# FINE MAPPING OF APHID RESISTANCE GENES IN SOYBEAN PLANT INTRODUCTION (PI) 567598B

**By**

**Carmille Joanna C. Bales**

**A DISSERTATION**

**Submitted to**
**Michigan State University**
**in partial fulfillment of the requirements**
**for the degree of**

**Plant Breeding, Genetics and Biotechnology – Crop and Soil Sciences –**
**Doctor of Philosophy**

**2013**

## ABSTRACT

## FINE MAPPING OF APHID RESISTANCE GENES IN SOYBEAN PLANT INTRODUCTION (PI) 567598B

**By**

**Carmille Joanna C. Bales**

The soybean aphid (*Aphis glycines* Matsumura) has become a major pest of soybean [*Glycine max* (L.) Merr.] in North America since it was first discovered in 2000. The best approach to effectively control evolving biotypes of soybean aphids is to stack resistant genes from different sources using marker-assisted strategies. Plant introduction (PI) 567598B possesses strong antibiosis resistance to soybean aphids. Previous work revealed that two recessive genes control the aphid resistance in PI 567598B.

The first objective of the study was to identify the location of the two recessive genes in the soybean genome by genetic linkage analysis of a mapping population consisting of 282 $F_{4:5}$ lines evaluated in the field and greenhouse. The QTLs controlling soybean aphid resistance were mapped on chromosomes 7 and 16 and named *rag1b* and *rag3*, respectively. The two QTLs explained over 30% of the total phenotypic variation. In all trials, *rag3* consistently conferred resistance while a soybean aphid isolate overcame *rag1b* in the field.

The next objective was to fine map the aphid resistance genes by selecting plants with recombination events within the QTL intervals. Fine mapping was conducted by screening 4,041 $BC_1F_2$-derived lines using SNP markers that flank the interval of *rag3*. Fifty-five recombinants were tested with high-density molecular markers using the SoySNP50K genotyping array. Progenies of each recombinant were rated for damage by the soybean aphid and tested using custom-designed SNP assays. Results of fine mapping delimited *rag3* into a 152-kb interval

between SNP markers on chromosome 16. The study showed that effectiveness of *rag1b* is dependent on the presence of alleles for *rag3*. Eight candidate genes found within the *rag3* interval are encoded as NBS-LRR genes; their potential role in soybean aphid resistance in PI 5657598B will be discussed.

**“I have fought the good fight, I have finished the race, I have kept the faith.”**
*2 Timothy 4:7*

**This piece of work is humbly dedicated to**
**Bai Joseph, Papsy, Mamsy and Dodong Alvin**

## ACKNOWLEDGEMENTS

My sincere gratitude to my adviser, Dr. Dechun Wang, for being the best mentor I can hope for. To my committee members: Drs. Amy Iezzoni, Chris DiFonzo and Robin Buell, for their valuable lessons, advice and teaching me the value of humility and success.

To my labmates: Kate, Desmi, John, Yang, Umesh, Wenyan, Shichen, Zixiang, Ruijuan, Paul, John Boyse, Randy, and Cherry for their camaraderie and team work.

Appreciation goes to my other mentors, Drs Russell Freed, Dave Douches, Guo-qing Song, Peter Guzman, and Ana Dias for all their support and guidance. I also want to extend my gratitude to the CSS (PSM) staff, PBGB faculty and graduate students.

My sincere appreciation to Monsanto and PBGB for their financial support for my PhD fellowship; and to United Soybean Board and Michigan Soybean Checkoff programs for my research funding.

To the MSU Filipino Club members and Filipino friends around Lansing who have made my stay here in Michigan a home away from home; most especially to friends who have come and moved on: ate Joy, ate Ann, ate Nelfa, ate Lia, Tintin, and Karla for their moral support and prayers.

To my family Mamsy, Papsy and Dodong for their unfailing love, support and prayers. To the most important person in my life, my husband Joseph, for his unconditional love, for holding my hand and urging me on, and for believing that we can achieve our dreams no matter the odds.

And lastly, to God Almighty for His infinite love and faithfulness to me. I am giving these all back to You.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER I

# LITERATURE REVIEW

## Economic importance of soybean

Soybean is the world's foremost provider of plant protein and oil for human and animal consumption. In 2012, 1.345 billion bushels (36.6 million metric tons) of soybeans were exported by the United States, which accounted for 37% of the world's soybean trade. Domestically, soybeans provided 66% of the edible consumption of fats and oils in the United States (www.soystats.com, 2012). Michigan is 12$^{th}$ in rank to contribute to the nation's total soybean production out of 31 other soybean-producing states. The total economic impact of Michigan soybean farming in 2011 was approximately $1.25 billion and was valued 23% of the total crop planted in the state (Michigan Soybean Promotion Committee, 2012).

## The soybean aphid, life cycle and distribution

The soybean aphid (SBA), *Aphis glycines* Matsumura, was initially known to be a soybean pest in Asia, but became a major threat to soybean production in the Midwest since its discovery in 2000 (Hartman et al., 2001). The impact of SBA on yield is by direct plant feeding; secondary effects are brought about by black sooty mold infestation and virus transmission which can result in plant stunting, leaf deformation and thus reduced pod set (Hill et al., 2001; Li et al., 2008). In Minnesota, high levels of aphid infestation in 2001 were associated with a 50% reduction in yield (Ostlie, 2002).

The soybean aphid has a complicated life cycle, with two morphs (wingless and winged) feeding on the cultivated soybean, its secondary host, during summer (Ragsdale et al., 2004). The buckthorn (*Rhamnus* spp.) was identified as the primary host (Voegtlin et al., 2004) where males and females mate, eggs overwinter, and nymphs hatch in spring to go in search for soybean fields to colonize. Soybean aphids commonly form colonies at the growing points of the soybeans and increase exponentially from the early vegetative stages up to R5 stage (Ragsdale et al., 2004). Ragsdale et al. (2007) found that SBA densities peak during soybean growth stages R3 (beginning pod formation) to R5 (full-size pod) and sometimes even up to the R6 stage (full-size green seed). They thrive best in temperatures from 22-27$^0$C during the months of June-August in Michigan. It has been estimated that at 25$^0$F, aphid populations double in 1.5 days (Mccornack et al., 2004).

Since its discovery in 2000, outbreak years have been reported in several North American states (DiFonzo and Hines, 2002; Losey et al., 2002; Ostlie, 2002; Ragsdale et al., 2004). The spread and population dynamics of soybean aphid alates (winged morphs) have been studied in seven Midwestern states over a period of four years using a suction trap network (Schmidt et al., 2012). In trying to understand movement patterns of alates from *Rhamnus* to soybeans in spring and back to the primary host in fall, it was found that most alates move in a latitudinal pattern, correlated to the distribution of the primary host in the investigated states. Peak flight among soybean fields was from late July to mid-August and another peak during fall in the last two weeks of September (Schmidt et al., 2012). Venette and Ragsdale (2004) reported that the soybean aphid was successful in invading US soybean fields because of the widespread distribution of its primary host in the surrounding states.

**Soybean aphid management**

Significant efforts are being made in the past half-decade to find solutions in keeping SBA populations down and mitigating yield losses caused by soybean aphid damage. Tactics including chemical, biological, and cultural methods have been recommended for SBA control. Integrated Pest Management (IPM) program for soybean aphid encourages the use of cultural and other methods to make decisions that will minimize the use of chemical control as a sound, ecological and economic approach. However, in very high outbreak situations, spraying with insecticides is recommended, although timing and method of application are most important (DiFonzo, 2005). Protecting soybean plants during the flowering to pod-forming stages (R1-R3) and during pod-filling (R5-R6) are key to preventing yield loss.

The most commonly applied insecticides that control SBA are organophosphates, pyrethroids and neonicotinoids available as foliar sprays and seed treatments (DiFonzo, 2005; Ohnesorg et al., 2009; McCarville and O'Neal, 2013). However, there is always the possibility that SBA can develop resistance against these commonly used insecticides and the continual use can harm beneficial insects and predators that naturally keep aphid populations down. The economic injury level (EIL) and economic threshold (ET) for soybean aphid was estimated to guide growers in timing insecticide application. Ragsdale et al. (2007) recommended an ET of 273 aphids per plant providing a lead-time of seven days before aphid populations reached an EIL at 674 aphids per plant. The use of an ET and regular scouting was recommended as a guide for farmers or producers to manage soybean aphids accurately, thus reducing the risk of the loss of insect biodiversity and earlier development of insecticide resistance (Ragsdale et al., 2007;

Catangui et al., 2009). Johnson et al. (2009) showed that the use of an IPM-based system improved the cost-effectiveness of soybean production and reduced the risk of aphid outbreak.

Natural enemies also control SBA. The most common and efficient biological control agents are predators such as the multicolored Asian lady beetle, insidious flower bug (*Orius insidiosus*), syrphid fly and the lacewing larvae. Parasitoids include the parasitic wasp, *Aphelinus albipodus*, that lays eggs inside the soybean aphid (Mahr et al., 2008; Ragsdale et al., 2011). However, the predators and parasitoids were found to be inconsistent and not as effective as chemical control methods (Nielsen and Hajek, 2005).

**Host plant resistance**

The use of insect resistant soybean genotypes is a significant component of IPM for soybean aphids. Thus far, this is cited to be the most effective and economical way of controlling SBA (Ragsdale et al., 2011; Hesler et al., 2013). Host plant resistance against insects comes in multiple forms, including: antibiosis (feeding on the plant reduces fecundity or leads to death), antixenosis (non-preference of the insect to infest the plant), and tolerance (the plant reaches its yield potential despite significant infestation of the pest) (Painter, 1968).

Several studies have identified new and useful plant resistant sources for soybean aphid management (summarized in Table 1.1). By screening commercial cultivars from Asia and the US ancestral pool, Hill et al. (2004) was the first to identify the ancestral genotypes Dowling and Jackson which are resistant against SBA. Expression of resistance in Dowling and Jackson was characterized as mainly antibiosis (Li et al., 2004) and genetic studies showed that Dowling resistance was controlled by a single dominant gene (Hill et al., 2006a). By soybean genetics committee convention, the gene was named *Rag1*, which stands for Resistance against *Aphis*

**Table 1.1** Soybean aphid resistance genes mapped on the soybean genome from identified accessions.

| Gene/ QTL | Source accession | Resistance mechanism | Chr/ Linkage Group | Flanking Markers | Size of interval (cM) | % Variation explained | Reference |
|---|---|---|---|---|---|---|---|
| *Rag* | 'Jackson' | Antibiosis | 07/M | Satt435~Satt463 | 6.1 | - | (Hill et al., 2006a; Li et al., 2006) |
| *Rag1* | 'Dowling' | Antibiosis | 07/M | Satt435~Satt463 | 3.7 | - | (Hill et al., 2006b; Li et al., 2006) |
| *Rag2* | PI 243540 | Antibiosis | 13/F | Satt334~Sct_033 | 4.5 | - | (Kang et al., 2008; Mian et al., 2008) |
| *Rag2* | PI 200538 | Antibiosis | 13/F | Satt334~Sct_033 | 10.0 | - | (Hill et al., 2009) |
| *rag1c* | PI 567541B | Antibiosis | 07/M | Satt150~Satt435 | 24.7 | 43~88 | (Mensah et al., 2008; Zhang et al., 2008) |
| *rag4* | PI 567541B | Antibiosis | 13/F | Satt649~Satt348 | 19.4 | 1.6~30 | |
| *Rag3* | PI 567543C | Antixenosis | 16/J | Satt339~Satt414 | 19.2 | 84~90 | (Zhang et al., 2010) |
| *qRa_1* | 'Zhongdou 27' | Antibiosis | 08/A2 | Satt470 | - | 25~35 | (Meng et al., 2011) |
| *qRa_2* | 'Zhongdou 27' | Antibiosis | 13/F | Satt144 | - | 7~11 | |
| *Rag2/5?* | PI 567301B | Antixenosis | 13/F | BARC- 060107-16382 | 10.0 | 90 | (Jun et al., 2012) |
| | PI 567301B | Antixenosis | 08/A2 | BARC-063283-18296 | 4.0 | 9.8 | |

**Table 1.1 (cont'd)**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Rag3b* | PI 567537 | Antibiosis | 16/J | Satt339~Satt654 | 10.2 | 87.4 | (Zhang et al., 2013) |
| *Rag6_P203* | P203 | Antixenosis | 08/A2 | SSR_08_75~ SSR_08_88 | 1.0 | - | (Xiao et al., 2013) |
| *QTL_6_1* | PI 567324 | Antixenosis | 06 | BARCSOYSSR_06_0998 | 5.6 | 13.1 | (Jun et al., 2013) |
| *QTL_13_1* | PI 567324 | Antixenosis | 13 | BARCSOYSSR_06_1139 | 20.8 | 70.6 | (Jun et al., 2013) |
| *QTL_13_2* | PI 567324 | Antixenosis | 13 | Satt649 | 0.7 | 13.1 | (Jun et al., 2013) |

*Rag* = dominant gene
*rag* = recessive gene

*glycines*. A similar dominant gene, *Rag*, was found in Jackson (Hill et al., 2006b). Li et al. (2006) mapped these resistance genes to the same linkage group M (chromosome 07) in soybean, suggesting that these two genes may be allelic.

In Ohio, PI 243540 was found to possess antibiosis resistance against biotype 2 (Kang et al., 2008). Genetic linkage map approach identified the source of this aphid resistance on linkage group F (chromosome 13), named *Rag2* (Mian et al., 2008). PI 200538 is another source of *Rag2*, with a QTL mapped on the same region (Hill et al., 2009). PI 567301B had a locus that mapped near the *Rag2* region, but it showed antixenosis resistance indicating that the resistance may not be controlled by the same gene as *Rag2*. It was tentatively named *Rag5* (Jun et al., 2012).

Mensah et al. (2005) identified different SBA resistance sources by screening 2,147 soybean Chinese accessions from maturity group (MG) 0 to III. In this study, two accessions PI 567541B and PI 567598B, conferred antibiosis resistance and two others, PI 567543C and PI 567597C, had antixenosis resistance. Genetic mapping studies on PI 567541B and PI 567598B showed that resistance in each accession was controlled by two recessive genes (Mensah et al., 2008). The recessive genes for PI 567541B mapped on chromosomes 7 and 13 (LG M and F), and were named *rag1c* and *rag4* (Zhang et al., 2008). The QTL found on chromosome 7 of PI 567541B mapped on the same region as the *Rag1* in Dowling and *Rag* in Jackson. This finding suggested that these three genes were tightly linked, but it still isn't known if they are multiple alleles for the same gene. A single dominant gene was found to control antixenosis resistance in PI 567543C. It was mapped on chromosome 16 (LG J) and was designated *Rag3* (Zhang et al., 2010). The same QTL on chromosome 16 region (*Rag3b*) was later found by Zhang et al. (2013) controlling aphid resistance in PI 567537.

In a Chinese line 'Zhongdou 27', two genes controlling SBA resistance (*qRa_1* and *qRa_2*) were mapped on chromosomes 8 and 13. Both loci are also highly associated with high isoflavone content (Meng et al., 2011). The QTL *qRa_2* was associated with a different region in chromosome 13 than the QTLs found by Mian et al. (2008) and Zhang et al. (2008).

Another line from China, P203, contained a major antixenosis QTL on chromosome 8 (Xiao et al., 2013). Using SSR markers, the gene on chromosome 8, named as [*Rag6*]*_P203*, was narrowed down into a physical distance of 192 kb. A single candidate gene annotated as serine/threonine protein kinase was found within the interval. The percent phenotypic variation controlled by [*Rag6*]*_P203* was not reported.

A study conducted by Jun et al (2013) found three QTL in PI 567324 which conferred resistance against biotypes 1 and 2. Two major QTL were identified on chromosome 13 and one minor QTL on chromosome 6. This is the first report for a QTL found on chromosome 6. The two QTL on chromosome 13 coincided with the location of previous known *Rag2* (Mian et al., 2008; Hill et al., 2009) and *rag4* intervals (Zhang et al., 2008).

Although several *rag* genes have been genetically mapped, significant efforts are still ongoing for screening and identifying new resistant sources (Bansal et al.; Bhusal et al., 2013; Fox et al., 2013; Hesler, 2013). Further genetic characterization and mapping of the new loci or genes that control the soybean aphid resistance will be valuable to identify novel genes for pyramiding, to counteract the identified aphid biotypes. Another important research area to investigate is the efficacy of the *rag* loci when combined or pyramided. Thus far, no studies reported an investigation of the allelic interaction of the common QTLs (i.e. chromosomes 7, 13 and 16) found on different sources.

**Soybean aphid biotypes**

Diehl and Bush (1984) defined insect biotypes as "parasite or parasitoids distinguished by survival and development on a particular host or by host preference for feeding, oviposition or both." The isolation of insect populations may involve genetic differences due to their host preference.

Kim et al. (2008) reported that several SBA biotypes existed in the northern US. They termed this the Illinois and Ohio biotypes. A list of known soybean aphid biotypes is given in Table 1.2. Biotype 1, collected in Illinois, is effectively controlled by *Rag1* and *Rag*; while biotype 2 originating from Ohio can overcome this resistance.

In 2010, biotype 3 was collected from an overwintering host glossy buckthorn (*Frangula alnus*) and was able to overcome *Rag2* but not the *Rag1* (Hill et al., 2010). Finally, a fourth biotype was recently identified that readily colonizes *Rag1* and *Rag2*, as well as soybean lines with a pyramid of both genes (Alt and Ryan-Mahmutagic, 2013). To date, no biotype has been identified which overcomes *Rag3*. The resistance found on PI 567598B and *rag1c*/*rag4* from PI 567541B has proven to be effective or moderately effective under all biotypes. (Alt and Ryan-Mahmutagic, 2013).

The presence of SBA biotypes demonstrates the importance of continually identifying new SBA resistance genes and pyramiding these into commercial cultivars to achieve durable resistance. Soybean aphids will continue to be a threat in soybean-producing states and sustainable solutions must be identified to manage the pest.

**Table 1.2** Current list of soybean aphid biotypes and their virulence relationships to published soybean aphid resistant lines.

| Biotype | Source isolate | Soybean genotype/ Virulence relationship | | | | | | | Reference |
|---|---|---|---|---|---|---|---|---|---|
| | | *Rag1* Dowling | *Rag2* PI 243540 | *Rag2* PI 200538 | *Rag2/5?* PI 567301B | *Rag3* PI 567543C | *rag3,rag1b* PI 567598B | *rag4,rag1c* PI 567541B | |
| 1 | Illinois | A | A | A | A | A | A | A | (Kim et al., 2008) |
| 2 | Ohio | V | A | A | A | A | A | A | (Kim et al., 2008) |
| 3 | Indiana | A | V | V | ND | A | A | MV | (Alt and Ryan-Mahmutagic, 2013; Hill et al., 2010) |
| 4 | Wisconsin | V | V | V | ND | A | MV | MV | (Alt and Ryan-Mahmutagic, 2013) |

A = avirulent
V = virulent
MV = moderately virulent
ND = no data

**Genetic approaches to improve soybean aphid resistance**

Since 2010, aphid-resistant varieties have been available to growers albeit a smaller proportion; understanding the genetic interaction between the soybean aphid and naturally-occurring resistant lines are key to the sustainable management of SBA (Ragsdale et al., 2011).

