The effect of MG on yield was further explored in a separate model and found to be significant ($P < 0.0001$). Maturity group effect estimates indicated that the impact of resistance on yield was greater in earlier maturity groups compared to later maturity groups in this study. In order to maintain adequate sample size within analyses, maturity groups were divided into early (MG00-II) and late (MGIII and IV) maturity subsets, and analyses were conducted to explore the impact of resistance on yield within the subsets. In Figure 5.3, the predicted percent of susceptible (y-axis) is graphed with the $\log_{10}$(Pi+1) (x-axis) of early and late subsets of data. Predicted responses of resistance ratings at Pi = 1,000 and 10,000 ($\log_{10}$(Pi+1) = 3 and 4) ranged from 104% to 131% in the early MGs (Fig. 5.3, Panel A) and 101% to 112% in the late MGs (Fig. 5.3, Panel B).

SCN virulence on PI 88788

The abundant use of PI 88788 as a source of resistance to SCN warrants investigations into the ability of breeding lines with this source of resistance to yield well when grown in fields with SCN populations able to reproduce on PI 88788 ($FI_{88788} \geq 10$). To gain insights into the effects of the ability of the nematode population to reproduce on PI 88788 in combination with field Pi on highly resistant (HR0), resistant (R0), and susceptible (S) breeding line yield, we analyzed entry yield data instead of percent of the susceptible average. This is due to a correlation found in HG type testing between the number of females on PI 88788 and the number of females on the susceptible check Lee74 ($r = 0.35$, $P < 0.0001$) for the SCN populations at locations. This suggests that virulent populations on PI 88788 are also more virulent on susceptible entries when compared to populations that are not virulent on PI 88788. We analyzed a model with initial egg counts ($\log_{10}$(Pi+1)), virulence on PI 88788 ($\log_{10}(FI_{88788}+1)$), the resistance ratings of susceptible and the combined ratings of highly resistant and resistant (HR0 and R0), and all interacting factors. The interacting factors of resistance rating × $\log_{10}(FI_{88788}+1)$ and the three-way interaction of resistance rating × $\log_{10}$(Pi+1) × $\log_{10}(FI_{88788}+1)$ were found to be non-significant (simultaneous F-test, $P = 0.0821$) and were removed from the optimal model. Polynomial squares were evaluated for inclusion in the model; however issues arose with inflated variances due to multicollinearity. Table 5.4 shows results from an optimal model which all factors were found to be significant at the $P = 0.05$ level, except virulence on PI 88788,

$\log_{10}(FI_{88788}+1)$ ($P = 0.0709$). The relationship found between $\log_{10}(Pi+1)$, $\log_{10}(FI_{88788}+1)$, and the predicted yield of highly resistant and resistant (HR0 and R0) and susceptible (S) entries is shown in Figure 5.4. Across all levels of virulence on PI 88788, resistant entries yielded more than susceptible entries and as Pi increased the yield advantage of resistant entries increased.

DISCUSSION AND CONCLUSION

This research offers a unique opportunity to study the yield impact of SCN resistance across a wide range of environments. The scope of MG, genotypes, number of environments, and overall characterization of SCN populations included in this dataset is unmatched within public soybean yield testing efforts. While previous studies have shown the value of resistant soybean cultivars in environments infested with SCN, this dataset offers more precision by greatly expanding the number of observations and the opportunity to explore relationships among MG and varying differences in the ability of field SCN populations to infect PI 88788.

As expected, we found that entries resistant to SCN, either HG type 0 or 2.5.7 populations, yielded greater than susceptible entries at high levels of initial egg counts ($\log_{10}(Pi+1) > 3$). It is important to note that even at Pi of 100 eggs cm^{-3} ($\log_{10}(Pi+1) = 2$), yields of HR0 and R257 were significantly higher than the susceptible average. In addition, just as other researchers have reported, we found that as SCN Pi increases so does the impact of resistant entries compared to susceptible entries (Chen et al., 2001; Niblack et al., 1992; Koenning, 2000). Although we found significant differences in yield at non-infested sites, our data is not well suited to evaluate a yield drag associated with resistant breeding lines. Susceptible check entries have been selected for high yield with multiple years of advanced yield testing, while breeding line entries have not.

While analyzing data across MGs, we were able to show the impact of resistance on yield is greater in early MGs (MG00-II) compared to later MGs (MGIII and IV). Care should be taken when interpreting our analysis, however. Since we have measured the yield as a percent of the susceptible entries within a test, there is the possibility that the differences among early versus late MGs could be due to high yielding resistant breeding lines, low yielding susceptible breeding lines, or both occurring in the early MGs compared to the late MGs. In addition, the SCN populations in early and late MGs could vary considerably. Soybean cyst nematode has

been identified in soils of the MG III and IV region for a longer time than earlier MGs (Riggs, 2004). Subsequently, PI 88788 has been used as a source of resistance for a longer timeframe in late MGs compared to early MGs (Faghihi et al., 2010). To explore the possibility that differences in SCN population from early vs. late MGs could cause the differences in yield, we investigated the data for a trend of an SCN population's virulence on PI 88788 ($\log_{10}(FI_{88788}+1)$) and MG. No trend was found. However, a trend was found among the Pi of early MG locations and late MG locations. The average Pi of observations in locations of early MGs is 1693 while the average Pi in locations of late MGs is 1072. When coding the MGs as an ordinal variable, a correlation of $r = -0.10$ ($P < 0.0001$) was found with Pi. Although higher Pi levels were found in sites of early MGs in this study, when evaluating specific levels of Pi, a greater percent of susceptible yields were found in early MGs. In addition to differences among SCN populations, differences may exist among susceptible entries grown in early and late MGs. Susceptible entries developed in regions where SCN has been present for many years may have built up a small level of resistance that is not differentiated by the rating of NR (FI > 60) in screening efforts. We compared the FI of susceptible entries screening with HG type 0 and 2.5.7 among early MG and late MGs. No differences were found.

Examination of the relationship between initial egg counts and virulence on PI 88788, FI_{88788}, revealed that growing resistant varieties remains important as Pi increases and also as FI_{88788} increases. Breeding lines with resistance from PI 88788 remain higher yielding than susceptible entries even at high FI_{88788} levels. One explanation for why there was limited yield loss on entries with PI 88788 resistance at $FI_{88788} > 10$, ($\log_{10}(FI_{88788}+1) > 1$), is that many of the SCN populations present at these environments are infecting PI 88788 at rates lower than the susceptible control Lee 74 within HG type testing. Evidence of this is shown in Figure 5.1 Panel B where the majority (97%) of SCN populations in this study are virulent on PI 88788 at $\log_{10}(FI_{88788}+1)$ of less than 1.8 ($FI_{88788} = 62$). This means that SCN populations with a $FI_{88788} = 62$ would infect the roots of PI 88788 at 62% of the level of infection on the susceptible Lee74. Therefore, it is expected that entries with resistance derived from PI 88788 and showing no resistance to the HG type 2.5.7 in greenhouse tests would then show some resistance to SCN populations found in the soil of most production fields in this study labeled with the presence of HG type 2 ($FI_{88788} > 10$).

We did not find criteria to recommend that resistance sources other than PI 88788 should be utilized on the sole basis of yield. While locations did exist where entries with sources of resistance other than PI 88788 yielded more than others, the highly variable nature of this data suggests that more observations are needed from breeding lines with resistance sources other than PI 88788. In addition, more observations are needed from all resistance sources in environments with FI_{88788} of greater than 40 ($\log_{10}(FI_{88788}+1) > 1.6$) (Fig. 5.1, Panel B). Within our dataset, only 21 of 408 environments exceed this level of FI_{88788}. The occurrence of environments capable of reproducing on PI 88788 ($FI_{88788} > 10\%$) has been reported to be increasing (Faghihi et al., 2010; Mitchum et al., 2007; and Niblack et al., 2008) which suggests that an increase in environments with $FI_{88788} > 40$ could be occurring as well. A focused effort to test multiple genotypes with varying resistance levels is needed to improve estimations of performance at environments that we expect PI 88788 derived resistance to break down. An additional consideration is that the use of only PI 88788 as a source of resistance will continue to expose SCN populations to high selection pressures. These pressures could result in SCN populations that are better able infect PI 88788 than found in this study. Rotation of resistant sources and non-host crops remains important.

LITERATURE CITED

Avendaño, F., O. Schabenberger, F.J. Pierce, and H. Melakeberhan. 2003. Geostatistical analysis of field spatial distribution patterns of soybean cyst nematode. Agron. J. 95:936–948.

Cary, T. and B. Diers (ed.) 2014. 2014 Northern regional soybean cyst nematode tests. Univ. of Illinois, Urbana.

Caviness, C. E., R. D. Riggs, H. J. Walters. 1975. Registration of Lee 74 soybean. Crop Sci. 15: 100.