Molecular markers have been used as effective tools to complement the selection process from thousands of plant breeding lines and down to only a few that contain the traits or genes of interest. To aid in precise marker-assisted selection, there is a need to develop a genetic fine map that will identify the location of markers that flank targeted genes and are within one or fewer centiMorgans (cM) apart (Bennetzen, 1999). Genetic fine mapping will also serve as the initial step in map-based gene cloning to be able to characterize each of the identified SBA resistant genes and eventually use them for soybean improvement by transgenic technology. Furthermore, this is also an important approach to determine the allelic relationships of the QTL from different sources that have been mapped on the same chromosome regions. Most populations used for fine mapping studies are backcross and $F_2$ lines due to the abundance of recombination events during the early generations than for advanced lines in later generations where most loci in the genome become more homozygous.

Other reliable and efficient approaches in fine mapping of QTLs are the use of near isogenic lines (NILs) (Kaeppler et al., 1993) and a mapping population derived from a residual heterozygous line (RHL) proposed initially in sorghum by Tuinstra et al. (1997) and used in soybeans by Yamanaka et al. (2005). The near isogenic line approach is most ideal to dissect the genetic basis of a trait controlled by a QTL which may be masked by epistatic interactions of other loci that are segregating in the genetic background. By backcrossing for several

generations, the genome of lines are almost 100% similar to the recurrent parent except for a small fragment introgressed from the donor parent expressing the phenotype of interest, thus near isogenic. However, the backcross procedure takes considerable amount of time and labor for marker-assisted selection.

An alternative approach to producing NILs is through a selfing scheme of a recombinant inbred line (RIL) population that is expected to be mostly homozygous for other regions of the genome (expected homozygosity in $F_9$ RILs: 99.6% homozygous) and can harbor a heterozygous region where the target QTL is located, although of lesser probability (Fehr et al., 1987; Allard, 1999). RHLs that are heterozygous for the QTL of interest, but homozygous at other loci, are propagated to produce families called heterozygous inbred families (HIFs) that differ only in the genotype at that QTL region (Anderson and Mitchell-Olds, 2011). RHL progenies forming these families will show a simple phenotypic segregation based on the effects of the target QTL at the heterozygous region (Watanabe et al., 2011). The RHL strategy has already been used to identify loci underlying pathogen resistance in soybean (Meksem et al., 1999; Njiti et al., 1998; Triwitayakorn et al., 2005).

**Fine mapping of *Rag* genes**

Among the identified QTLs for SBA resistance, three have been fine-mapped so far. Kim et al. (2009) fine-mapped *Rag1* from Dowling to a 115-kb interval from 824 $BC_4F_2$ and 1,000 $BC_4F_3$ plants. *Rag2* from PI 200538, on the other hand, was mapped to a 54-kb interval using lines derived from 5,782 $F_2$ plants (Kim et al., 2010). From the fine mapping studies, SNP

markers were discovered and found useful for marker-assisted selection of aphid resistant lines derived from Dowling and PI 200538. SSR markers have been used to delimit *[Rag6]_P203* into a 192-kb interval (Xiao et al., 2013). Candidate genes were identified and map-based cloning can be done to characterize and compare the function of these SBA putative resistant genes. Two out of the 13 predicted genes in the 115-kb interval of *Rag1* (Kim et al., 2009) were identified to be potential candidate genes while one out of seven predicted genes deduced from the 54-kb interval of *Rag2* was a potential candidate gene (Kim et al., 2010).

**Candidate genes for aphid resistance**

All of the identified candidate genes of *Rag1*, *Rag2* and *[Rag6]_P203* are nucleotide-binding site leucine-rich repeat (NBS-LRR) genes. This is similar to the *Mi* gene from tomato which confers resistance against potato aphid (Rossi et al., 1998; Kaloshian et al., 2000; Cooper et al., 2004); other examples include the *Vat* gene from melons against the melon aphid (Chen et al., 1997; Villada et al., 2009) and the *AKR* gene underlying resistance to blue-green aphids in *Medicago truncatula* (Klingler et al., 2005).

In a transcriptome response study done by Studham and MacIntosh (2013), several hundred transcripts were involved in response to aphid infestation in the susceptible plant. The resistant plant containing *Rag1* only had one transcript induced by aphid infestation. The study concluded that defense-related transcripts are expressed constitutively in resistant plants while they are suppressed during aphid infestation in susceptible plants. The transcripts induced by soybean aphids in susceptible plants were reported to regulate hormone signaling pathways such as abscissic acid (ABA) and ethylene (ET) that can suppress defense-signaling pathways controlled by salicylic acid (SA) and jasmonic acid (JA).

Soybean aphid resistance may involve negative feedback regulation but molecular genetic characterization studies will have to be done to prove this. In Arabidopsis, CPR30 (an F-box gene) was found to negatively regulate defense-related genes EDS1, PAD4 and NDR1 against pathogens (Gou et al., 2009). The loss of function of the CPR1/CPR30 gene increased the levels of SNC1 proteins, controlled by a NBS-LRR gene, and triggering a defense-signaling cascade; this interaction implies that the F-box gene is involved in the stability of NBS-LRR genes (Gou et al., 2009, 2012). It was also reported that CPR1/CPR30 regulated both SA-dependent and SA-independent pathways and may not only interact with SNC1 but other R proteins as well (Gou et al., 2012).

Other crops with aphid-resistant varieties are *Medicago*, barley, peach, peanut, lettuce, apple, maize and wheat. Only a few genes have been mapped and characterized: *Nr* gene in lettuce/lettuce aphid (van Helden and Tjallingii, 1993; McCreight, 2008), *Sd1* in apple /rosy-leaf curling aphid (Roche et al., 1997), *Aph* gene in maize/ corn leaf aphid (So et al., 2010) and eight *Dn* genes in wheat/Russian wheat aphid (Nkongolo et al., 1991; Boyko et al., 2006; Valdez et al., 2012). The *AKR* gene conferring resistance against blue-green aphid in *Medicago truncatula* 'Jester' was mapped to a region flanked by NBS-LRR genes and F-box domain-coding genes (Klingler et al., 2005). Other QTLs in *M*. *truncatula* confer resistance against cowpea aphid (Kamphuis et al., 2012) and spotted alfalfa aphid (Kamphuis et al., 2013).

**Genetic and genomic tools for soybean improvement**

The advent of the genomic revolution has paved the way to discovering new tools to make plant breeding methods more precise and time-efficient. The use of molecular markers are important in the discovery of new traits leading to the characterization of genes that control a

phenotype or traits of interest. Molecular markers are also important tools in the efficient integration of traits and genes into breeding lines or elite cultivars. In the soybean breeding community, simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) are the most popular molecular markers being used (www.soybase.org). Abundant SSR markers have been used for constructing a consensus genetic linkage map of soybean and understanding the genetic basis of important soybean traits for breeding (Song et al., 2004, 2010; Ott et al., 2011; Sayama et al., 2011).

In 2010, a high-quality soybean reference genome was sequenced using a whole-genome shotgun approach (Schmutz et al., 2010). With the availability of a reference genome, re-sequencing of other soybean genome cultivars or breeding lines are being done for lower cost and shorter time to identify new SNPs unique to the germplasm, such as the development of the SoySNP50 genotyping array (Song et al., 2013). Significant progress were done in the discovery of SNP markers using next generation sequence analysis for mapping of QTLs controlling a trait, genome-wide association mapping studies, and for marker-assisted breeding (Chaisan et al., 2010; Lam et al., 2010; Wu et al., 2010; Song et al., 2013).

NGS methods are also being used to develop transcriptomic and genomic resources for identifying SSRs and SNPs unique to the soybean aphid genome (Bai et al., 2010). Biotype-specific markers can be used to potentially diagnose and differentiate SBA biotypes between locations or fields. This can be a useful tool for recommendations of soybean seeds containing different *rag* genes.

The present study aimed to locate the recessive genes in PI 567598B, a broader source of aphid resistance, using linkage map analysis and fine mapping approaches. Using available SSR and SNP molecular markers from genotyping assays and SNP arrays, a genetic map of the

genomic location of the two recessive genes was constructed. Two approaches for the fine mapping study were used: preliminary fine mapping utilized lines derived from $F_2$ plants and the other approach used a mapping population derived from residual heterozygous lines (RHL). Candidate genes for *rag1b* and *rag3* were identified through these fine mapping approaches. This process also identified SNP markers that co-segregated with soybean aphid resistance found in PI 567598B that can be used for precise marker-assisted breeding and stacking with other sources of *Rag*/*rag* genes for durable resistance in the field.

**REFERENCES**

## REFERENCES


Allard, R.W. (1999). Principles of plant breeding (John Wiley & Sons, Inc: New York, NY).

Alt, J. and Ryan-Mahmutagic, M. (2013). Soybean aphid biotype 4 identified. Crop Sci. 53: 1491.

Anderson, J.T. and Mitchell-Olds, T. (2011). Ecological genetics and genomics of plant defences: evidence and approaches. Funct. Ecol. 25: 312–324.

Bai, X., Zhang, W., Orantes, L., Jun, T.-H., Mittapalli, O., Mian, M.R., and Michel, A.P. (2010). Combining next-generation sequencing strategies for rapid molecular resource development from an invasive aphid species, Aphis glycines. PLoS One 5: e11370.

Bansal, R., Mian, M.A.R., and Michel, A.P. Identification of novel sources of host plant resistance to known soybean aphid biotypes. J. Econ. Entomol. 106: 1479–1485.

Bennetzen, J. (1999). Genetic fine-mapping technology (In:Training manual for a seminar held by International Crops Research for the Semi-Arid Tropics (ICRISAT). Haussmann, B.I.G., H.H. Geiger, D.E. Hess, C.T. Hash and P. Bramel-Cox (eds): Patancheru 502 324, Andhra Pradesh, India).

Bhusal, S.J., Jiang, G.-L., Tilmon, K.J., and Hesler, L.S. (2013). Identification of soybean aphid resistance in early maturing genotypes of soybean. Crop Sci. 53: 491–499.

Boyko, E.V., Smith, C.M., Thara, V.K., Bruno, J.M., Deng, Y., Starkey, S.R., and Klaahsen, D.L. (2006). Molecular basis of plant gene expression during aphid invasion: wheat *Pto*- and *Pti*-like sequences are involved in interactions between wheat and Russian wheat aphid (Homoptera: *Aphididae*). J. Econ. Entomol. 99: 1430–1445.

Catangui, M.A., Beckendorf, E.A., and Riedell, W.E. (2009). Soybean aphid population dynamics, soybean yield loss, and development of stage-specific economic injury levels. Agron. J. 101: 1080.

Chaisan, T., Van, K., Kim, M.Y., Kim, K.D., Choi, B.-S., and Lee, S.-H. (2010). *In silico* single nucleotide polymorphism discovery and application to marker-assisted selection in soybean. Mol. Breed. 29: 221–233.

Chen, J.-Q., Martin B, Rahbe Y, and Fereres A (1997). Early intracellular punctures by two aphid species on near-isogenic melon lines with and without the virus aphid transmission (*Vat*) resistance gene. Eur. J. Plant Pathol. 103: 521–536.

Cooper, W.C., Jia, L., and Goggin, F.L. (2004). Acquired and R-gene-mediated resistance against the potato aphid in tomato. J. Chem. Ecol. 30: 2527–2542.

Diehl, S.R. and Bush, G.L. (1984). An evolutionary and applied perspective of insect biotypes. Annu. Rev. Entomol. 29: 471–504.

DiFonzo, C. (2005). Soybean aphid control using insecticides (Michigan State University: Michigan).

DiFonzo, C.D. and Hines, R. (2002). Soybean aphid in Michigan: update from the 2001 season (Michigan State University: East Lansing, MI).

Fehr, W.R., Fehr, E.L., and Jessen, H.J. (1987). Principles of Cultivar Development: Theory and Technique (Macmillan Pub Co).

Fox, C.M., Kim, K.-S., Cregan, P.B., Hill, C.B., Hartman, G.L., and Diers, B.W. (2013). Inheritance of soybean aphid resistance in 21 soybean plant introductions. Theor. Appl. Genet.: 1–8.

Gou, M., Shi, Z., Zhu, Y., Bao, Z., Wang, G., and Hua, J. (2012). The F-box protein CPR1/CPR30 negatively regulates R protein SNC1 accumulation. Plant J. 69: 411–420.

Gou, M., Su, N., Zheng, J., Huai, J., Wu, G., Zhao, J., He, J., Tang, D., Yang, S., and Wang, G. (2009). An F-box gene, CPR30, functions as a negative regulator of the defense response in Arabidopsis. Plant J. 60: 757–770.

Hartman, G.L., Domier, L.L., Wax, L.M., Helm, C.G., and Onstad, D.W. (2001). Occurrence and distribution of *Aphis glycines* on soybeans in Illinois in 2000 and its potential control.

Van Helden, M. and Tjallingii, W.F. (1993). Tissue localisation of lettuce resistance to the aphid *Nasonovia ribisnigri* using electrical penetration graphs. Entomol. Exp. Appl. 68: 269–278.

Hesler, L.S. et al. (2013). Performance and prospects of Rag genes for management of soybean aphid. Entomol. Exp. Appl. 147: 201–216.

Hesler, L.S. (2013). Resistance to soybean aphid among wild soybean lines under controlled conditions. Crop Prot. 53: 139–146.

Hill, C.B., Crull, L., Herman, T.K., Voegtlin, D.J., and Hartman, G.L. (2010). A new soybean aphid (Hemiptera: *Aphididae*) biotype identified. J. Econ. Entomol. 103: 509–515.

Hill, C.B., Kim, K.-S., Crull, L., Diers, B.W., and Hartman, G.L. (2009). Inheritance of resistance to the soybean aphid in soybean PI 200538. Crop Sci. 49: 1193.

Hill, C.B., Li, Y., and Hartman, G.L. (2006a). A single dominant gene for resistance to the soybean aphid in the soybean cultivar dowling. Crop Sci. 46: 1601.

Hill, C.B., Li, Y., and Hartman, G.L. (2004). Resistance to the soybean aphid in soybean germplasm. Crop Sci 44: 98–106.

Hill, C.B., Li, Y., and Hartman, G.L. (2006b). Soybean aphid resistance in soybean jackson is controlled by a single dominant gene. Crop Sci. 46: 1606.

Hill, J.H., Alleman, R., Hogg, D.B., and Grau, C.R. (2001). First report of transmission of soybean mosaic virus and alfalfa mosaic virus by *Aphis glycines* in the new world. Plant Dis. 85: 561–561.

Johnson, K.D., O'Neal, M.E., Ragsdale, D.W., Difonzo, C.D., Swinton, S.M., Dixon, P.M., Potter, B.D., Hodgson, E.W., and Costamagna, A.C. (2009). Probability of cost-effective management of soybean aphid (Hemiptera: *Aphididae*) in North America. J. Econ. Entomol. 102: 2101–2108.

Jun, T.-H., Mian, M.R., and Michel, A.P. (2012). Genetic mapping revealed two loci for soybean aphid resistance in PI 567301B. Theor. Appl. Genet. 124: 13–22.

Jun, T.-H., Rouf Mian, M.A., and Michel, A.P. (2013). Genetic mapping of three quantitative trait loci for soybean aphid resistance in PI 567324. Heredity 111: 16–22.

Kaeppler, S.M., Phillips, R.L., and Kim, T.S. (1993). Use of near-isogenic lines derived by backcrossing or selfing to map qualitative traits. Theor. Appl. Genet. 87: 233–237.

Kaloshian, I., Kinsey, M.G., Williamson, V.M., and Ullman, D.E. (2000). *Mi*-mediated resistance against the potato aphid *Macrosiphum euphorbiae* (Hemiptera: *Aphididae*) limits sieve element ingestion. Environ. Entomol. 29: 690–695.

Kamphuis, L.G., Gao, L., and Singh, K.B. (2012). Identification and characterization of resistance to cowpea aphid (*Aphis craccivora* Koch) in *Medicago truncatula*. BMC Plant Biol. 12: 101.

Kamphuis, L.G., Lichtenzveig, J., Peng, K., Guo, S.-M., Klingler, J.P., Siddique, K.H.M., Gao, L.-L., and Singh, K.B. (2013). Characterization and genetic dissection of resistance to spotted alfalfa aphid (*Therioaphis trifolii)* in Medicago truncatula. J. Exp. Bot.: ert305.

Kang, S.-T., Rouf Mian, M.A., and Hammond, R.B. (2008). Soybean aphid resistance in PI 243540 is controlled by a single dominant gene. Crop Sci. 48: 1744.

Kim, K.-S., Bellendir, S., Hudson, K.A., Hill, C.B., Hartman, G.L., Hyten, D.L., Hudson, M.E., and Diers, B.W. (2009). Fine mapping the soybean aphid resistance gene *Rag1* in soybean. Theor. Appl. Genet. 120: 1063–1071.

Kim, K.-S., Hill, C.B., Hartman, G.L., Hyten, D.L., Hudson, M.E., and Diers, B.W. (2010). Fine mapping of the soybean aphid-resistance gene *Rag2* in soybean PI 200538. Theor. Appl. Genet. 121: 599–610.

Kim, K.-S., Hill, C.B., Hartman, G.L., Mian, M.A.R., and Diers, B.W. (2008). Discovery of Soybean Aphid Biotypes. Crop Sci. 48: 923–928.

Klingler, J., Creasy, R., Gao, L., Nair, R.M., Calix, A.S., Jacob, H.S., Edwards, O.R., and Singh, K.B. (2005). Aphid resistance in *Medicago truncatula* involves antixenosis and phloem-specific, inducible antibiosis, and maps to a single locus flanked by NBS-LRR resistance gene analogs. Plant Physiol. 137: 1445–1455.

Lam, H.-M. et al. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat Genet 42: 1053–1059.

Li, Y., Hill, C.B., Carlson, S.R., Diers, B.W., and Hartman, G.L. (2006). Soybean aphid resistance genes in the soybean cultivars Dowling and Jackson map to linkage group M. Mol. Breed. 19: 25–34.

Li, Y., Hill, C.B., and Hartman, G.L. (2004). Effect of three resistant soybean genotypes on the fecundity, mortality, and maturation of soybean aphid (Homoptera: *Aphididae*). J. Econ. Entomol. 97: 1106–1111.