Chen, S. Y., P. M. Porter, J. H. Orf, C. D. Reese, W. C. Stienstra, N.D. Young, D. D. Walgenbach, P. J. Schaus, T. J. Arlt, and F. R. Breitenbach. 2001. Soybean Cyst Nematode Population Development and Associated Soybean Yields of Resistant and Susceptible Cultivars in Minnesota. Plant Dis. 85:760-766.

DeBruin, J. L. and P. Pedersen. 2008. Response of old and new soybean cultivars to *Heterodera glycines* Ichinohe. Agron. J. 100:1347-1353.

Donald, P.A., W.W. Donald, A.J. Keaster, R.J. Kremer, J.A. Kendig, B.S. Sims, and J. Mihail. 1999. Changes in *Heterodera glycines* egg population density in continuous *Glycine max* over four years. J. Nematol. 31:45–53.

Donald, P. A., P. E. Pierson, S. K. St. Martin, P. R. Sellers, G. R. Noel, A. E. MacGuidwin, J. Faghihi, V. R. Ferris, C. R. Grau, D. J. Jardine, H. Melakeberhan, T. L. Niblack, W. C. Stienstra, G. L. Tylka, T. A. Wheeler, and D. S. Wysong. 2006. Assessing *Heterodera glycines*-resistant and susceptible cultivar yield response. J. Nematol. 38:76-82.

Faghihi, J. and J. M. Ferris. 2000. An efficient new device to release eggs from *Heterodera glycines*. J. Nematol. 32:411-413.

Faghihi, J., P. A. Donald, G. Noel, T. W. Welacky, and V. R. Ferris. 2010. Soybean resistance to field populations of *Heterodera glycines* in selected geographic areas. Plant Health Prog. doi:10.1094/PHP-2010-0426-01-RS.

Francl, L.J. 1986. *Heterodera glycines* population dynamics and relation of initial population to soybean yield. Plant Dis. 70:791–795.

Joos, D. K., R. W. Esgar, B. R. Henry, E. D. Nafziger. 2013. Soybean Variety Test Results in Illinois-2013. Crop Sciences Special Report 2013-04. Univ. of Illinois, Urbana.

Koenning, S. R. 2000. Density-dependent yield of *Heterodera glycines*-resistant and -susceptible cultivars. J. Nematol. 32:502-507.

Koenning, S. R. 2004. Population biology. In: D.P. Schmitt et al. editors. Biology and management of soybean cyst nematode, 2nd ed. Schmitt and Associates of Marceline. Marceline, MO. pp. 73-88.

Koenning, S. R. and J. A. Wrather. 2010. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. Plant Health Prog. doi:10.1094/PHP-2010-1122-01-RS.

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

Niblack, T. L. 2005. Soybean Cyst Nematode Management Reconsidered. Plant Dis. 89:1020-1026.

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

Niblack, T. L., N. K. Baker, and D. C. Norton. 1992. Soybean yield losses due to *Heterodera glycines* in Iowa. Plant Dis. 76:943-948.

Niblack, T. L., and S. Y. Chen. 2004. Cropping systems and crop management practices. In: D.P. Schmitt et al. editors. Biology and management of soybean cyst nematode, 2nd ed. Schmitt and Associates of Marceline. Marceline, MO. pp. 181-206.

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

Riggs, R.D. 2004. History and Distribution. In: D.P. Schmitt et al. editors. Biology and management of soybean cyst nematode, 2nd ed. Schmitt and Associates of Marceline. Marceline, MO. pp. 9-40.

SAS Institute. 2011. The SAS system for Microsoft Windows. Release 9.3. SAS Inst., Cary, NC.

Smith, T. J. and H. M. Camper, 1973. Registration of Essex soybean. Crop Sci. 13:495.

Wang, J., T. L. Niblack, J. A. Tremain, W. J. Wiebold, G. L. Tylka, C. C. Marett, G. R. Noel, O. Meyers, and M. E. Schmidt. 2003. Soybean cyst nematode reduces soybean yield without causing obvious aboveground symptoms. Plant Dis. 87:623-628.

Wheeler, T. A., P. E. Pierson, C. E. Young, R. M. Riedel, H. R. Wilson, J. B. Eisley, A. F. Schmitthenner, and P. E. Lipps. 1997 Effect of soybean cyst nematode (*Heterodera glycines*) on yield of resistant and susceptible soybean cultivars grown in Ohio. J. Nematol. 29:703-709.

Wrather, J. A. and S. R. Koenning. 2006. Estimates of disease effects on soybean yields in the United States 2003 to 2005. J. Nematol. 38:173-180.

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

TABLES AND FIGURES

Table 5.1. Count of environment and test combinations† within Northern Regional Soybean Cyst Nematode Tests (2004-2014). Counts are summarized by state or province and different soybean cyst nematode infestation classifications within maturity groups.

State or province	Maturity group					
	00	0	I	II	III	IV
ND	2	8				
MN	6	37	48	39		
ON	6	32	16	24		
SD			10	17		
NE			24	59	74	
IA			24	33	38	
IL			39	66	147	114
MI				11		
OH				17	28	
IN				13	24	1
TN				3	3	24
KS					49	39
MO					80	77
DE						2
KY						13
Environment classification						
Non infested	8	41	18	45	112	88
Infested, HG type 2 present	3	25	92	146	211	118
Infested, HG type 2 not present	3	5	32	53	47	17
Infested ‡	0	6	19	38	73	47
Total	14	77	161	282	443	270

† Locations were removed with overall low yield (< 673 kg ha^{-1}) and high coefficient of variation (>20%).
‡ Infested environments are lacking full HG type testing of soybean cyst nematode population present in the soil.

Table 5.2. ANOVA results for percentage of susceptible average yield across all environments in the Northern Regional Soybean Cyst Nematode Tests from 2004-2014.

Source	DF	Mean Square	F Value	Pr > F
Resistance rating	5	2,020	6.4	<.0001
$\log_{10}(Pi+1)$	1	245	0.8	0.3793
$(\log_{10}(Pi+1))^2$	1	18,954	59.8	<.0001
$\log_{10}(Pi+1) \times$ Resistance rating	5	2,839	9.0	<.0001
$(\log_{10}(Pi+1))^2 \times$ Resistance rating	5	5,034	15.9	<.0001
Residual	23,259	317		

Table 5.3. Least squared means (LSM) of percent susceptible average for resistance ratings at sites without soybean cyst nematode infestation in the Northern Regional Soybean Cyst Nematode Tests (2004-2014).

Resistance Rating	(LSM) Estimate	Standard Error	Number of Observations	Contrast with S rating, Dunnett Adj-P
S	100.0	0.46	892	-
HR257	96.1	0.93	211	0.0007
R257	98.6	1.36	99	0.8208
HR0	97.1	0.28	2,296	<.0001
R0	98.8	0.38	1,249	0.1582
mid-level	96.0	0.55	597	<0.0001

Table 5.4. ANOVA results for entry yield of breeding lines grown in Northern Regional Soybean Cyst Nematode Tests from 2004-2014. Only entries with resistance from PI 88788 and susceptible entries at infested sites were included.

Source	DF	Mean Square	F Value	Pr > F
Resistance rating	1	6,469	35.9	<.0001
$\log_{10}(Pi{+}1)$	1	1,697	9.4	0.0022
$\log_{10}(FI_{88788}{+}1)$	1	588	3.3	0.0709
Resistance rating × $\log_{10}(Pi{+}1)$	1	17,746	98.4	<.0001
$\log_{10}(Pi{+}1) \times \log_{10}(FI_{88788}{+}1)$	1	4,038	22.4	<.0001
Residual	12,321	180		

Figure 5.1. Histograms of soybean cyst nematode (SCN) population characteristics included in Northern Regional Soybean Cyst Nematode Tests (2004-2014). The percent occurrence is graphed on the y-axes. (A) Initial egg counts (Pi) of test locations are represented by the transformation $\log_{10}(Pi + 1)$. (B) Virulence phenotype on PI 88788 (FI_{88788}) of SCN samples at test locations determined by HG type testing are represented by the transformed variable $\log_{10}(FI_{88788} + 1)$. Values of $\log_{10}(FI_{88788} + 1) = 0$ represent sites without infestation and no virulence on PI 88788. Back transformed values are shown in parentheses along the x-axis.