Li, Y., Zou, J., Li, M., Bilgin, D.D., Vodkin, L.O., Hartman, G.L., and Clough, S.J. (2008). Soybean defense responses to the soybean aphid. New Phytol. 179: 185–195.

Losey, J.E., Waldron, J.K., Hoebeke, E.R., Macomber, L.E., and Scott, B.N. (2002). First record of the soybean aphid, *Aphis glycines* Matsumura (Hemiptera: Sternorrhyncha: Aphididae), in New York. Gt. Lakes Entomol. 35: 101–105.

Mahr, D.L., Whitaker, P., and Ridgway, N.M. (2008). Biological control of insects and mites: An introduction to beneficial natural enemies and their use in pest management (University of Wisconsin).

McCarville, M.T. and O'Neal, M.E. (2013). Soybean aphid (Aphididae: *Hemiptera*) population growth as affected by host plant resistance and an insecticidal seed treatment. J. Econ. Entomol. 106: 1302–1309.

Mccornack, B.P., Ragsdale, D.W., and Venette, R.C. (2004). Demography of soybean aphid (Homoptera: *Aphididae*) at summer temperatures. J. Econ. Entomol. 97: 854–861.

McCreight, J.D. (2008). Potential sources of genetic resistance in lactuca spp. To the lettuce aphid, *Nasanovia ribisnigri* (mosely) (Homoptera: *Aphididae*). HortScience 43: 1355–1358.

Meksem, K., Doubler, T.W., Chancharoenchai, K., Nijti, N., Chang, S.J., Arelli, A.P.R., Cregan, P.E., Gray, L.E., Gibson, P.T., and Lightfoot, D.A. (1999). Clustering among loci underlying soybean resistance to *Fusarium solani*, SDS and SCN in near-isogenic lines. Theor. Appl. Genet. 99: 1131–1142.

Meng, F., Han, Y., Teng, W., Li, Y., and Li, W. (2011). QTL underlying the resistance to soybean aphid (*Aphis glycines* Matsumura) through isoflavone-mediated antibiosis in soybean cultivar "Zhongdou 27." Theor. Appl. Genet.

Mensah, C., DiFonzo, C., Nelson, R.L., and Wang, D. (2005). Resistance to soybean aphid in early maturing soybean germplasm. Crop Sci. 45: 2228.

Mensah, C., DiFonzo, C., and Wang, D. (2008). Inheritance of soybean aphid resistance in PI 567541B and PI 567598B. Crop Sci. 48: 1759.

Mian, M.A.R., Kang, S.-T., Beil, S.E., and Hammond, R.B. (2008). Genetic linkage mapping of the soybean aphid resistance gene in PI 243540. Theor. Appl. Genet. 117: 955–962.

Michigan Soybean Promotion Committee (2012). Michigan Soybean Checkoff.

Nielsen, C. and Hajek, A.E. (2005). Control of invasive soybean aphid, Aphis glycines (Hemiptera: Aphididae), populations by existing natural enemies in New York state, with emphasis on entomopathogenic fungi. Environ. Entomol. 34: 1036–1047.

Njiti, V.N., Doubler, T.W., Suttner, R.J., Gray, L.E., Gibson, P.T., and Lightfoot, D.A. (1998). Resistance to soybean sudden death syndrome and root colonization by *Fusarium solani* f. sp. *glycine* in near-isogenic lines. Crop Sci. 38: 472.

Nkongolo, K.K., Quick, J.S., Limin, A.E., and Fowler, D.B. (1991). Sources and inheritance of resistance to Russian wheat aphid in *Triticum* species amphiploids and *Triticum tauschii*. Can. J. Plant Sci. 71: 703–708.

Ohnesorg, W.J., Johnson, K.D., and O'Neal, M.E. (2009). Impact of reduced-risk insecticides on soybean aphid and associated natural enemies. J. Econ. Entomol. 102: 1816–1826.

Ostlie, K. (2002). Managing Soybean Aphid. University of Minnesota Extension report. www.soybeans.umn.edu

Ott, A., Trautschold, B., and Sandhu, D. (2011). Using microsatellites to understand the physical distribution of recombination on soybean chromosomes. PLoS One 6(7):e22306.

Painter, R.H. (1968). Insect resistance in crop plants. Annu. Rev. Entomol. 3: 276–290.

Ragsdale, D.W. et al. (2007). Economic threshold for soybean aphid (Hemiptera: *Aphididae*). J. Econ. Entomol. 100: 1258–1267.

Ragsdale, D.W., Landis, D.A., Brodeur, J., Heimpel, G.E., and Desneux, N. (2011). Ecology and management of the soybean aphid in North America. Annu. Rev. Entomol. 56: 375–399.

Ragsdale, D.W., Voegtlin, D.J., and O'Neil, R.J. (2004). Soybean Aphid Biology in North America. Ann. Entomol. Soc. Am. 97: 204–208.

Roche, P., Alston, F.H., Maliepaard, C., Evans, K.M., Vrielink, R., Dunemann, F., Markussen, T., Tartarini, S., Brown, L.M., Ryder, C., and King, G.J. (1997). RFLP and RAPD markers linked to the rosy leaf curling aphid resistance gene (*Sd1*) in apple. Theor. Appl. Genet. 94: 528–533.

Rossi, M., Goggin, F.L., Milligan, S.B., Kaloshian, I., Ullman, D.E., and Williamson, V.M. (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc. Natl. Acad. Sci. 95: 9750 –9754.

Sayama, T., Hwang, T.-Y., Komatsu, K., Takada, Y., Takahashi, M., Kato, S., Sasama, H., Higashi, A., Nakamoto, Y., Funatsuki, H., and Ishimoto, M. (2011). Development and application of a whole-genome simple sequence repeat panel for high-throughput genotyping in soybean. DNA Res. 18: 107 –115.

Schmidt, N.P. et al. (2012). Spatial Distribution of Aphis glycines (Hemiptera: Aphididae): A Summary of the Suction Trap Network. J. Econ. Entomol. 105: 259–271.

Schmutz, J. et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183.

So, Y.-S., Ji, H.C., and Brewbaker, J.L. (2010). Resistance to corn leaf aphid (*Rhopalosiphum maidis* Fitch) in tropical corn *(Zea mays* L.). Euphytica 172: 373–381.

Song, Q., Hyten, D.L., Jia, G., Quigley, C.V., Fickus, E.W., Nelson, R.L., and Cregan, P.B. (2013). Development and evaluation of SOYSNP50k, a high-density genotyping array for soybean. PLoS ONE 8: e54985.

Song, Q., Jia, G., Zhu, Y., Grant, D., Nelson, R.T., Hwang, E.-Y., Hyten, D.L., and Cregan, P.B. (2010). Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. Crop Sci. 50: 1950–1960.

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

Studham, M.E. and MacIntosh, G.C. (2013). Multiple phytohormone signals control the transcriptional response to soybean aphid infestation in susceptible and resistant soybean plants. Mol. Plant. Microbe Interact. 26: 116–129.

Triwitayakorn, K., Njiti, V.N., Iqbal, M.J., Yaegashi, S., Town, C., and Lightfoot, D.A. (2005). Genomic analysis of a region encompassing QRfs1 and QRfs2: genes that underlie soybean resistance to sudden death syndrome. Genome Natl. Res. Counc. Can. Génome Cons. Natl. Rech. Can. 48: 125–138.

Tuinstra, M.R., Ejeta, G., and Goldsbrough, P.B. (1997). Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. Theor. Appl. Genet. 95: 1005–1011.

Valdez, V.A., Byrne, P.F., Lapitan, N.L.V., Peairs, F.B., Bernardo, A., Bai, G., and Haley, S.D. (2012). Inheritance and genetic mapping of russian wheat aphid resistance in iranian wheat landrace accession PI 626580. Crop Sci. 52: 676.

Venette, R.C. and Ragsdale, D.W. (2004). Assessing the invasion by soybean aphid (Homoptera: *Aphididae*): Where Will It End? Ann. Entomol. Soc. Am. 97: 219–226.

Villada, E.S., González, E.G., López-Sesé, A.I., Castiel, A.F., and Gómez-Guillamón, M.L. (2009). Hypersensitive response to *Aphis gossypii* Glover in melon genotypes carrying the Vat gene. J. Exp. Bot. 60: 3269 –3277.

Voegtlin, D.J., O'Neil, R.J., and Graves, W.R. (2004). Tests of suitability of overwintering hosts of aphis glycines: identification of a new host association with *Rhamnus alnifolia* L'Héritier. Ann. Entomol. Soc. Am. 97: 233–234.

Watanabe, S., Xia, Z., Hideshima, R., Tsubokura, Y., Sato, S., Yamanaka, N., Takahashi, R., Anai, T., Tabata, S., Kitamura, K., and Harada, K. (2011). A map-based cloning strategy employing a residual heterozygous line reveals that the GIGANTEA gene is involved in soybean maturity and flowering. Genetics 188: 395–407.

Wu, X., Ren, C., Joshi, T., Vuong, T., Xu, D., and Nguyen, H. (2010). SNP discovery by high-throughput sequencing in soybean. BMC Genomics 11: 469.

www.soystats.com (2012). www.soystats.com. Soy Stats 2012.

Xiao, L., Hu, Y., Wang, B., and Wu, T. (2013). Genetic mapping of a novel gene for soybean aphid resistance in soybean (Glycine max [L.] Merr.) line P203 from China. Theor. Appl. Genet.126(9): 2279-87.

Yamanaka, N., Watanabe, S., Toda, K., Hayashi, M., Fuchigami, H., Takahashi, R., and Harada, K. (2005). Fine mapping of the FT1 locus for soybean flowering time using a residual heterozygous line derived from a recombinant inbred line. Theor. Appl. Genet. 110: 634–639.

Zhang, G., Gu, C., and Wang, D. (2010). A novel locus for soybean aphid resistance. Theor. Appl. Genet. 120: 1183–1191.

Zhang, G., Gu, C., and Wang, D. (2013). Mapping and validation of a gene for soybean aphid resistance in PI 567537. Mol. Breed. 32: 131–138.

Zhang, G., Gu, C., and Wang, D. (2008). Molecular mapping of soybean aphid resistance genes in PI 567541B. Theor. Appl. Genet. 118: 473–482.

# CHAPTER II

## MAPPING SOYBEAN APHID RESISTANCE GENES IN PI 567598B

Revised from: Bales C, Zhang G, Liu M, Mensah C, Gu C, Song Q, Hyten D, Cregan P, Wang D (2013) Mapping soybean aphid resistance genes in PI 567598B. Theoretical and Applied Genetics 126:2081–2091

**Abstract**

The soybean aphid (*Aphis glycines* Matsumura) has become a major pest of soybean [*Glycine max* (L.) Merr.] in North America since it was first reported in 2000. Our previous study revealed that the strong aphid resistance of plant introduction (PI) 567598B was controlled by two recessive genes. The objective of this study was to locate these two genes on the soybean genetic linkage map using molecular markers. A mapping population of 282 $F_{4:5}$ lines derived from IA2070 x E06902 was evaluated for aphid resistance in a field trial in 2009 and a greenhouse trial in 2010. Two quantitative trait loci (QTLs) were identified using the composite and multiple interval mapping methods, and were mapped on chromosomes 7 (linkage group M) and 16 (linkage group J), respectively. E06902, a parent derived from PI 567598B, conferred resistance at both loci. In the 2010 greenhouse trial, each of the two QTLs explained over 30 % of the phenotypic variation. Significant epistatic interaction was also found between these two QTLs. However, in the 2009 field trial, only the QTL on chromosome 16 was found and it explained 56.1 % of the phenotypic variation. These two QTLs and their interaction were confirmed with another population consisting of 94 $F_{2:5}$ lines in the 2008 and 2009 greenhouse

trials. For both trials in the alternative population, these two loci explained about 50 and 80.4 % of the total phenotypic variation, respectively. Our study shows that the soybean aphid isolate used in the 2009 field trial overcame the QTL found on chromosome 7. Presence of the QTL on chromosome 16 conferred soybean aphid resistance in all trials. The markers linked to the aphid-resistant QTLs in PI 567598B or its derived lines can be used in marker-assisted breeding for aphid resistance.

**Introduction**

The soybean aphid is one of the most damaging pests on soybean. It can reduce yield either by feeding directly on soybean or transmitting various viruses (Wu et al. 2004); such loss can reach up to 88 %. Soybean aphids can also affect seed quality by reducing the oil content (Beckendorf et al. 2008).

Host resistance is considered an effective, economical, and environmentally friendly means for pest control. There are two types of host resistance to insects: antibiosis and antixenosis (Painter 1951). Antibiosis affects insect biology and reduces insect populations. Antixenosis affects insect behavior and is expressed as non-preference for certain plants. Researchers in the US have identified several aphid-resistant germplasm accessions (Hill et al. 2004; Mensah et al. 2005; Diaz-Montano et al. 2006; Hesler et al. 2007; Hesler and Dashiell 2008; Mian et al. 2008a). Genetic studies have shown that the antibiosis resistance in Dowling and Jackson were both controlled by a single dominant gene (Hill et al. 2006a, b). The gene in Dowling was named *Rag1* (Hill et al. 2006a). Later, *Rag1* and the resistance gene (*Rag*) in Jackson were both mapped in the same genomic region on chromosome 7 [linkage group (LG) M] (Li et al. 2007). Similarly, a single dominant gene, *Rag2*, controlled antibiosis resistance in PI 243540 (Kang et al. 2008) was mapped on chromosome 13 (LG F) (Mian et al. 2008b). A

single dominant gene controlling the antibiosis resistance in PI 200538 was mapped in the same genomic region as *Rag2* (Hill et al. 2009). A codominant gene, *Rag3*, on chromosome 16 (LG J) controls the antixenosis resistance in PI 567543C (Zhang et al. 2010). However, the antibiosis resistance in both PI 567541B and PI 567598B is controlled by two recessive genes (Mensah et al. 2008). A genetic mapping study located the two genes in PI 567541B on chromosomes 7 and 13 (LG M and F) (Zhang et al. 2009). The gene on chromosome 7 (LG M) was mapped in the same genomic region as *Rag1* and was later designated *rag1_c*. The gene on chromosome 13 (LG F) was located far from *Rag2* and was later designated *rag4* (Zhang et al. 2009). Significant epistatic interaction was also found between the two genes identified in PI 567541B (Zhang et al. 2009).

Dominant and recessive genes were found to control aphid resistance in other crops such as cowpea, barley, peach, wheat, corn and peanut. The aphid (*Aphis craccivora* Koch) resistance in cowpea (*Vigna unguiculata* L.) involves a single dominant gene (Pathak 1988). The aphid resistance in spring barley (*Hordeum vulgare* L.) is controlled by two dominant genes (Mornhinweg et al. 2002). In the peach cultivar 'Rubira', the resistance to the green peach aphid (*Myzus persicae*) is controlled by a single dominant gene (Pascal et al. 2002). In wheat (*Triticum spp*.), eight independent dominant genes each confer resistance to the Russian wheat aphid (*Diuraphis noxia*) from different resistance sources (Liu et al. 2005), while one recessive gene contributes to the resistance in *Triticum tauschii* line SQ24 (Nkongolo et al. 1991). A single recessive gene was also found to control resistance to corn leaf aphid (*Rhopalosiphum maidis* Fitch) (So et al. 2010) and the groundnut rosette disease vector, *Aphis craccivora*, infesting peanut (Herselman et al. 2004).

Commercial varieties with *Rag1* have been available in the US. However, at least three

biotypes of soybean aphid have been identified and *Rag1* is only effective against biotype 1 (Kim et al. 2008, Hill et al. 2010). The soybean aphid in Michigan might be another biotype, since it overcame both *Rag1* and *Rag2* (unpublished data). Using microsatellite markers, Michel et al. (2009) found that the population genetic structure of soybean aphids in Michigan differed from those collected from the other eight states. Therefore, finding new sources of resistance and new resistance genes is necessary to control the newly discovered or evolved biotypes. PI 567598B and its derived lines have strong antibiosis resistance (Mensah et al. 2005) and resistance to soybean aphids as shown by biotype studies conducted in several states (Cooper 2012; Mian et al. 2008a). However, little is known about the genomic locations of the two recessive genes for the aphid resistance in PI 567598B; this could hinder its utilization. Quantitative trait loci (QTL) analysis is a powerful tool to explore the genetic mechanisms, since it not only identifies the loci, but also determines their effects. The objective of this study was to map the aphid resistance loci in PI 567598B with molecular markers.

## Materials and methods

### *Plant materials*

A population with 282 $F_{4:5}$ recombinant inbred lines (RILs) was developed from a cross between E06902 and IA2070 through the single seed descent method and used for the mapping study. E06902 is an elite advanced breeding line derived from Titan x PI 567598B and possesses aphid resistance similar to that of PI 567598B in field evaluations (unpublished data). Titan (Diers et al. 1999) is susceptible to soybean aphids. IA2070 is an experimental line from Iowa State University and is susceptible to soybean aphids.

### *Aphid resistance evaluation*

The $F_{4:5}$ RILs, parent, and the grandparent PI 567598B were evaluated for aphid damage without replication in the field in the summer of 2009. Evaluation was carried out in a 12.2 x 18.3 m aphid- and predator-proof cage (Redwood Empire Awning Co., Santa Rosa, CA, USA) on the Agronomy Farm at Michigan State University (MSU). Each line was planted in a single-row plot, 60 cm long with a row spacing of 60 cm. The average number of plants per plot was more than 10 with most lines having 12 plants.

Greenhouse evaluations were conducted for the mapping population ($F_{4:6}$ lines) in the fall of 2010 without replication, while the parental lines and PI 567598B were replicated three times. Eight seeds per line were planted in a large plastic pot 105 mm in diameter and 125 mm deep. The greenhouse was maintained at 26/15 °C day/night temperature and sodium vapor lights were used to supplement light intensity during the day (14 h).

In both field and greenhouse trials, each plant was inoculated at the V2 stage with two wingless soybean aphids. All aphid resistance evaluation trials were choice tests, which identified resistance genotypes with either antibiosis or antixenosis resistance. The aphids used for infestation in the field trial were collected from a naturally infested field on the MSU Agronomy Farm during the summer of 2009. The aphids used in the greenhouse infestation in the fall of 2010 were from greenhouse-maintained aphids originally collected from a naturally infested field on the MSU Agronomy Farm in the summer of 2010.

Aphid resistance was visually rated for each plant 3 weeks after infestation in the summer 2009 test and 3 and 4 weeks after infestation in the fall 2010 test, using a scale of 0–4 developed by Mensah et al. (2005, 2008). The following criteria were used: 0 = no aphids; 0.5 = less than 10 aphids per plant, no colony formed; 1 = 11–100 aphids per plant, plant appears healthy; 1.5 =

101–150 aphids per plant, plant appears healthy; 2 = 151–300 aphids per plant, mostly on the young leaves or tender stems, plant appears healthy; 2.5 = 301–500 aphids per plant, plant appears healthy; 3 = 501–800 aphids per plant, young leaves and tender stems covered with aphids, leaves slightly curly and shiny; 3.5 = More than 800 aphids per plant, plants stunted, leaves curled and slightly yellow, no sooty mold and few cast skins; 4 = more than 800 aphids per plant, plant stunted, leaves severely curled and yellow, covered with sooty mold and cast skins.

A damage index (DI) for each line was calculated by the following formula (Mensah et al. 2005): DI = $\Sigma$ (Scale value x No. of plants in the category)/(4 x Total no. of plants) x 100. The DI ranges between 0 for no infestation and 100 for the most severe damage. The DI was used as an indicator of aphid resistance and was applied in the following analyses.

*DNA extraction and marker analysis*

Before infestation, the non-expanded trifoliate leaves from each line were bulk harvested for genomic DNA isolation. The DNA was extracted with the CTAB (hexadecyltrimethylammonium bromide) method as described by Kisha et al. (1997), and the concentration was determined with a ND-1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA). Simple sequence repeat (SSR) markers (Song et al. 2004) were used to amplify the genomic DNA according to the PCR protocol described by Cregan and Quigley (1997), using a MJ Tetrad $^{TM}$ thermal cycler (MJ Research, Waltham, MA). PCR products were detected on 6 % non-denaturing polyacrylamide gels using a DASG-400-50 electrophoresis system (C.B.S. Scientific Co., Del Mar, CA, USA) as described by Wang et al. (2003). Gels stained with ethidium bromide were photographed, and scored under UV light.

To accelerate the location of the loci associated with the aphid resistance, the bulked segregant analysis method described by Michelmore et al. (1991) was used. Based on the 2010 phenotypic data, 10 resistant lines with the lowest DI values and 10 susceptible lines with the highest DI values were selected to form a resistant pool and a susceptible pool, respectively. Parental polymorphic SSR markers at approximately every 15 cM of the integrated soybean map of Song et al. (2004) were selected to test the polymorphism between the two bulked DNA pools. The polymorphic markers between the two pools were chosen to genotype the individual lines in the two pools together with the two parents. The markers that appeared to be associated with the aphid resistance were genotyped on the remaining lines of the whole mapping population. The genomic regions associated with the aphid resistance were then saturated with additional markers. Additional SSR markers within the candidate region were selected from 33,065 BARCSOYSSR_1.0 database (Song et al. 2010) and were screened. Primers and hybridization probes for single nucleotide polymorphic (SNP) markers were developed for TaqMan® endpoint genotyping assay (Life Technologies, Carlsbad, CA, USA) performed using Lightcycler® 480 (Roche Applied Science, Indianapolis, IN, USA). The SNP markers were selected from the SoySNP50K genotyping array for Illumina Infinium II assay (Song et al. 2013). Flanking sequences were extracted based on the genomic physical position from the soybean whole genome sequence assembly, Glyma v1.0 (www. phytozome.net/soybean).

*Statistical and QTL analysis*

Pearson correlation for the aphid resistance between trials was calculated using R statistical software (R Development Core Team 2008). A linkage map was constructed with the Kosambi function and a LOD score of 3 using JoinMap 4.0 (Van Ooijen 2006). Then, linkage groups were

assigned to specific chromosomes according to the soybean consensus map (Song et al. 2004). The maps and QTL intervals were drawn using MapChart (Voorrips 2002). Composite interval mapping (CIM) was performed to locate the aphid resistance QTLs using QTL Cartographer V2.5 with the standard model Zmapqtl 6 (Wang et al. 2008). The CIM analysis uses markers other than the interval being tested as cofactors to control the genetic background (Zeng 1994). The forward and backward regression method was used to select markers as cofactors. The walking speed chosen for CIM was 1 cM. The empirical LOD threshold at 5 % probability level was determined by a 1,000-permutation test (Churchill and Doerge 1994). The QTL x QTL interaction was further determined using the multiple interval mapping (MIM) method of QTL Cartographer.

**Results**

*Phenotypic analysis*

The phenotypic values of the 282 $F_4$-derived RILs and its parents, and the resistant source PI 567598B are summarized in Table 2.1. In both field and greenhouse trials, the susceptible parent, IA2070, was severely damaged by the aphids, while the resistant parent E06902 and PI 567598B were not. There was no significant difference in aphid resistance between E06902 and PI 567598B. Correlation between the 3- and 4-week ratings from the 2010 greenhouse trial was strong ($r = 0.88$, $P < 0.0001$). However, ratings from the 2010 greenhouse trial were not strongly correlated with the 2009 field ratings (0.37 and 0.44 for the week 3 and 4 ratings, respectively, $P < 0.0001$). The distributions for the population ratings in both field and greenhouse trials were continuous, but not normal ($W = 0.80$, 0.92 and 0.92, respectively at $P < 0.0001$), and the distribution in the field trial appeared bimodal (Fig. 2.1a, b, c). This indicates

that a limited number of major genes might control the aphid resistance in PI 567598B.

**Table 2.1** Mean damage index of the $F_4$-derived main mapping population and its parental lines and grandparent, PI 567598B, in the field trial in summer 2009 and greenhouse trial in fall 2010

| Trial | Parents[a] | | Grandparent[a] | $F_4$-derived lines | | |
|---|---|---|---|---|---|---|
| | IA2070 | E06902 | PI 567598B | Mean | Range | SE |
| Field 2009 | | | | | | |
| 3-week rating | 87.5b | 16.8a | 12.5a | 57.0 | 12.5~87.5 | 24.0 |
| Greenhouse 2010 | | | | | | |
| 3-week rating | 85.5b | 19.0a | 16.3a | 46.5 | 12.5~87.5 | 23.3 |
| 4-week rating | 87.5b | 15.5a | 23.3a | 46.8 | 12.5~97.5 | 30.5 |

[a] Within trials and ratings, means followed by the same letters are not significantly different according to Fisher's Protected LSD ($P = 0.05$)

**Figure 2.1** Frequency distribution of soybean aphid damage index (%) taken from $F_4$-derived lines of the cross IA2070 x E06902. Parental lines, PI 567598B, and Dowling ratings are shown by arrows. **a** Three-week rating in the field trial in summer 2009, **b** Three-week rating in the greenhouse trial in fall 2010, **c** Four-week rating in the greenhouse trial in fall 2010

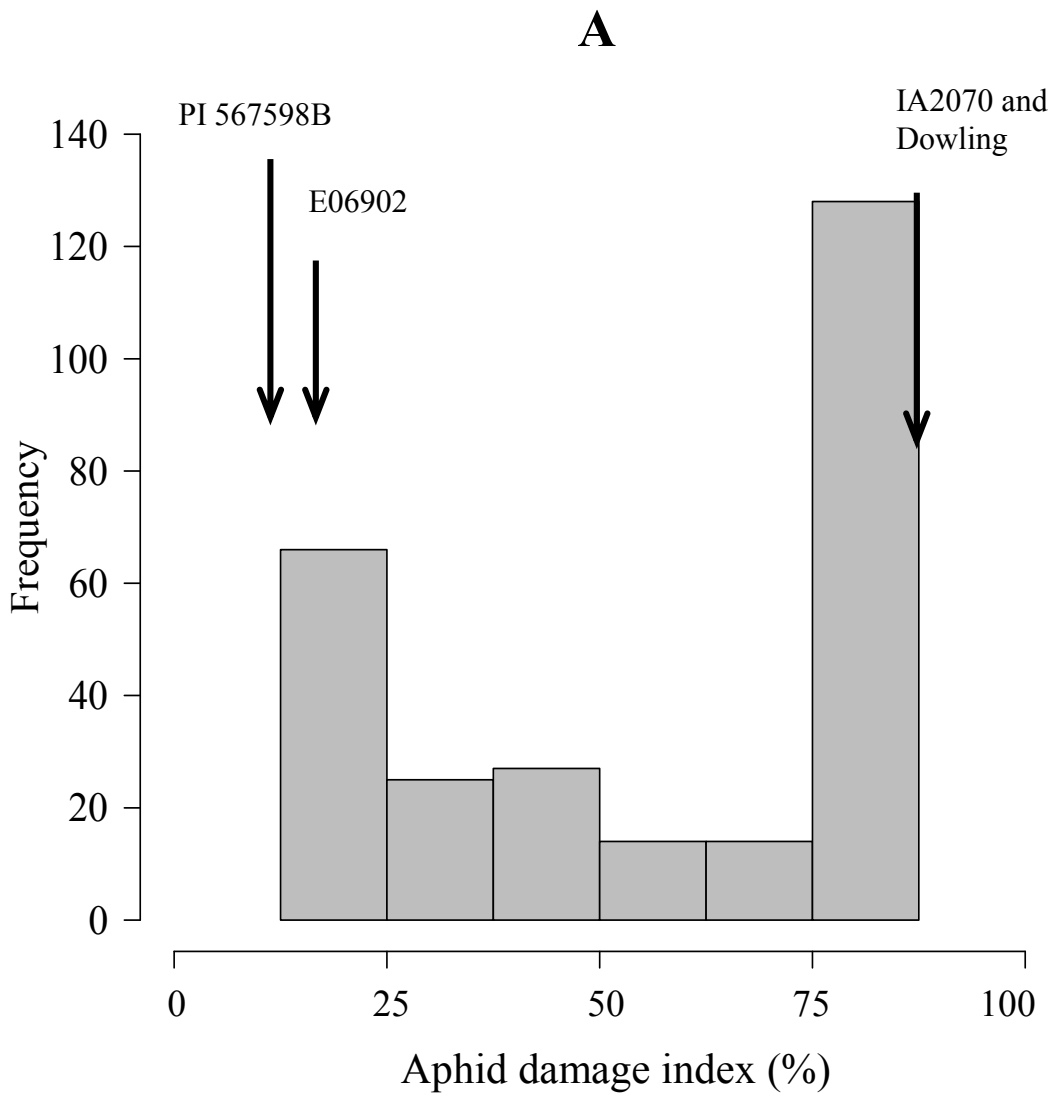


**Figure 2.1 (cont'd)**

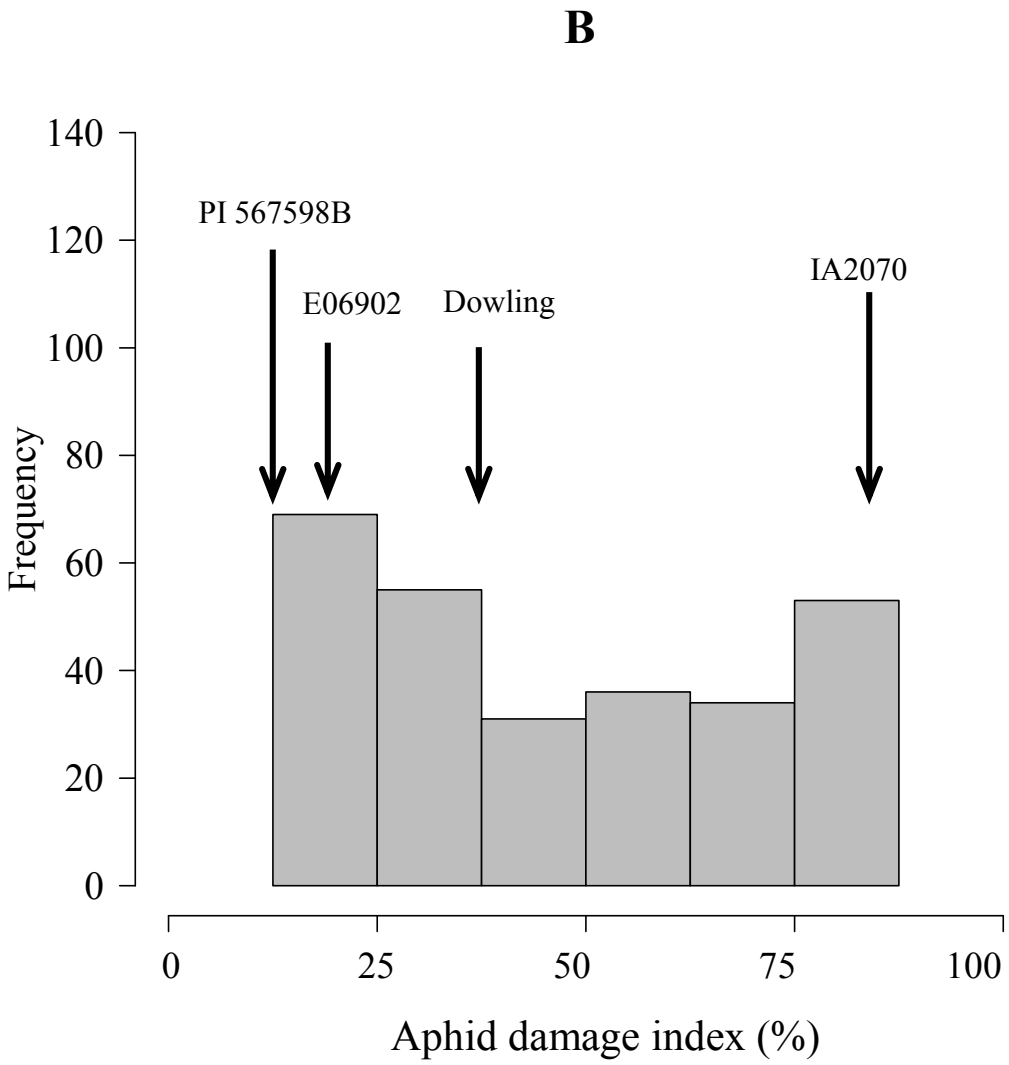
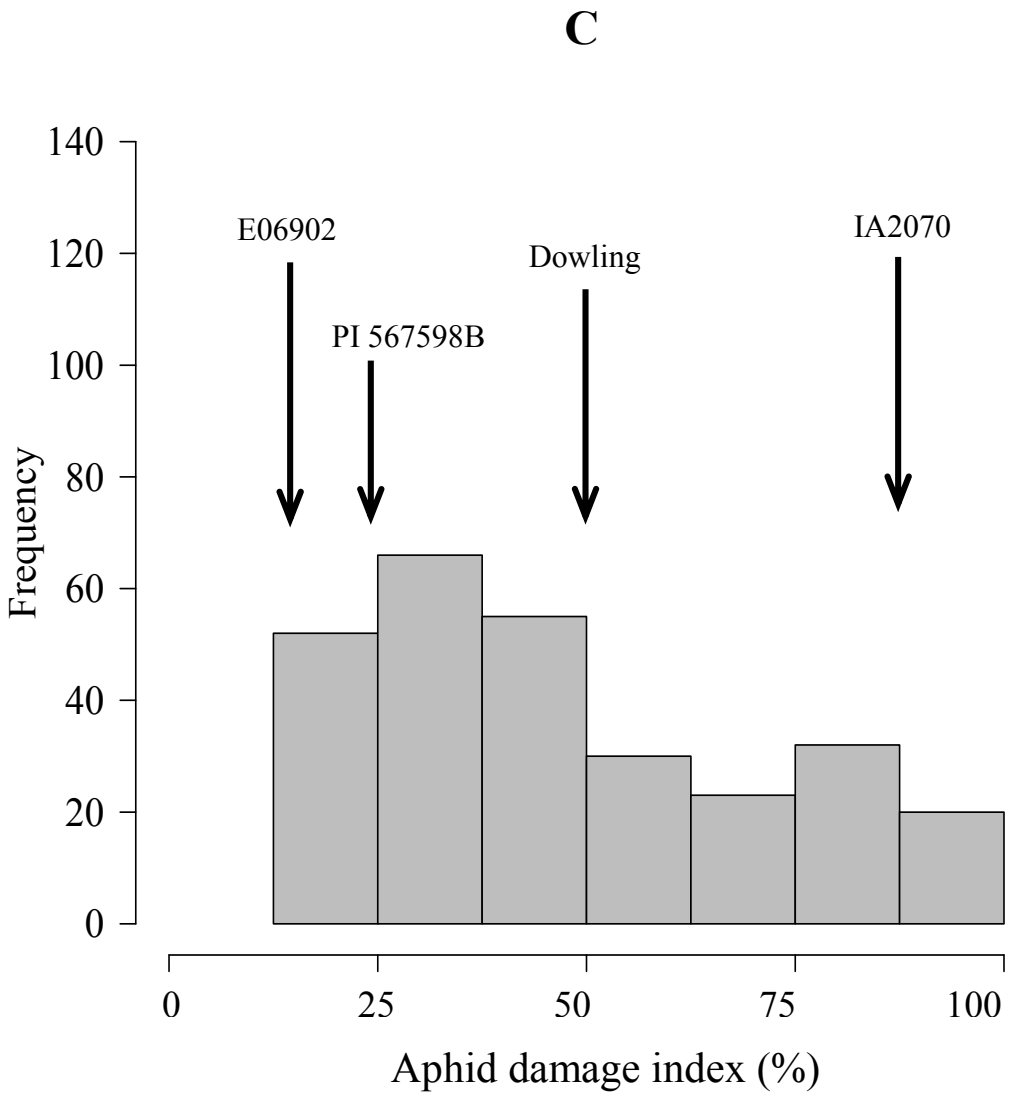


**Figure 2.1 (cont'd)**



*QTL analysis*

Among 1056 SSR markers, 38 revealed polymorphism between the resistant and the susceptible bulk DNA samples. These 38 markers were from chromosomes 1, 3, 7, 13, 16 and 18 (LGs D1a, N, M, F, J and G). Only Satt654 and Sct_001 on chromosome 16 (LG J) and Satt435 on chromosome 7 (LG M) appeared to be associated with aphid resistance when the individual lines from the DNA pools were genotyped. Therefore, these two regions were saturated with parental polymorphic markers within ±20 cM in the consensus map (Song et al. 2004) using the whole population. Based on the markers from BARCSOYSSR 1.0 database (Song et al. 2010), 48 additional markers were screened for polymorphism within the identified intervals. BARCSOYSSR16_0366 on chromosome 16 was found to be associated with aphid resistance, while four other markers between Satt435 and Satt323 in the chromosome 7 interval were found to be polymorphic. SNP markers in these two chromosome intervals were also extracted from the SNP list in SoySNP50K genotyping array (Song et al. 2013) and designed for TaqMan® endpoint genotyping assay.

A total of eight SSR and four SNP markers were mapped to the interval on chromosome 16, spanning a total of 43.5 cM (Fig. 2.2a); while seven SSR and one SNP marker were mapped to the interval on chromosome 7, spanning a total of 45.9 cM (Fig. 2.2d).