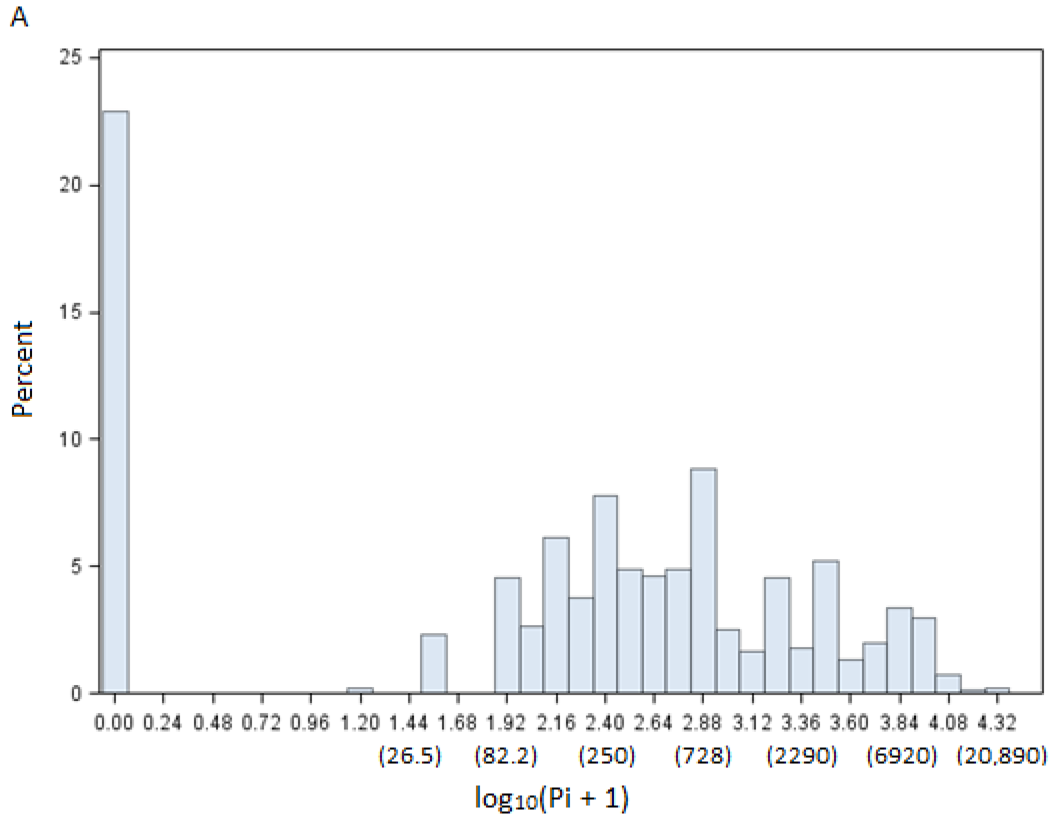


Figure 5.1 (cont.)

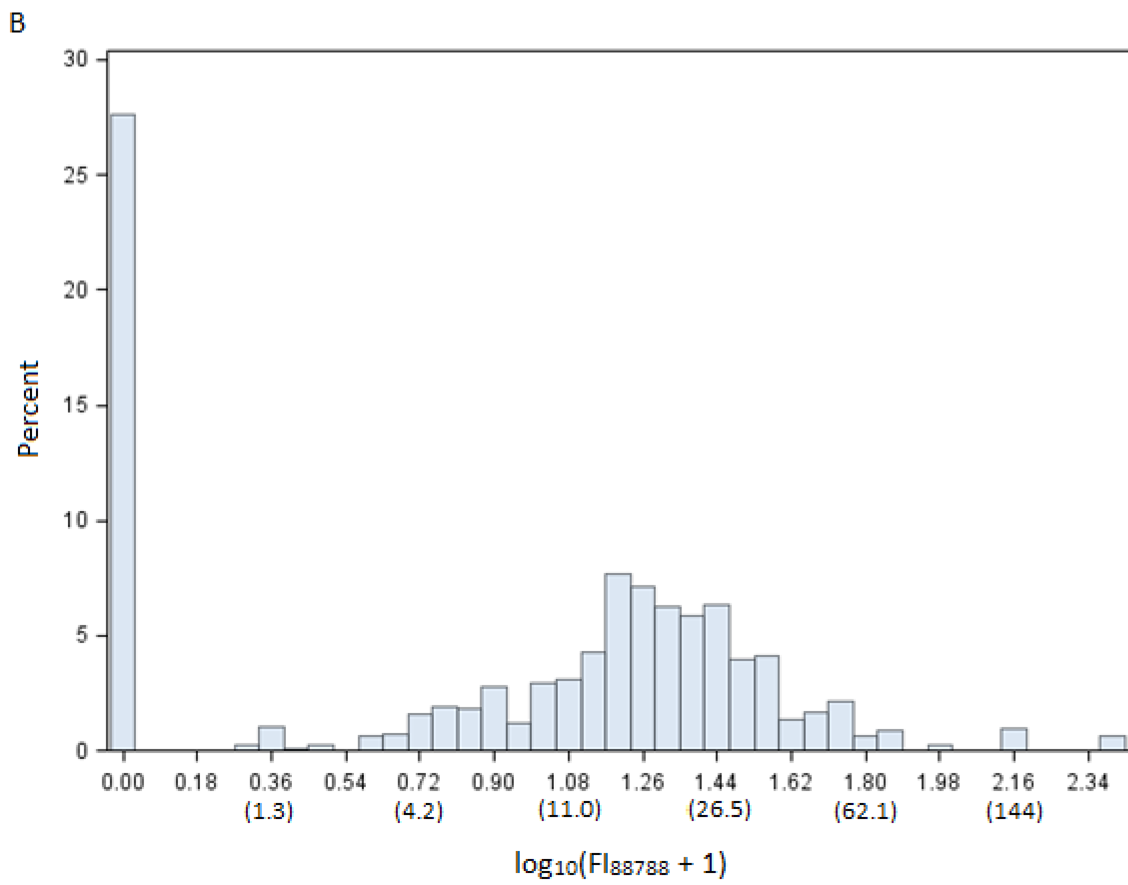


Figure 5.2. Predicted percent susceptible average graphed over the transformed initial egg counts (Pi) represented as $\log_{10}(Pi + 1)$. Predictions are based on a model with Northern Regional Soybean Cyst Nematode Test data from 2004 to 2014 across maturity groups 00-IV.