The QTL analysis detected two QTLs based on the greenhouse trial, while only the one on chromosome 16 was significant in the field trial. In both trials, the allele from E06902 conferred resistance against soybean aphids at the identified QTLs. Using the CIM method, the QTL on chromosome 16 was consistently mapped between Gm16_6262227_C_T and Gm16_6424067_A_G and explained 30.7–45.8 % of the phenotypic variation, with the field trial

having the highest percentage (Table 2.2; Fig. 2.2a). The QTL on chromosome 7 was only detected in the greenhouse trials and located between Satt435 and BARCSOYSSR_07_0309, explaining over 30 % of the phenotypic variation (Table 2.2; Fig. 2.2d).

The MIM (multiple interval mapping) method was further conducted to determine whether there was significant QTL interaction. The MIM method detected the same QTLs as CIM (composite interval mapping) method with two QTLs in the greenhouse trial and one QTL in the field trial (Table 2.3). For the week 4 ratings in the greenhouse trial, MIM method detected a significant additive x additive interaction between the two QTLs located on chromosome 7 and 16, but this was not the case for the week 3 ratings. The LOD score of the QTL interaction is 3.4 and explained 1.2 % of the total phenotypic variations. The two QTLs together with their interaction explained 41.7 % of the total phenotypic variation. For the week 3 ratings, these two QTLs together explained 33.6 % of the phenotypic variation. The QTL on chromosome 16 detected in the field trial explained the highest phenotypic variation, 56.1 %.

Since the QTLs from this study were mapped to similar regions as *Rag1* (Li et al. 2007) and *Rag3* (Zhang et al. 2009), we named the locus on chromosome 7 as *rag1b* and the locus on chromosome 16 as *rag3*, according to the conventions of the Soybean Genetics Committee.

**Table 2.2** Summary of QTLs for soybean aphid resistance detected in the main mapping population (IA2070 x E06902) and alternative population (PI 567598B x Titan) using the composite interval mapping method

| Trials | Chr/ LG[a] | Peak pos.[b] | Flanking markers [c] | QTL [d] | | |
|---|---|---|---|---|---|---|
| | | | | LOD | $R^2$ [e] | a [f] |
| **IA2070 x E06902 population** | | | | | | |
| Field 2009 | | | | | | |
| 3-week rating | 16/J | 7.5 | Gm16_6424067_A_G[g] | 42.5 | 45.8 | -22.1 |
| Greenhouse 2010 | | | | | | |
| 3-week rating | 7/M | 3.6 | Satt435-BARCSOYSSR_07_0295 | 16.6 | 35.5 | -11.2 |
| | 16/J | 5.5 | Gm16_6262227_C_T - Gm16_6423098_G_A | 12.5 | 35.9 | -9.2 |
| 4-week rating | 7/M | 5.3 | BARCSOYSSR_07_0295 - BARCSOYSSR_07_0309 | 16.7 | 31.2 | -10.3 |
| | 16/J | 7.3 | Gm16_6423098_G_A - Gm16_6424067_A_G | 15.9 | 30.7 | -9.9 |

[a] Chromosome/Linkage group. The chromosome number and linkage group name are according to the SoyBase (Grant et al. 2010)

[b] QTL peak position is expressed in cM

[c] Markers flanking the peak position

[d] The LOD thresholds are 3.89, 1.78, 2.3, 2.4, and 5.1 for the field 2009 rating, the three week rating in 2010, the four-week rating in 2009, the three week rating in 2008 and the 2009 rating, respectively.

[e] $R^2$, percentage of phenotypic variation explained by a QTL

[f] Additive effect. The negative value implies that the IA2070 allele increases the phenotypic value

[g] Marker on the peak position

**Table 2.3** Summary of QTLs for soybean aphid resistance detected in the main mapping population (IA2070 x E06902) and alternative population (PI 567598B x Titan) using the multiple interval mapping method

| Trials | Chr/ LG[a] | Peak pos.[b] | Flanking markers[c] | Genetic effect | | |
|---|---|---|---|---|---|---|
| | | | | LOD[d] | $R^{2e}$ | a[f] |
| **IA2070 x E06902 population** | | | | | | |
| Field 2009 | | | | | | |
| 3-week rating | 16/J | 7.5 | Gm16_6424067_A_G[g] | 56.0 | 56.1 | -26.8 |
| Greenhouse 2010 | | | | | | |
| 3-week rating | 07/M | 5.3 | BARCSOYSSR_07_0295 - BARCSOYSSR_07_0309 | 17.3 | 20.4 | -12.9 |
| | 16/J | 7.5 | Gm16_6424067_A_G | 12.3 | 13.2 | -10.5 |
| | | | Total | | 33.6 | |
| 4-week rating | 07/M | 6.3 | BARCSOYSSR_07_0295 - BARCSOYSSR_07_0309 | 19.9 | 21.2 | -13.0 |
| | 16/J | 7.5 | Gm16_6424067_A_G | 20.0 | 19.3 | -12.6 |
| | | | Interaction | 3.4 | 1.2 | -4.7 |
| | | | Total | | 41.7 | |

[a] Chromosome/Linkage group. The chromosome number and linkage group name are according to the SoyBase (Grant et al. 2010)

[b] QTL peak position is expressed in cM

[c] Markers flanking the peak position or the marker at the peak position

[d] Using the same LOD thresholds as in the composite interval mapping method (Table 2.3)

[e] $R^2$, percentage of phenotypic variation explained by a QTL

[f] Additive effect. The negative value implies that the IA2070 allele increases the phenotypic value

[g] Marker on the peak position

**Figure 2.2** Locations of soybean aphid resistance QTLs using composite interval mapping method. Solid bars represent QTLs for the three-week rating in the 2009 field trial (*2009Field-Wk3*). Diagonally-hatched bars represent QTLs for the three-week rating in the 2010 trial (*2010GH-Wk3*). Open bars represent QTLs for the four-week rating in the 2010 trial (*2010GH-Wk4*). **a** and **c** Maps of chromosome 16 (LG J) and 7 (LG M) in the mapping population, the QTL positions are listed at its left side; **b** and **d** Consensus maps of chromosome 16 (LG J) and 7 (LG M) (Song et al. 2004)

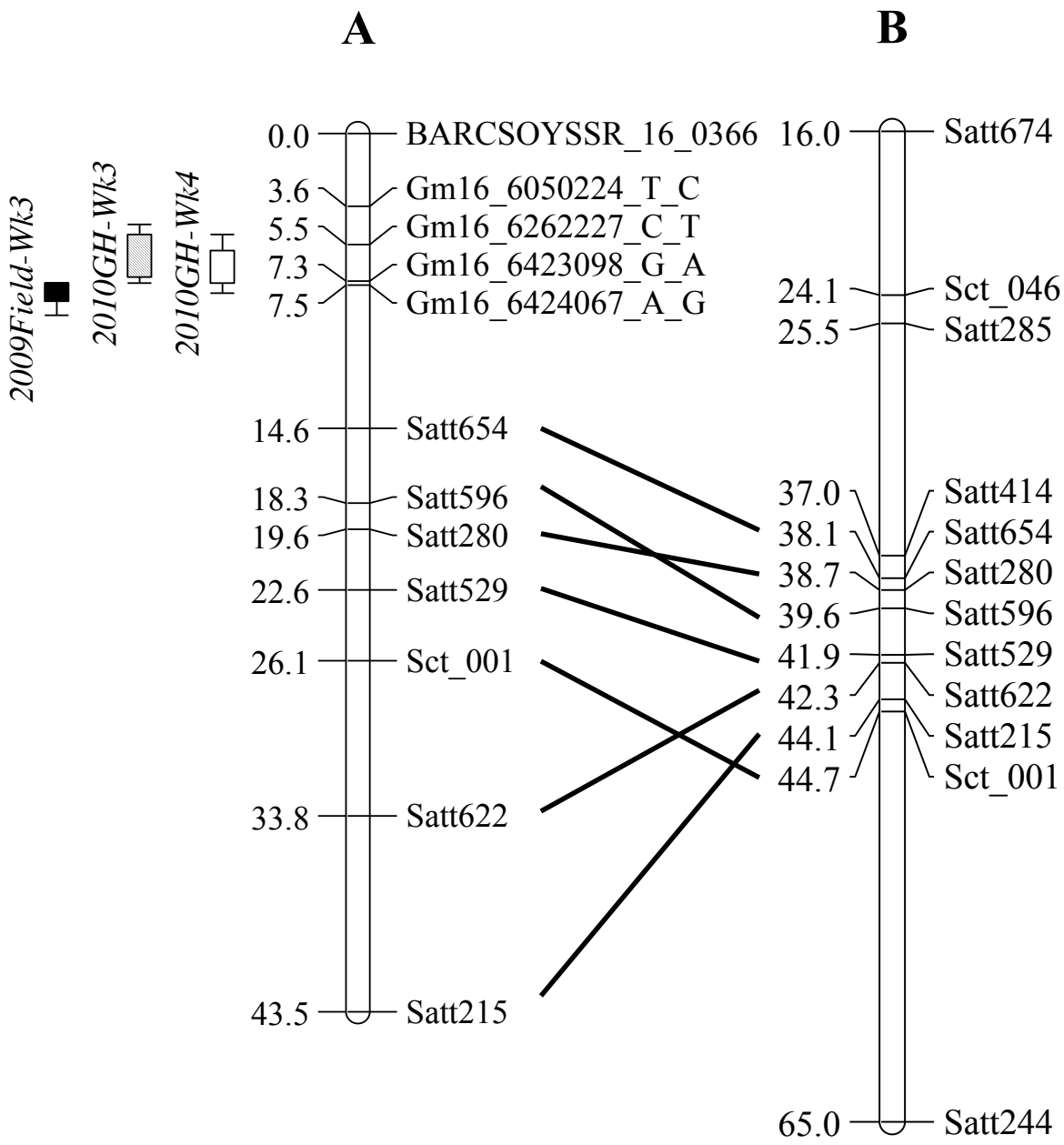
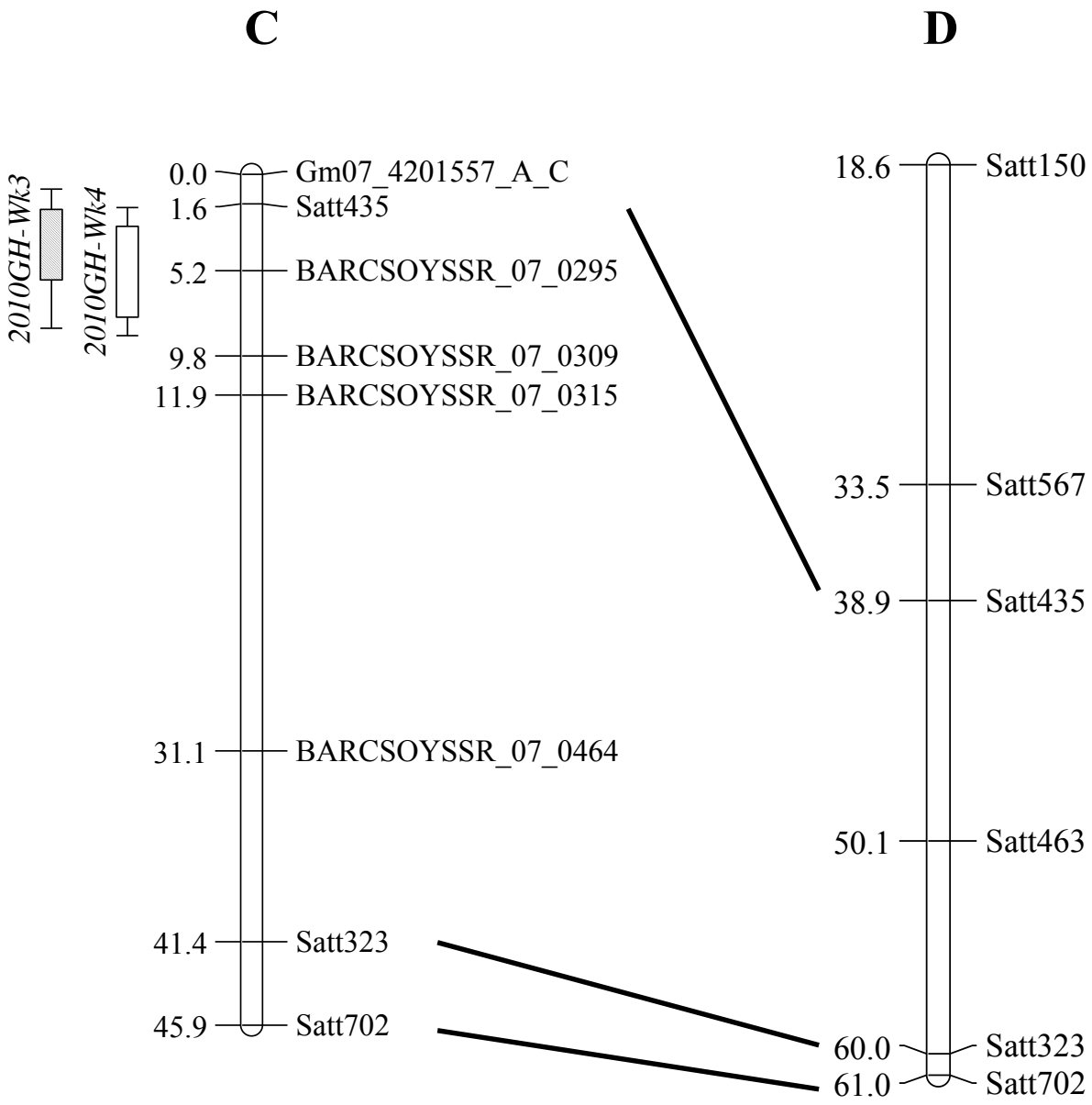


**Fig. 2.2 (cont'd)**


C
D
2010GH-Wk3
2010GH-Wk4
0.0
Gm07_4201557_A_C
1.6
Satt435
5.2
BARCSOYSSR_07_0295
9.8
BARCSOYSSR_07_0309
11.9
BARCSOYSSR_07_0315
31.1
BARCSOYSSR_07_0464
41.4
Satt323
45.9
Satt702
18.6
Satt150
33.5
Satt567
38.9
Satt435
50.1
Satt463
60.0
Satt323
61.0
Satt702

*Effect of the combination of QTL alternative alleles*

The F4-derived lines from the mapping population were classified based on the SNP and SSR alleles within the QTL regions identified in E06902. Four distinct genotypes were defined by the presence or absence of the allele from E06902 for those QTL-associated markers on chromosomes 7 and 16 (Table 2.4). A total of 139 lines were grouped into the defined genotypes and only individual lines with complete and unambiguous genotype data for all loci were included. Mean soybean aphid damage index for all lines within each genotypic group was obtained for each of the trials in 2009 and 2010. In the 2010 greenhouse trial, the presence of E06902 alleles at both *rag1b* and *rag3* gave the lowest aphid damage, while absence of alleles at both QTLs made lines very susceptible (Fig. 2.3a). The absence of E06902 allele at one QTL (either *rag1b* or *rag3*) gave intermediate resistance against aphids. However, in the 2009 field cage trial, the lines without *rag3* were as susceptible as those with none of the two QTL alleles from E06902 (Fig. 2.3b). On the other hand, genotypes containing only *rag3* gave resistant phenotypes that were comparable to the lines that had both resistant alleles. It seems that the QTL on chromosome 7 (*rag1b*) failed to confer resistance in the field trial, while only QTL on chromosome 16 (*rag3*) conferred resistance. This shows that the QTLs identified in this study confer differential reactions against the soybean aphids in the field and greenhouse trials.

**Table 2.4** Genotypic groups of 139 $F_4$-derived lines from the mapping population IA2070 x E06902 containing alternative alleles of the associated markers on chromosome 7 (*rag1b*) and 16 (*rag3*)

| Genotype[a] | No. of lines | SSR and SNP Markers | | | | | |
|---|---|---|---|---|---|---|---|
| | | Chromosome 7 | | | Chromosome 16 | | |
| | | Satt435 | BARCSOYSSR_07_0295 | BARCSOYSSR_07_0309 | Gm16_6262227_C_T | Gm16_6423098_G_A | Gm16_6424067_A_G |
| *rag1b* / *rag3* | 43 | + | + | + | + | + | + |
| *rag1b* / - | 35 | + | + | + | - | - | - |
| - / *rag3* | 40 | - | - | - | + | + | + |
| - / - | 21 | - | - | - | - | - | - |

[a] (+) Implies homozygous allele from the E06902 resistant source. (-) Implies homozygous allele from susceptible parent.