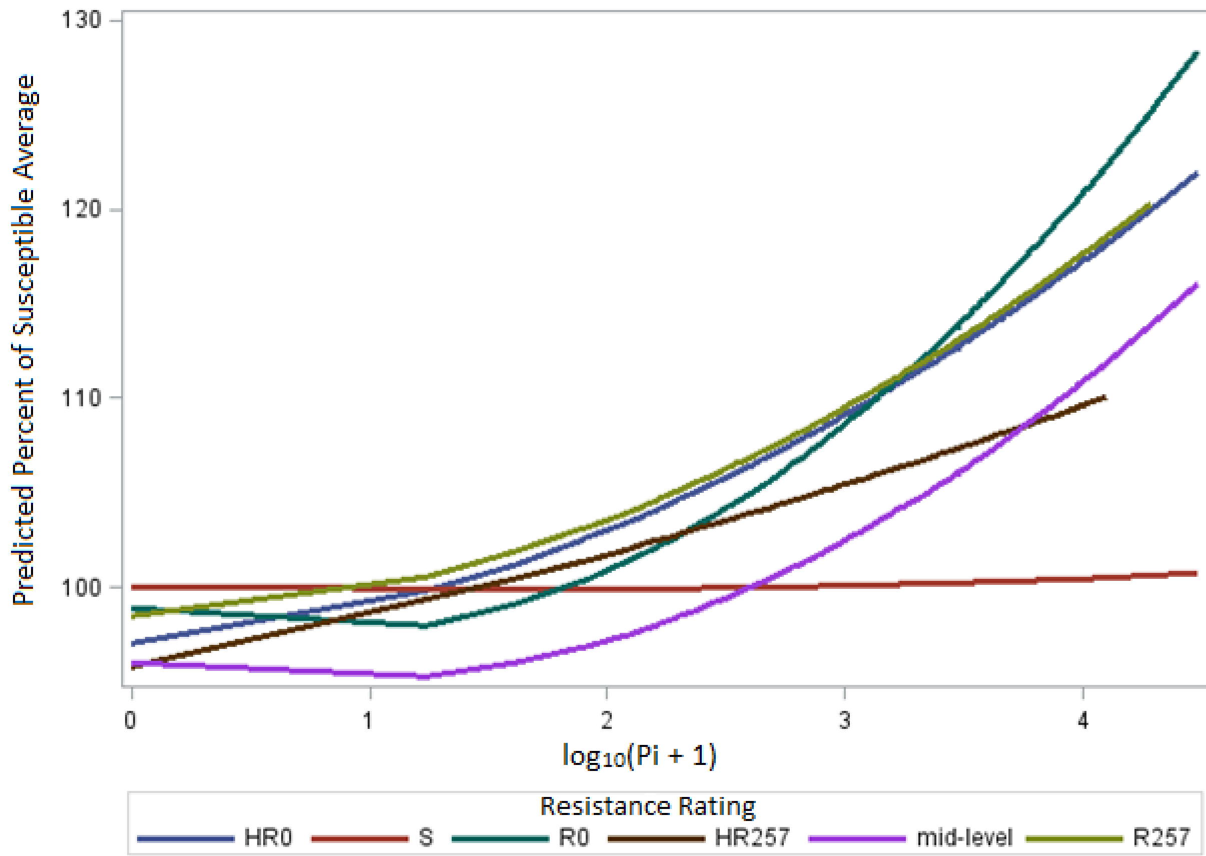


Figure 5.3 Predicted percent susceptible average graphed over the transformed initial egg counts (Pi) represented as $\log_{10}(Pi + 1)$. Predictions are based on a model with Northern Regional Soybean Cyst Nematode Test data from 2004 to 2014 across maturity groups 00-II (A) and maturity groups III-IV (B).

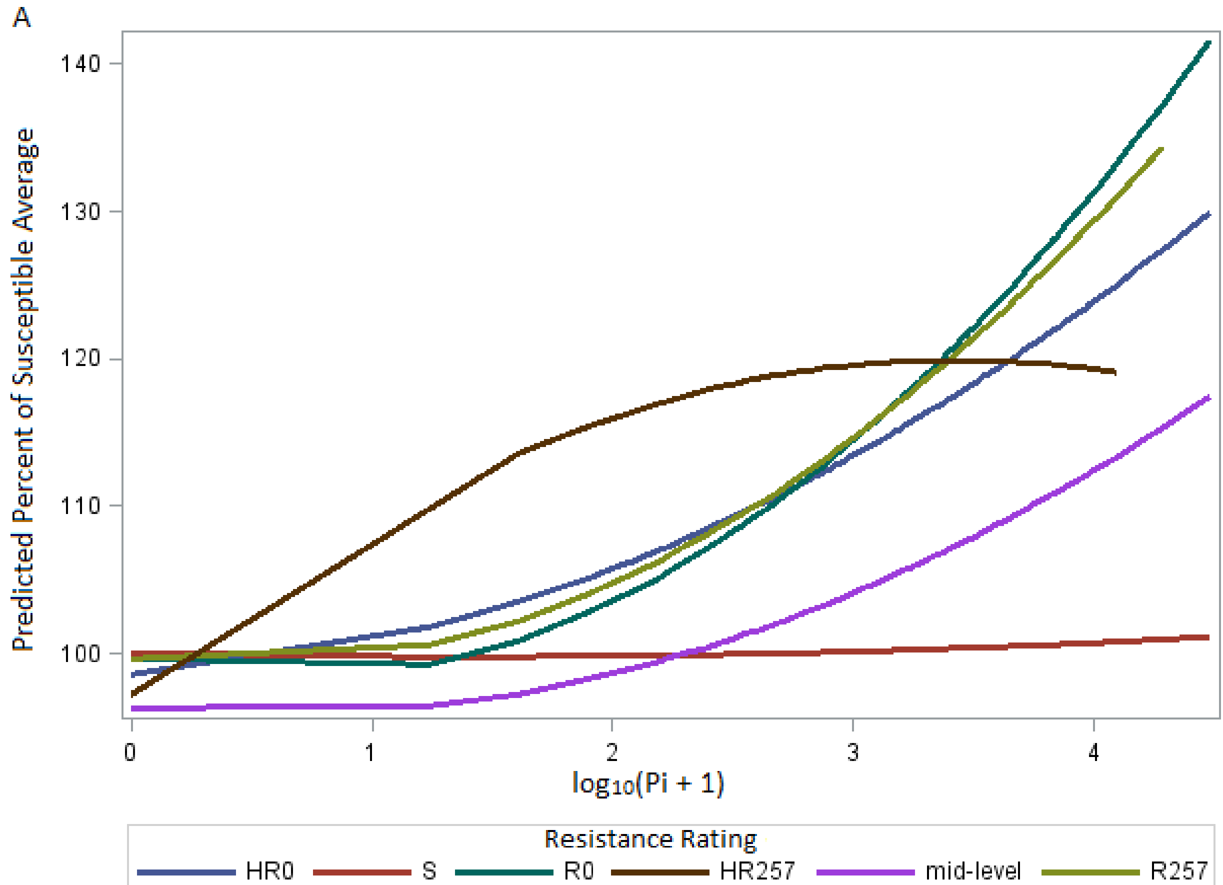


Figure 5.3 (cont.)