**Figure 2.3** Mean soybean aphid damage index (%) for selected lines having alternate alleles within the intervals of Satt435 and BARCSOYSSR_07_0309 (chromosome 7) and with Gm16_6262227_C_T and Gm16_6424067_A_G (chromosome 16) in the mapping population IA2070 x E06902. **a** Three-week and four-week rating in the greenhouse trial in fall 2010, **b** Three-week rating in the field trial in summer 2009. Lines shown are standard error. Bars with the same letter are not significantly different according to Fisher's Protected LSD ($P = 0.05$)

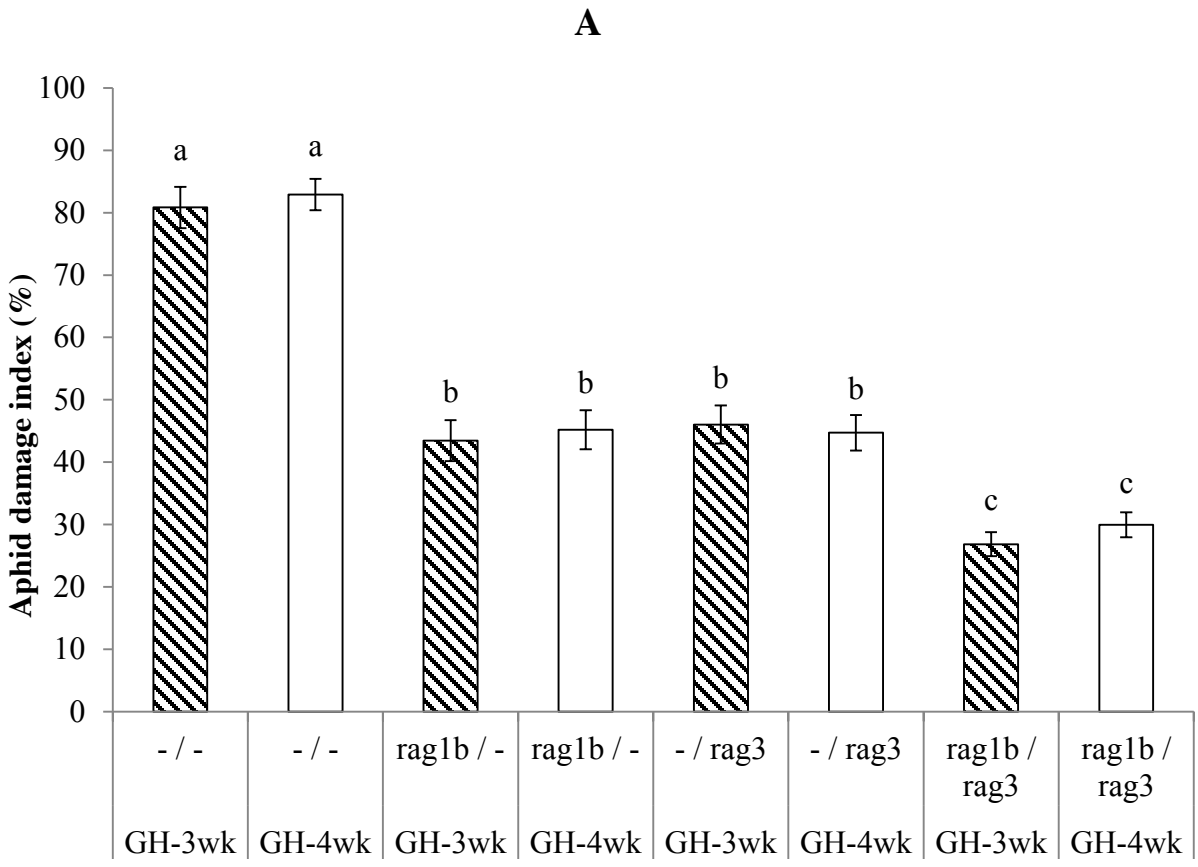
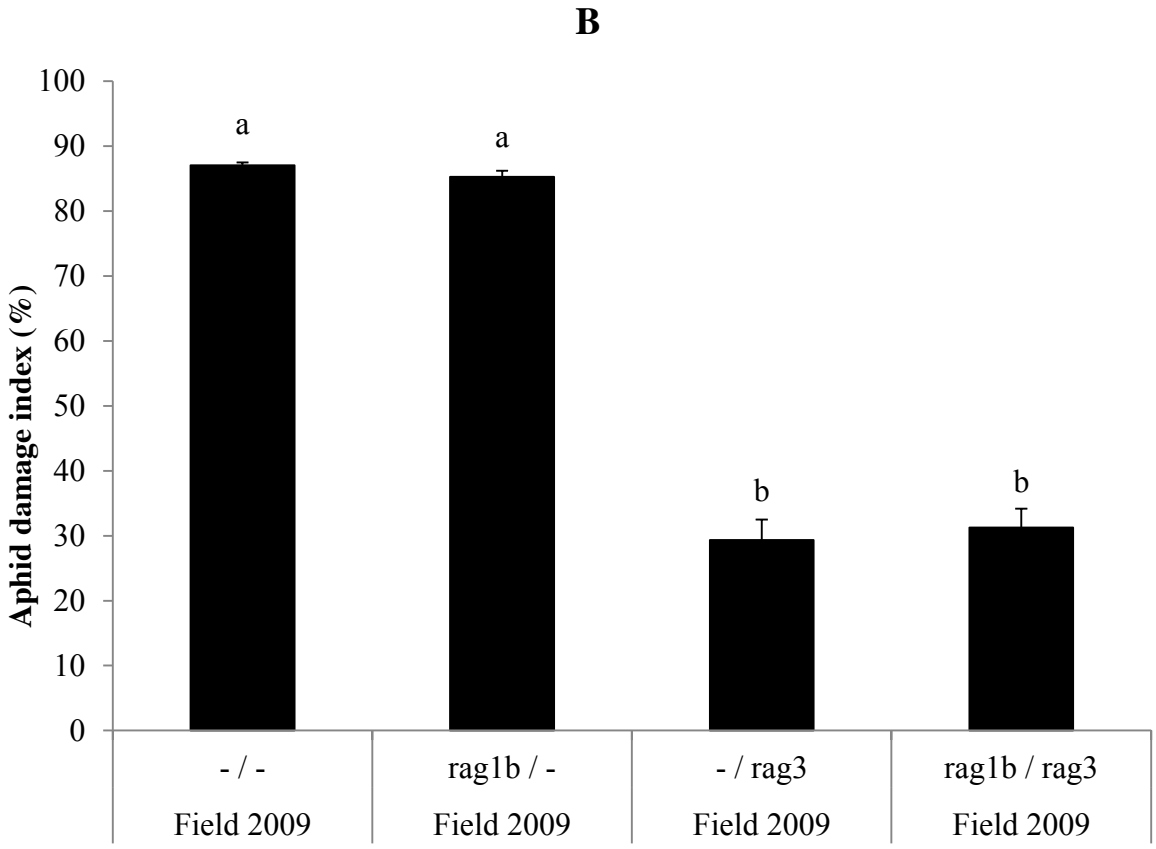


**Fig. 2.3 cont'd**




**Discussion**

In this study, two QTLs for controlling the aphid resistance in PI 567598B or its derived line were consistently detected in all three years. These two QTLs explained most of the phenotypic variation, indicating that two major genes control the aphid resistance in PI 567598B. This finding is consistent with the conclusion of Mensah et al. (2008), who conducted a genetic study and suggested a two-gene model for the aphid resistance in PI 567598B. Other than soybeans, single recessive genes controlling aphid resistance have been previously reported for wheat (Nkongolo et al. 1991), peanut (Herselman et al. 2004), and corn (So et al. 2010). The difference or similarity of action between dominant and recessive aphid resistance genes in

soybeans remains to be investigated.

The QTLs detected in this study are located in similar genomic regions as *Rag1* on chromosome 7 (Li et al. 2007, Kim et al. 2010) and *Rag3* on chromosome 16 (Zhang et al. 2010). Although the two resistant genes in PI 567598B were considered recessive (Mensah et al. 2008), and *Rag1* or *Rag3* was considered dominant or co-dominant (Hill et al. 2006a; Zhang et al. 2010), they may still be the same genes as *Rag1* and *Rag3*, since the susceptible parent in this study was different from the ones used for characterizing *Rag1* and *Rag3*. It is also possible that the two genes discovered in this study are allelic to *Rag1* or *Rag3*, or different genes, but tightly linked to *Rag1* or *Rag3*. *Rag1* in Dowling can be overcome by the Michigan aphids and *Rag3* in PI 567543C did not provide antibiosis resistance (Mensah et al. 2005). PI 567598B had a relatively lower DI value than PI 567543C (unpublished data). The better resistance of PI 567598B compared with Dowling and PI 567543C might be due to one or more of the following factors: (1) the stacking of resistant genes *rag1b* and *rag3*; (2) different resistant alleles at *rag1b* or *rag3*, or both loci conferring better resistance than the alleles in Dowling and PI 567543C; (3) one or two new genes closely linked to *Rag1* and/or *Rag3* have better resistance than *Rag1* and *Rag3*. Further investigations, such as fine mapping or gene cloning, might be necessary to elucidate their relationships with *Rag1* and *Rag3*.

Different sources of aphids could determine the resistant reaction of a soybean plant containing any of the *Rag* genes. Our QTL analysis revealed that only *rag3* was detected in the field trial. Zhang et al. (2009) also found that the two resistance genes from PI 567541B were expressed differently in the field and greenhouse trials, which was explained by the different aphid biotypes. In fact, Mensah et al. (2007) found that the *Rag1* in Dowling was first overcome by the Michigan aphids in 2006. The mixture of aphids used in the 2009 field trial in this study

was collected from the Michigan field, which infested Dowling (unpublished data) in that year. This may explain why the QTL on chromosome 7 was not detected in our field trial. However, in the 2010 greenhouse trial, the single aphid clone collected from the field in 2010 was used, and Dowling was resistant to this clone (unpublished data). Consequently, the QTL on chromosome 7 was significant in this trial. The present study demonstrated that PI 567598B can still give some effective tolerance to aphids even if one of the resistance genes is overcome, indicating that the presence of both *rag1b* and *rag3* can confer broader aphid resistance. This supports the hypothesis that stacking more than one aphid resistance gene will provide durable resistance against soybean aphids.

PI 567598B possesses strong and broad resistance to soybean aphids; therefore, it is a promising resistant source for improving aphid resistance in soybean. The localization of the two resistance genes in PI 567598B using molecular markers in this study could be useful to breeders in marker-assisted selection for aphid resistance lines.

# REFERENCES

## REFERENCES


Beckendorf EA, Catangui MA, Riedell WE (2008) Soybean aphid feeding injury and soybean yield, yield components, and seed composition. Agron J 100:237–246

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Cooper S (2012) Soybean aphid biotypes: understanding geographic distribution (2008–2012), Monsanto Company. In: Conference presentation for Soybean Breeders' Workshop held February 27–29, 2012, St. Louis, MO, USA

Cregan PB, Quigley CV (1997) Simple sequence repeat DNA marker analysis. In: Caetano-Anolles G, Gresshoff PM (eds) DNA markers: protocols, applications and overviews. Wiley, New York, pp 173–185

Diaz-Montano J, Reese JC, Schapaugh WT, Campbell LR (2006) Characterization of antibiosis and antixenosis to the soybean aphid in several soybean genotypes. J Econ Entomol 999:1884–1889

Diers BW, Isleib TG, Sneller CH, Boyse JF (1999) Titan soybean. Crop Sci 39(5):1534

Grant D, Imsande MI, Shoemaker RC (2010) SoyBase, The USDA- ARS soybean genome database. http://soybase.agron.iastate.edu

Herselman L, Thwaites R, Kimmins FM et al (2004) Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosettedisease. Theor Appl Genet 109:1426–1433

Hesler LS, Dashiell KE (2008) Identification and characterization of new sources of resistance to *Aphis glycines* Matsumura (Hemiptera: *Aphididae*) in soybean lines. Appl Entomol Zool 43:197–206

Hesler LS, Dashiell KE, Lundgren JG (2007) Characterization of resistance to *Aphis glycines* in soybean accessions. Euphytica 154:91–94

Hill CB, Li Y, Hartman GL (2004) Resistance to the soybean aphid in soybean germplasm. Crop Sci 44:98–106

Hill CB, Li Y, Hartman GL (2006a) A single dominant gene for resistance to the soybean aphid in the soybean cultivar Dowling. Crop Sci 46:1601–1605

Hill CB, Li Y, Hartman GL (2006b) Soybean aphid resistance in soybean Jackson is controlled by a single dominant gene. Crop Sci 46:1606–1608

Hill CB, Kim KS, Crull L, Diers BW, Hartman GL (2009) Inheritance of resistance to the soybean aphid in soybean PI 200538. Crop Sci 49:1193–1200

Hill CB, Crull L, Herman TK, Voegtlin JD, Hartman GL (2010) A new soybean aphid (Hemiptera: Aphididae) biotype identified. J Econ Entomol 103:509–515

Kang ST, Mian MAR, Hammond RB (2008) Soybean aphid resistance in PI 243540 is controlled by a single dominant Gene. Crop Sci 48:1744–1748

Kim K, Hill CB, Hartman GL, Mian MR, Diers BW (2008) Discovery of soybean aphid biotypes. Crop Sci 48:923–928

Kim K, Bellendir S, Hudson KA, Hill CB, Hartman GL, Hyten DL, Hudson ME, Diers BW (2010) Fine mapping the soybean aphid resistance gene *Rag1* in soybean. Theor Appl Genet 120:1063–1071

Kisha T, Sneller CH, Diers BW (1997) Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci 37:1317–1325

Li Y, Hill CB, Carlson SR, Diers BW, Hartman GL (2007) Soybean aphid resistance genes in the soybean cultivars Dowling and Jackson map to linkage group M. Mol Breed 19:25–34

Liu XM, Smith CM, Friebe BR, Gill BS (2005) Molecular mapping and allelic relationships of Russian wheat aphid–resistance genes. Crop Sci 45:2273

Mensah C, Difonzo C, Nelson RL, Wang D (2005) Resistance to soybean aphid in early maturing soybean germplasm. Crop Sci 45:2228–2233

Mensah C, DiFonzo C, Wang D (2007) A case for the presence of soybean aphid biotypes in Michigan. In: 2007 annual meeting abstracts [CD-ROM]. ASA, CSSA, and SSSA, Madison, WI

Mensah C, Difonzo C, Wang D (2008) Inheritance of soybean aphid resistance in PI 567541B and PI 567598B. Crop Sci 48:1759–1763

Mian MAR, Hammond RB, St Martin SK (2008a) New plant introductions with resistance to the soybean aphid. Crop Sci 48:1055–1061

Mian MAR, Kang ST, Beil SE, Hammond RB (2008b) Genetic linkage mapping of the soybean aphid resistance gene in PI 243540. Theor Appl Genet 117:955–962

Michel AP, Zhang W, Jung JK, Kang ST, Mian MAR (2009) Population genetic structure of *Aphis glycines*. Environ Entomol 38:1301–1311

Michelmore RW, Paran J, Kesseli RV (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88:9828–9832

Mornhinweg DW, Porter DR, Webster JA (2002) Inheritance of russian wheat aphid resistance in spring barley germplasm line STARS-9577B. Crop Sci 42:1891

Nkongolo KK, Quick JS, Limin AE, Fowler DB (1991) Sources and inheritance of resistance to Russian wheat aphid in *Triticum* species amphiploids and *Triticum tauschii*. Can J Plant Sci 71:703–708

Painter RH (1951) Insect Resistance in Crop Plants. Macmillan Publishing Co, New York

Pascal T, Pfeiffer F, Kervella J et al (2002) Inheritance of green peach aphid resistance in the peach cultivar ‘‘Rubira’’. Plant Breeding 121:459–461

Pathak RS (1988) Genetics of resistance to aphid in cowpea. Crop Sci 28:474–476

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

So Y-S, Ji HC, Brewbaker JL (2010) Resistance to corn leaf aphid (*Rhopalosiphum maidis* Fitch) in tropical corn (Zea mays L.). Euphytica 172:373–381

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128

Song Q, Jia G, Zhu Y, Grant D, Nelson R, Hwang E, Hyten D, Cregan P (2010) Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. Crop Sci 50:1950–1960

Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB (2013) Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS ONE 8(1):e54985

Van Ooijen JW (2006) JoinMap$^{TM}$ 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, Netherlands

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Heredity 93:77–78

Wang D, Shi J, Carlson SR, Cregan PB, Ward RW, Diers BW (2003) A low-cost and high-throughput system for high-resolution genotyping with microsatellite DNA markers. Crop Sci 43:1828–1832

Wang S, Basten CJ, Zeng ZB (2008) Windows QTL Cartographer 2.5. Department of Statistic, North Carolina State Univ, Raleigh. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm

Wu Z, Schenk-Hamlin D, Zhan W, Ragsdale DW, Heimpel GE (2004) The soybean aphid in China: a historical review. Ann Entomol Soc Am 97:209–218

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468

Zhang G, Gu C, Wang D (2009) Molecular mapping of soybean aphid resistance in PI 567541B.

Theor Appl Genet 118:473–482

Zhang G, Gu C, Wang D (2010) A novel locus for soybean aphid resistance. Theor Appl Genet 120:1183–1191

# CHAPTER III

## FINE MAPPING OF APHID RESISTANCE GENES *RAG3* AND *RAG1B* IN SOYBEAN PI 567598B


### Abstract

Soybean production in the Midwestern U.S. is increasingly more challenging by the soybean aphid (*Aphis glycines* Matsumura) after its discovery in 2000. Four soybean aphid biotypes have been reported in the field since then. One of the best approach to effectively control evolving biotypes of soybean aphids is to pyramid resistant genes from different sources using marker-assisted breeding strategies. Plant introduction (PI) 567598B has strong antibiosis resistance against the soybean aphid. Previous studies revealed that two QTLs on chromosome 7 (*rag1b*) and chromosome 16 (*rag3*) control the aphid resistance in PI 567598B. The objective of this study was to fine map the aphid resistance QTLs on chromosome 7 and 16 by selecting plants with recombination events within the QTL interval. Fine mapping was conducted by screening 4,041 $BC_1F_2$-derived lines using SNP markers that flank the interval of *rag1b* and *rag3*. Fifty-five recombinants were tested with high-density molecular markers using the SoySNP50K genotyping array. Progenies of each recombinant were rated for damage by the soybean aphid and tested using custom-designed SNP assays. Results of fine mapping delimited *rag3* into a 152-kb interval between SNP markers on chromosome 16. An attempt to fine-map *rag1b* from three populations of different genetic background has proven that *rag1b* is dependent with the presence of alleles for *rag3*. From the soybean annotation database, there are eleven candidate genes found within the *rag3* interval where eight are encoded as NBS-LRR genes.

## Introduction

Soybean is the world's foremost provider of plant protein and oil for human and animal consumption. In 2012, 1.345 billion bushels (36.6 million metric tons) of soybeans were exported by the United States, accounting for 37% of the world's soybean trade. Domestically, soybeans provided 66% of the edible consumption of fats and oils in the United States (www.soystats.com, 2012). The soybean aphid (SBA), *Aphis glycines* Matsumura, a pest in Asia became a major threat to soybean production in the U.S. when it was discovered in the Midwest in 2000 (Hartman et al., 2001). The impact of SBA on yield is mainly caused by direct plant feeding and secondary effects are brought about by black sooty mold infestation and virus transmission which can result in plant stunting, leaf deformation and thus reduced pod set (Hill et al., 2001; Li et al., 2008).

Three QTLs for SBA resistance have been fine-mapped so far. Kim et al. (2009) fine-mapped *Rag1* from Dowling to a 115-kb interval from 824 $BC_4F_2$ and 1,000 $BC_4F_3$ plants. *Rag2* from PI 200538, on the other hand, was mapped to a 54-kb interval using lines derived from 5,782 $F_2$ plants (Kim et al., 2010). From the fine mapping studies, SNP markers were discovered and used for marker-assisted selection of aphid resistant lines derived from Dowling and PI 200538. SSR markers were used to delimit *[Rag6]_P203* into a 192-kbp interval (Xiao et al., 2013).

Two out of the 13 predicted genes in the 115-kb interval of *Rag1* (Kim et al., 2009) were identified as potential candidate genes, while one out of seven predicted genes were deduced from the 54-kb interval of *Rag2* (Kim et al., 2010). All of the identified candidate genes of *Rag1* and *Rag2* were nucleotide-binding site leucine-rich repeat (NBS-LRR) genes, which is similar to findings on *Mi* gene that deters potato aphid feeding on tomato (Rossi et al., 1998; Kaloshian et

al., 2000; Cooper et al., 2004); also on *Vat* gene in melons against the melon aphid (Chen et al., 1997; Villada et al., 2009) and resistance against blue-green aphid in *Medicago truncatula* (Klingler et al., 2005).

Since only two SBA resistant genes were fine-mapped and subsequent cloning has to be done to be able to identify functional genes, there is a need to further fine map the SBA resistance genes from other identified sources to identify tightly linked markers and be able to pyramid aphid resistance genes for durable resistance. Thus, the present study aimed to fine map the recessive genes found in PI 567598B.

## Materials and methods

### *Plant materials and screening of recombinants*

Two backcross populations were first used to determine recombinants that delimited the position of *rag3* in PI 567598B (Table 3.1). The first population, designated as 090004, was composed of 2,214 $BC_1F_2$ lines from a cross between E00003 and PI 567598B, where E00003 was the recurrent parent. E00003 had resistance against soybean root rot caused by *Phytophthora sojae* but was susceptible to soybean aphid. The second population, named as 090068, consisted of 1,827 $BC_1F_2$ lines from a cross between Skylla and PI 567598B. Skylla was the recurrent and soybean aphid-susceptible parent. The recurrent parents were cultivars developed at Michigan State University (Wang et al., 2006).

All 4,041 $F_2$ plants from the two populations were screened for recombination breakpoints near *rag3* using TaqMan SNP assays, MSU16-04 (Gm16_6050224_T_C) and MSU16-13 SNP (Gm16_6424067_A_G), identified from a previous genetic QTL mapping study

**Table 3.1** Populations derived from crosses with soybean aphid-resistant line PI 567598B that were used for screening of recombinants to delimit the locations of *rag3* and *rag1b*.

| Population | Female Parent | Male Parent (Pedigree) | Generation | Number of lines |
|---|---|---|---|---|
| 090004 | E00003[S] | E09902[R] (E00003 x PI 567598B) | $BC_1F_2$ | 2,214 |
| 090068 | Skylla[S] | E09904[R] (Skylla x PI 567598B) | $BC_1F_2$ | 1,827 |
| E10011 | IA2064[S] | E06906[R] (Titan x PI 567598B) | $F_{7:8}$ | 391 |
| Additional population for summer field 2013 | | | | |
| 070063 | IA2070[S] | E06902[R] (Titan x PI 567598B) | $F_{7:8}$ | 46 |
| 090039 | E08907 (PI 567598B) [R] | E09907 (PI 567541B) [R] | $F_{3:5}$ | 190 |

[S] Parent susceptible to soybean aphid

[R] Parent resistant to soybean aphid

(Bales et al., 2013) in the spring of 2011. Seeds from identified recombinants were individually harvested and planted in the field in summer of 2011 to screen for aphid resistance. The $F_2$-derived lines were analyzed for marker association with segregation of the trait.

*Residual heterozygous lines*

Additional screening was done using advanced generation lines obtained from four $F_{6:7}$ plants from a cross between IA2064 and E06906. E06906 was an elite line developed from a cross between Titan and PI 567598B, which inherited the resistant allele from PI 567598B (Mensah et al., 2005). The four $F_{6:7}$ lines possessed residual heterozygosity at the region where *rag3* was mapped by initially screening 941 recombinant inbred $F_{4:6}$ lines with the flanking markers, MSU16-04 (Gm16_6050224_T_C) and MSU16-13 SNP (Gm16_6424067_A_G). Figure 3.1 shows the schematic diagram of the residual heterozygote selection. All seeds from the four RHLs were advanced by genotyping for residual heterozygosity selecting only the line (s) with the shortest heterozygous genotype within the interval. At $F_{9:10}$ generation, four heterozygous lines were selected and allowed to set at least 300 seeds in the field (Fig 3.2). The $F_{10}$ families were genotyped and evaluated for soybean aphid resistance for linkage map analysis (Fig 3.3).

*Evaluation for soybean aphid resistance*

A greenhouse-based aphid resistance screen was conducted in the Plant Science Greenhouse on the MSU campus in the spring of 2011. The greenhouse was optimized at 26/15

**Figure 3.1** Schematic diagram of the selection of residual heterozygous lines for linkage map analysis

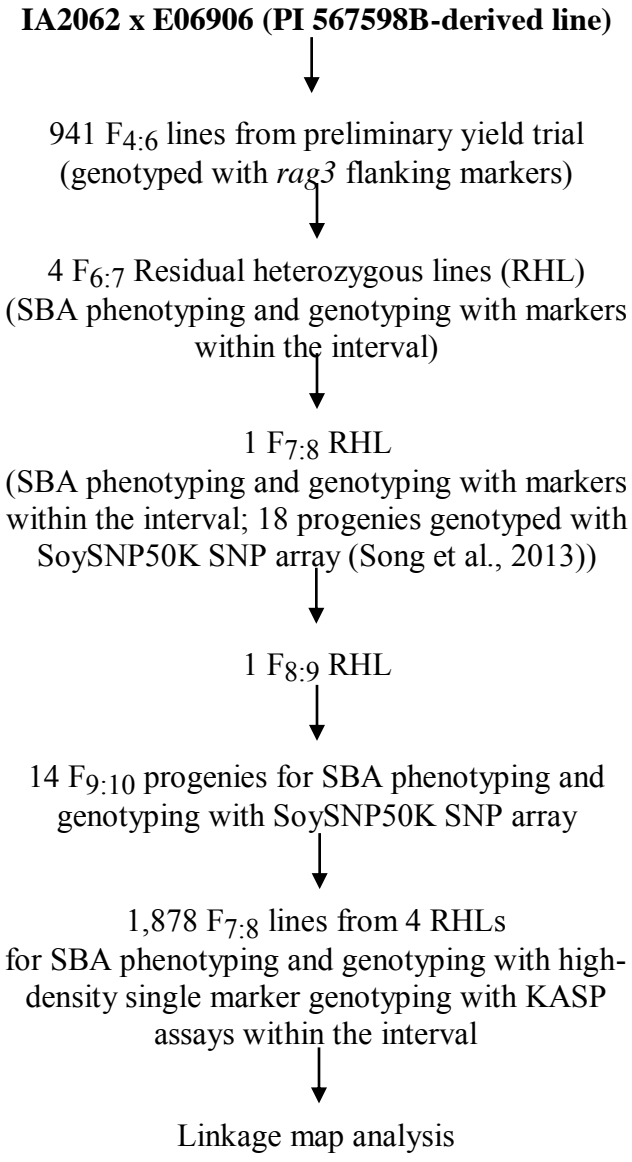


**Figure 3.2** Graphical genotype of 14 $F_{9:10}$ lines with residual heterozygosity. **A**, Genoype of chromosome 16 and **B**, heterozygous region containing the *rag3* interval. **White bars** represent homozygote genotype for the allele from the resistant parent; **gray bars** represent heterozygote genotype; **black bars** represent homozygote genotype for the allele from the susceptible parent.

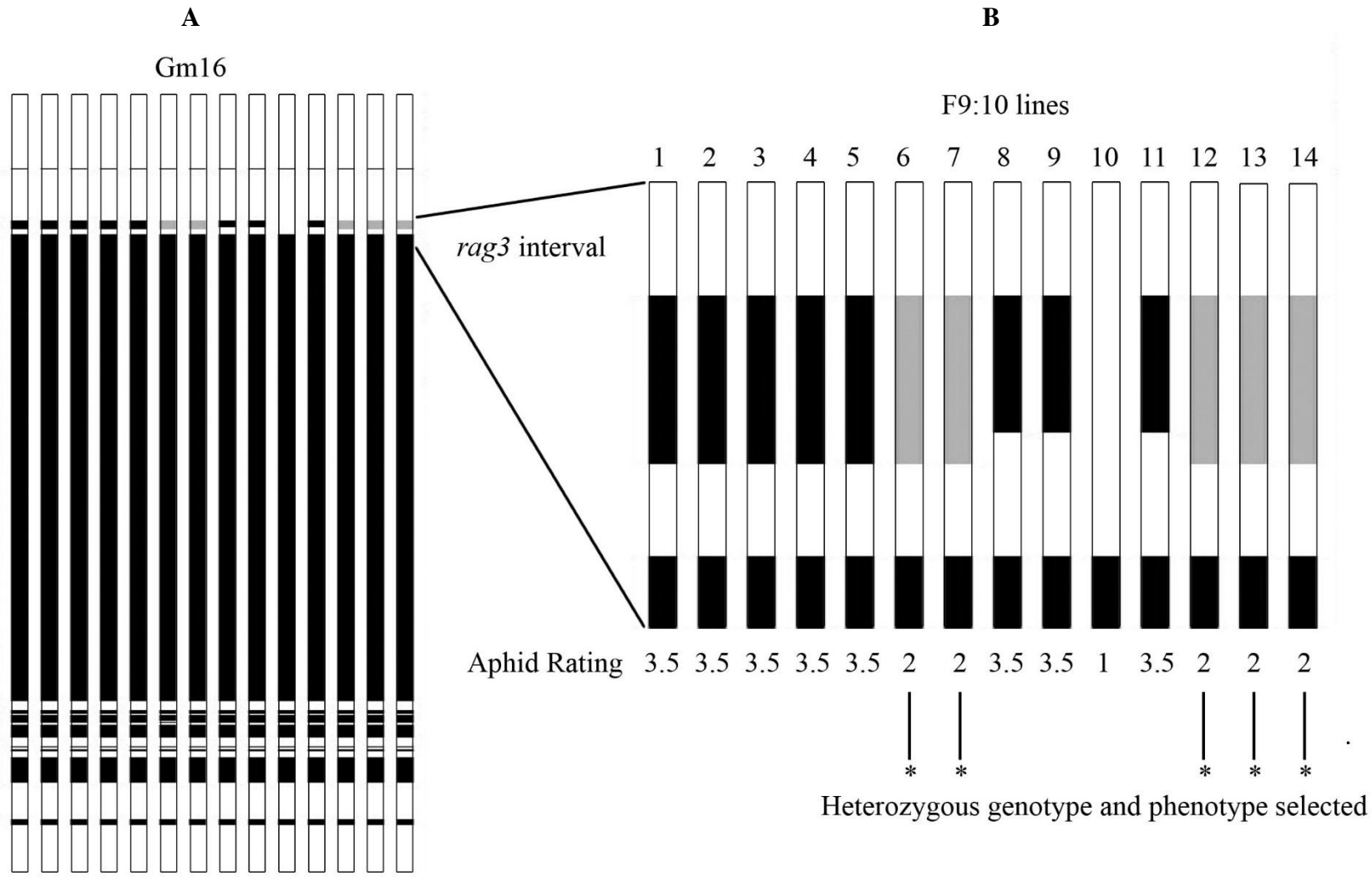
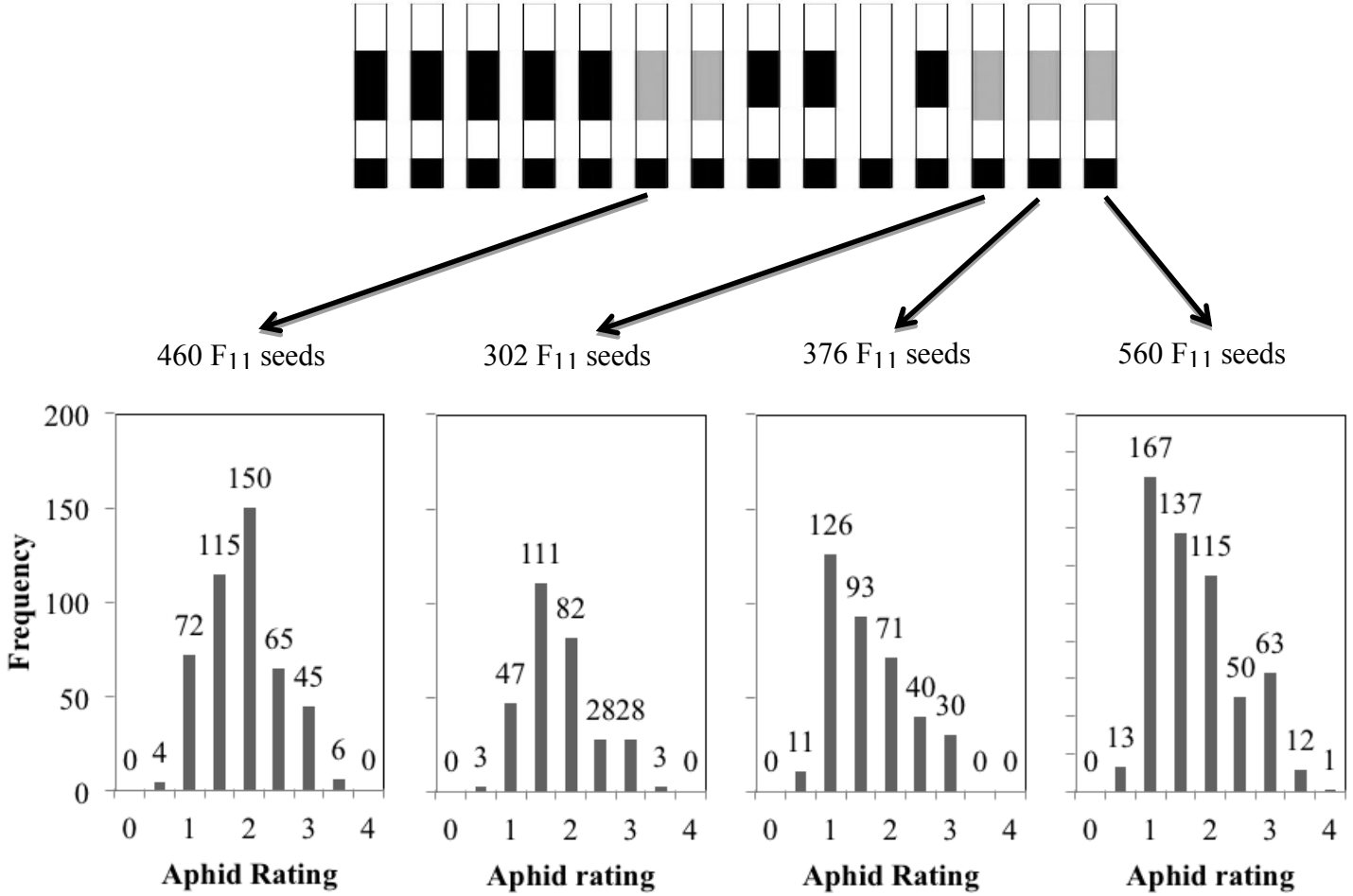


**Figure 3.3** Frequency distribution of the soybean aphid rating phenotype of $F_{10:11}$ lines with residual heterozygosity within the *rag3* interval. **White bars** represent homozygote genotype for the allele from the resistant parent; **gray bars** represent heterozygote genotype; **black bars** represent homozygote genotype for the allele from the susceptible parent.



$^{o}$C day/night temperature and light intensity was extended during the day (14 hr) using sodium vapor bulbs. All $BC_1F_2$ lines from the two populations were planted at eight seeds per pot (105 mm diameter x 125 mm depth). The corresponding parents for each population and resistant checks each had three pots as replicates and arranged randomly within the bench. Two wingless soybean aphids were placed on each soybean plant at the V2 stage. Soybean aphids used for infestation were obtained from a greenhouse colony that originated from the field in 2010 at the MSU Agronomy Farm.

A field-based aphid resistance screen was conducted at the Michigan State University (MSU) Agronomy Farm in the summer of 2011 to assess the phenotype of the progenies of each of the $F_2$ recombinants. Single row plots were set up inside 12.8 x 19.5 m aphid- and predator-proof cages (Redwood Empire Awning Co., Santa Rosa, CA). Each $F_2$ was planted as a single plot having at least 15 plants per line. The same procedure was used to infest aphids greenhouse trials, although the soybean aphid source was a naturally infested field at the MSU Agronomy Farm.

For both trials, visual ratings and calculation of damage index (%DI) were done after four weeks of aphid infestation as previously described by Mensah et al. (2005; 2008). Plants having 0% damage index were considered resistant while 100% damage index indicates susceptibility to soybean aphids.

*DNA extraction and marker analysis*

A quick DNA extraction method was employed to screen the genotypes of all $F_2$ lines at a rapid rate. The youngest non-expanding trifoliate leaves (5-8mm long) were collected for

extraction before soybean aphid infestation. Tissues from each $F_2$ plant were placed in individual wells of a 96-well PCR plate and 100ul 1x TE buffer (10mM Tris-HCl and 0.5M EDTA, pH 7.5) was added. The plate was sealed with foil (3M™) using a heat sealer to retain the moisture. The plate was placed in a pre-heated oven to heat at $94^0$C for 30mins, then centrifuged at 1500rpm for 3min to condense the leaf tissue to the bottom of each well. After centrifugation, the plate was stored in $4^0$C for 30min to overnight. Lysate was aspirated from the plate and diluted ten times with 0.1x TE buffer (pH 7.5) prior to genotyping.

More than 52,000 SNP markers on the SoySNP50 iSelect Infinium assay (Song et al., 2013) for Illumina Bead Chip arrays (Illumina, Inc) were screened for polymorphism between the parental lines for each of the fine mapping populations. Custom primers and hybridization probes for TaqMan SNP genotyping assays were designed through the Custom Taqman® Assays Design tool (www.appliedbiosystems.com). From the Williams 82 genome assembly (Glyma1) available at www.phytozome.net/soybean (Schmutz et al., 2010), 60 bp upstream (5' end) and downstream (3' end) of the identified SNP position were used as target sequences for custom design. This was based on the genomic physical position of the SNP screened for polymorphism among parents (Table 3.2). Endpoint genotyping was performed on the LightCycler® 480 system (Roche Applied Science). For one DNA sample, a total of 3 ul reaction volume was analyzed, comprised of 1.50 ul 2x TaqMan Universal PCR Master Mix, 0.15 ul 10x working stock of SNP genotyping assay and 1.35 ul DNA lysate sample. The parameters used to perform PCR were as follows: $95^0$C for 10 mins to activate enzyme followed by 45 cycles of denaturation at $92^0$C for 15s; and annealing and extension at $60^0$C for 1 min. Genotype calling was performed using the Endpoint Analysis module of LightCycler® 480 Software version1.5.