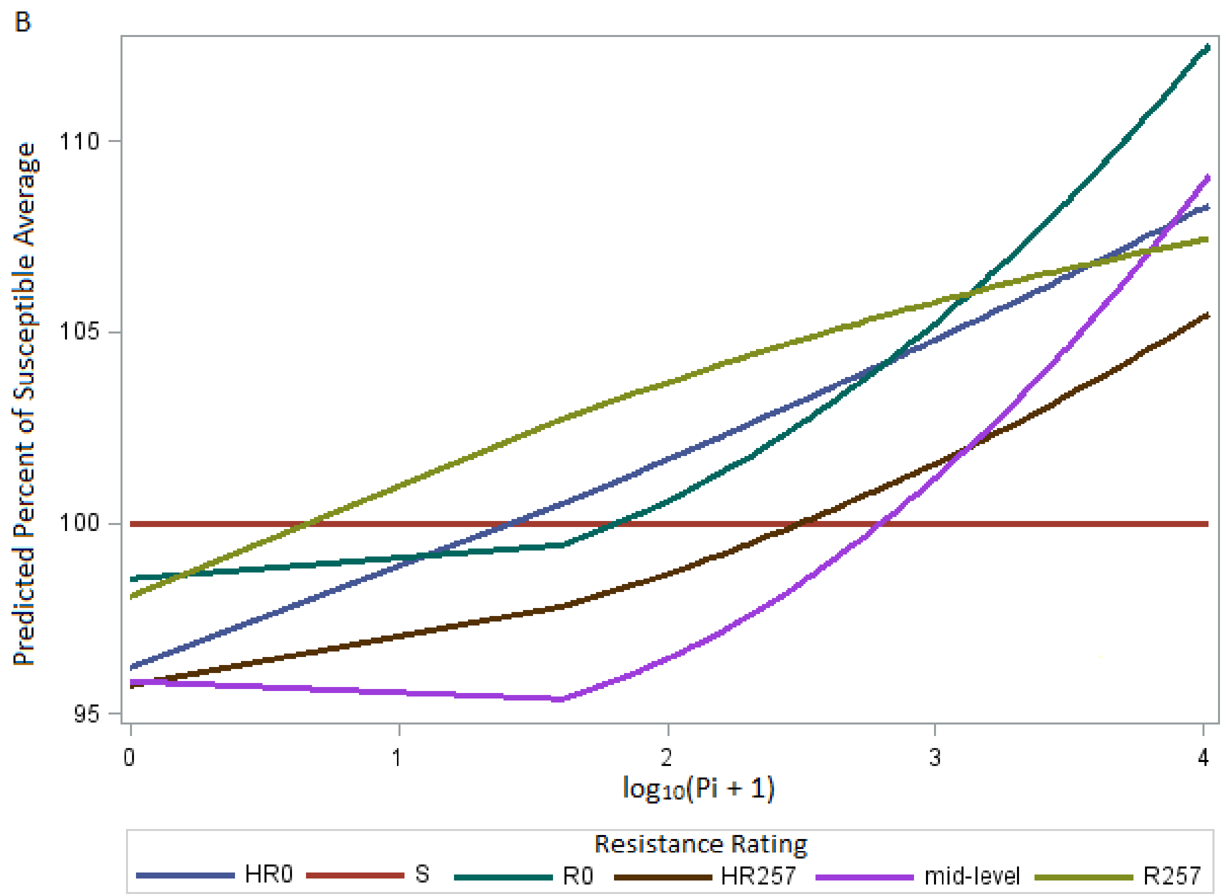


Figure 5.4. Predicted entry yields (kg ha^{-1}) of soybean entries are shown at multiple levels of virulence on PI 88788 (FI_{88788}) and initial egg counts (Pi). A model was fitted to data from infested sites and included Pi, represented by the transformation, $\log_{10}(Pi + 1)$; FI_{88788}, represented by the transformation, $\log_{10}(FI_{88788} + 1)$; and the resistance rating of soybean entries. The relationship of susceptible entries and entries with resistance derived from PI88788 (HR0 and R0) are shown at three levels of $\log_{10}(FI_{88788} + 1)$. These levels are represented at $\log_{10}(FI_{88788} + 1) = 1.8$ (A), 1.2 (B), and 0.6 (C). Back transformed values are shown in parentheses along the x-axis and for each panel.