**Table 3.2** TaqMan SNP assay (www.appliedbiosystems.com) information designed from SoySNP50 iSelect BeadChips.

| TaqMan/KASPar SNP assay | SoySNP50 Infinium assay | Chromosome | Genomic position (bp)[a] | Target sequence[b] |
|---|---|---|---|---|
| MSUSNP16-04 | Gm16_6050224_T_C | 16 | 6,050,224 | 5'CTAGTGGTCGCGCCTGGCAGGCCACCACTTTCACCTCTGTCCCATCGTCCTGTCAAGTCA**[T/C]**GACATGTGTCGCGTTCTGGTGGAATGCGCCCCTCAGAAAAGCGCTTTGTAGTAAAATAAC-3' |
| MSUSNP16_26 | Gm16_6052831_T_C | 16 | 6,052,831 | 5'CTGGCAGGCTACCACTAGTGGTCGCGCCTGGGGCCCACCACTAGTGGTCGCGCCTGGCAGGCCACCACTTTCACCTCTGTCCCATCGTCCTGTCAAGTCA**[T/C]**GACATGTGTCGCGTTCTGGTGGAATGCGCCCCTCAGAAAAGCGCTTTGTAGTAAAATAACAGACCCCCTTGATAAATAAAAATGAAACAGACCCATTTTA-3' |
| MSUSNP16_27 | Gm16_6061510_C_T | 16 | 6,061,510 | 5'ATACAAATACATATAACATATATGTTTTGTGCCTTATAATTACCTCTGCTGCTGGGAAAGCAACCTTTTCTCCCCCTGAAGCAATAGGATTTCCACTCGT**[C/T]**GGAGTCTCGGTTCCCATGGTTGTGAAGTATATGCACTTTTGTGTCTATGGTTGTTAGCAGTATTAAGTAATTAAGAGGAGGCTATGAAATTTTCTTCAAT-3' |
| MSUSNP16_28 | Gm16_6079769_A_G | 16 | 6,079,769 | 5'CATAGAGGGCTTGAGCGATGTCTTGATCGTTGACCGAAAGTTGTGATCAAGTGTGGTAGTGTACGTCATCTCTCTCAGTTCCCCCACGATTCCTAATAAC**[A/G]**CATCAATATTCTTCTCTTTTGAAACACACCAATTATATATATTTTCTTTTCTTTTTATTCATTTTCTTTCATCGTCATAATTTTAATTTTTTTATCTAAA-3' |
| MSUSNP16_35 | Gm16_6139859_A_G | 16 | 6,139,859 | 5'CCAAATCGTATCTATAAGTCTCATAAGCTCAATCGGTAAGCCTCATAAATTTACGGATTGATCAATCCATAAGCCTCGGCCAATTTGTATGACACTTATG**[A/G]**ATTGTCCAATCTATATCAGTTTTATGGAAAAGTAA |

**Table 3.2**
**(cont'd)**

| | | | | |
|---|---|---|---|---|
| | | | | AACAGGAATCGCGCACAATTGCTGAGTGTACCAG |
| | | | | CAATTTTGCTGGGTGCGCGTAGCAAGAGCTC-3' |
| MSUSNP16_36 | Gm16_6179363_T_G | 16 | 6,179,363 | 5'GCGGTGGTTTGTTTAATCGAATGCGATCTGGTTC |
| | | | | CGGCGATTGAATATATTTTTTTTAGGTGCGTGGTG |
| | | | | TAGTTCAACTACTTGAAGCTCTGTTTCAATT**[T/G]**T |
| | | | | GTTTGATTTAGTACATTGGTTGAGGTGAAATTTGC |
| | | | | AGTTATTTGAAGACTTTGGAGTAGAAGCTTCGGTT |
| | | | | GAGAATGCCGGGTTTTAGGGCGAAAATTA-3' |
| MSUSNP16_37 | Gm16_6184915_A_G | 16 | 6,184,915 | 5'CTTATTCAAGGCTCACGTTGAGGACTAGACATCT |
| | | | | TGAGCGTGAAGTTTGCAGGATTGGACATTTGCGG |
| | | | | GTGGTCCAATAACAGCTCAACTCACTAGGATA**[A/** |
| | | | | **G]**GCTCTGATACCATCTTAGAAAGTGGTTATGGGT |
| | | | | CTAACTCAACTCTACAAAATGGCTTGTAAGGTGA |
| | | | | GGGTTGTCCTCCACTTATATACACTTTTAAGGC-3' |
| MSUSNP16_38 | Gm16_6192576_T_G | 16 | 6,192,576 | 5'ATCACATTAATTTCATAGTACTTAATTATGTTAA |
| | | | | ATTCACTTTTGTAGAAACGATCTAAGGAATTTTTT |
| | | | | TTTTAGTTACTTTAGGCCATGCTTGATGGAG**[T/G]**A |
| | | | | TTGTTACTGAGGGGATTGATGCTTGAATCTTCACA |
| | | | | CAGGATTCATTGAAAGTCAAGCAGGCAGACATTG |
| | | | | AAGCCTATTGATTATCTCATTTTAGTTGTA-3' |
| MSUSNP16_39 | Gm16_6214642_C_T | 16 | 6,214,642 | 5'ACACGATTGAAGAAAATTGAAAAAGAAATACTA |
| | | | | CTACTATTGAAGAAAGTTGAAAAAGAAATACCAG |
| | | | | TACCTTGTCCATTTCTGCTTCTTTTGCGGGTTG**[C/T** |
| | | | | **]**GGAAAGTGTTCCAATTTAATGCTGGTCCTAAGTC |
| | | | | CTAACCAAGTTATCAAGATTCAAGCCGTGGCTTGC |
| | | | | AGGTAATATTTAACTCTCTGTTTAGACTTTA-3' |
| MSUSNP16-10 | Gm16_6262227_C_T | 16 | 6,262,227 | 5'5'CCCATGATGTCATGAGGTGTAAACTTGTTAAG |
| | | | | ACATATCAAACTTAGGGTTTAAGTTAAC**[C/T]**AGA |
| | | | | TCCGAAAAAGCTGCCACTATAGTGCCTTCTCTTTG |
| | | | | AGTATGTGGTAATTATTGATTG-3' |
| MSUSNP16_18 | - | 16 | 6,270,557 | 5'TCACAATTGATATCACCTCCTTATCAATAGGGCT |

**Table 3.2**
**(cont'd)**

| | | | | |
|---|---|---|---|---|
| | | | | TGTTGGNTGTGGAAGACGTTCGTCCAATTTAACCA TCAATGACATATGATCAAGTGTTGAGGTCAT**[A/T]** GTGGACGAAGATGATAGCAATAGAGAAGATGTAA CATCACCAGGGTGCTCTCCAAACAATATTTCCAAC GCAAAGACCCCAAAACTATACACATCACATT-3' |
| MSUSNP16_22 | - | 16 | 6,342,000 | 5'GACTTTTAATACCACTGTTGGGAAAAACTCGAT GGGAGGAAATACTGGGGAGATTCCTGTGGACCAC GAGCCACCACATAAGAAAACCTAGTACCACACT**[A /C]**TAACCCAAAACCTTAAGGCTTAGGTTTATGAGT CTTCTCTTCACTTATATGGTGCTCAGCCTTTCCACT TCTACCCGATGTAGGACTTCACCTCACACTT-3' |
| MSUSNP16-11 | Gm16_6413214_A_G | 16 | 6,413,214 | 5'GCATGGCGCGTGACACATTCAACAATGTTCATT GGGTAGCCCGTCTTAGTAGGTTACGCA**[A/G]**CAGG TAAGTTAAGACGATGTATTTGAAAACACTAGAAA TTTTGAATGTTAACGACGTTTT-3' |
| MSUSNP16-12 | Gm16_6423098_G_A | 16 | 6,423,098 | 5'AAATTATGACCCAATTAGATGCAAATGTCCTTG CTTCCTGTATTGAAACACCCCCTACGA[**G/A**]TCCT AACACCCCATTGTGTACGTCCCTTTTCAAGCCCAC CTCATACCATAAAGATGTAAC-3' |
| MSUSNP16-13 | Gm16_6424067_A_G | 16 | 6,424,067 | 5'CAACTTCCTGACACCACTCGCAGTCCCTGAGATT CGGCGGCGGCTAGCGTCGGTGGCGGC**[A/G]**GCGG CGGACGAGGACCCTCCGCAATCGCCGTCGTCGTTC ACTTTCTCGTCGGAGGGGGAG-3' |
| MSUSNP16_43 | Gm16_6431101_A_C | 16 | 6,431,101 | 5'GCATCAAAGAATGTATTAACAATACAATGACAA ACATAGAATTCAGCAAAGAGCTCTCTTATACTAGC TGAAGTACAAAGCATAGCACCAAGAACAGCAG**[A/ C]**TGAGATATTATGCTAGTATTAACAAGTAAATCA ACAAAACACCATGCATATATCATGGCAGGTGTGA ACTATATAATCCTTTCTCAACCCAAGCTTCAAA-3' |
| MSUSNP16_46 | Gm16_6469551_A_C | 16 | 6,469,551 | 5'ATTAAAAGGGGAGTTACATGAGATTAGGTTCTG |

**Table 3.2**
**(cont'd)**

| | | | | |
|---|---|---|---|---|
| | | | | TCGTTTTCAACCTCAACAAATGAGGAGATTGATTG<br>CTCATAATTTGAGTACAATAGCTCAGAGAAAT**[A/**<br>**C]**ATACAAAATGAATTTAGTCTAACAACCTAGAAA<br>TAATACTCTTTCTCTCTTAAAGAATCGCAACTTCA<br>ATTTGTGCTAAGATGATTCTTCCTTCACAAAC-3' |
| MSUSNP16_48 | Gm16_6680549_G_A | 16 | 6,680,549 | 5'TGCAAGCCAATGAAACATGAAATAGAAAGGCC<br>AACTGAAGTAAGATAAGAAACATAAAAAGGGTAC<br>TTATTTATTGTAGATTGTCCAACTACATGATCAC**[G**<br>**/A]**CATAGTTGGCTGACTTGTCATCACTCCTGTATT<br>GAATCAACGTTAACTCATCAATTGTAGCAAGCAC<br>ACCAATCAACATCTACATAAAAAATTACACTTG-3' |
| MSUSNP16_49 | Gm16_6713173_T_G | 16 | 6,713,173 | 5'ACTTTCTTCAGCATTCATTCTGTAGCAGCAGGAA<br>GTATTAGATTGCAGCGATGGCTCTCATCGGGTGGA<br>TTTTCCTAACTTGGTGTGTCTGGCACCCCAG**[T/G]**A<br>AAAGGTTGGTCTTCTCAAATGGCAGCCTAAATATT<br>AACAAGCTGATGGAGGACGCTATTATCTTCTGCTG<br>GACTTGGCTGAGAAACCTCCAAAAAAGAT-3' |

[a] Genomic position of single nucleotide polymorphism on the Williams 82 genome assembly, Glyma1 (Schmutz et al., 2010).

[b] Target sequence for TaqMan/KASPar custom design with 100-bp upstream and downstream of the single nucleotide polymorphism. SNPs in corresponding wild-type and mutant-alleles are in brackets [ ].

For the identified recombinants, good quality DNA was needed to run whole genome SNP genotyping analysis with the SoySNP50 iSelect Inifinium assay (Song et al., 2013). The CTAB (hexadecyltrimethyl ammonium bromide) method was used to extract DNA from $F_{2:3}$ samples as described by Kisha et al. (1997). Determination of DNA concentration was done using the Quant-iT™ Picogreen® dsDNA Assay Kit (Invitrogen, USA) and quantified using BioTek Multi-Detection Microplate Reader (Biotek, USA). Each DNA sample was normalized to 50 ng/ul for Infinium assay and performed following manufacturer's protocol. Infinium BeadChip data analysis for SoySNP50 iSelect was performed using the GenomeStudio Genotyping module.

Polymorphic markers were determined between susceptible and resistant parents and used to genotype each progeny of the identified recombinants. For this phase, KASPar-On-Demand (KOD) assays were utilized by custom-designed primers for the SNP genotyping assays (www.lgcgenomics.com) using 100-bp upstream (5' end) and 100-bp downstream (3' end) of the identified SNP position as target sequences. Endpoint genotyping was performed on the LightCycler® 480 system (Roche Applied Science). For one DNA sample, a total of 3 ul reaction volume was analyzed, comprised of 1.50 ul 2x KASP Master Mix, 0.045 ul 1x SNP genotyping primer mix and 1.455 ul ~20ng/ul DNA sample (from CTAB extraction protocol). The parameters used to perform PCR were as follows: $94^0$C for 15 mins to activate enzyme followed by 20 cycles of denaturation at $94^0$C for 10s, annealing at $57^0$C for 5s, extension at $72^0$C for 10s and another 18 cycles of $94^0$C for 10s, $57^0$C for 20s, and extension at $72^0$C for 40s. Similarly, genotype calling was performed using the Endpoint Analysis module of LightCycler® 480 Software version1.5 (Roche Applied Science, Indianapolis, IN, USA).

*Whole genome re-sequencing and SNP discovery*

Leaf tissue samples were collected from young soybean seedlings of PI 567598B at VC stage grown in the greenhouse (27°C/24°C day/night and 16 h/8 h light/dark). DNA extraction was performed using Promega's Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) and quantified using Quanti-IT$^{TM}$ Picogreen® dsDNA Quantitation reagent (Life Technologies, Carlsbad, CA, USA). DNA concentration was normalized at 50ng/ul and a total of 5ug of DNA per sample was submitted for pooling and library preparation at the Research Technology Support Facility at Michigan State University (East Lansing, MI). The sample was indexed during library preparation using Illumina TruSeq DNA Sample Prep kit (Illumina, Inc., San Diego, CA, USA). Three other soybean accessions were pooled with PI 567598B into one lane of the Illumina HiSeq 2000 flow cell (Illumina, Inc., San Diego, CA, USA).

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used for quality control check to visually examine sequence quality. FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) was used to remove Illumina adapter sequences (fastx_clipper program) and requiring a minimum sequence length of 20bp after trimming.

Bowtie version 0.12.7 was used to map the cleaned short reads into the reference genome, Williams 82 (Gmax_109 assembly data obtained from www.phytozome.com/soybean). The parameters used include using the paired-end mode for paired end reads and –v mode for single end reads. Only 2 mismatches were allowed for a read to map to the reference sequence. Reads that map only once (unique alignments) were processed for SNP calling.

Alignments for reads that mapped uniquely to the chromosomes were processed using the sort, index, and pileup programs within SAMtools version 0.1.12a to generate unfiltered pileup files that are then filtered for quality using the varFilter option. The SAMTools varFilter

parameters considered for high quality SNPs are: a) should at least have 3 read depth coverage (minimum), b) should at most have 20 read depth coverage (maximum), c) per base SNP quality should be more than 20 phred score (at least 1/100 error rate).

Custom KASP-by-design (LGC Genomics LLC, Beverly, MA, USA) SNP assays were submitted by obtaining 100-bp flanking sequences of the identified SNPs from the above NGS data. Flanking sequences of the SNPs were obtained from www.phytozome.com/soybean. KASPar SNP genotyping was performed using the 'Endpoint genotyping' module of LightCycler®480 system (Roche Applied Science, Indianapolis, IN, USA).

*Statistical and linkage map analysis*

Pearson correlation computation and one-way analysis of variance for the phenotype and genotype of the progenies of recombinants was calculated with the R Statistical package (R Development Core Team, 2011). A linkage map of the residual heterozygous line families was constructed with the Kosambi function and a LOD score of 3 using JoinMap 4.0 (Van Ooijen 2006) utilizing SNP markers designed within the *rag3* interval. Composite interval mapping (CIM) was done to validate the *rag3* aphid resistance loci using an RHL families as a single population using QTL Cartographer V2.5 (Wang et al. 2008). Forward and backward regression method was used to select markers as cofactors with a walking speed of 0.5 cM. The LOD threshold was empirically determined by running a 1,000-permutation test at 5 % probability level (Churchill and Doerge 1994).

## Results and Discussion

*Screening for recombination breakpoints*

This study reports fine-mapping of the aphid-resistance loci, *rag3* and *rag1b*, located on chromosomes 16 and 7, respectively. A similar approach using $F_2$ plants was done for *Rag1* (Kim et al., 2009) and *Rag2* (Kim et al., 2010). From 4,104 $BC_1F_2$ plants in the recombinant screening, 107 (from population 090004) and 94 (from population 090068) lines were selected for planting in the field and soybean aphid evaluation of progenies. These were lines with crossovers between MSU16-04 (Gm16_6050224_T_C) and MSU16-13 SNP (Gm16_6424067_A_G) markers using the quick DNA extraction method and assayed with TaqMan. The quick DNA extraction method proved to be a good protocol to screen thousands of lines using TaqMan assay, but it was not robust enough for SSR markers and Infinium assay analyses. Thus, a DNA pool of all $F_3$ progenies for each $F_2$ line was collected in the field for CTAB extraction and re-analyzed with TaqMan assay. Confirmed genotypes were subsequently used for SoySNP50 Infinium assay. Out of the 107 and 94 lines, 34 and 21 $F_2$ lines, respectively, were selected for SoySNP50 Infinium assay. Among the $F_2$ lines genotyped with the Infinium assay, six lines from the 090004 population and three lines from the 090068 population had recombination events in the r*ag3* region (Table 3.3).

**Table 3.3** Recombination breakpoints among identified recombinants that mapped the position of *rag3* on Chromosome 16. Bold letters represent the breakpoints and italicized letter are loci with TaqMan SNP assays used for marker association with progeny phenotype.

| Line ID | Gen | Ph[a] | Taqman/KASPar assay[b] SNP marker (MSUSNP16- ) and physical positions[c] | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | -04 | -26 | -27 | -28 | -35 | -36 | -37 | -38 | -39 | -10 | -18 | -22 | -11 | -12 | -13 | -43 | -46 | -48 | -49 |
| 04-2-653 | $F_2$ | Seg | S | **S** | ***H*** | H | H | H | H | H | H | H | H | H | H | H | H | H | H | H | H |
| 04-2-229 | $F_2$ | Seg | R | R | R | **R** | ***H*** | H | H | H | H | H | H | H | H | H | H | H | H | H | H |
| 04-2-471 | $F_2$ | R | H | H | H | ***H*** | **R** | R | R | R | R | R | R | R | R | R | R | R | R | **R** | **S** |
| 04-2-742 | $F_2$ | R | H | H | H | ***H*** | **R** | R | R | R | R | R | R | R | R | R | R | R | R | **R** | **S** |
| 04-2-466 | $F_2$ | R | H | H | H | H | H | ***H*** | **R** | R | R | R | R | R | R | R | R | R | R | **R** | **S** |
| 68-1rem-39 | $F_2$ | Seg | H | H | H | H | H | H | H | H | H | H | H | H | H | H | H | ***H*** | **R** | R | R |
| 68-1rem-168 | $F_2$ | R | H | H | H | H | H | H | H | ***H*** | **R** | R | R | R | R | R | R | R | R | R | R |
| 68-5-146 | $F_2$ | Seg | R | R | R | R | R | R | R | **R** | ***H*** | H | H | H | H | H | H | **H** | **R** | R | R |
| 11-831-7 | $F_{7:8}$ | Seg | H | H | H | H | H | H | H | H | H | H | H | H | ***H*** | **R** | R | R | R | R | R |
| 04-1rem-30 | $F_{4:5}$ | R | H | H | H | H | H | H | H | H | H | H | ***H*** | **R** | R | R | R | R | R | R | R |

[a] Phenotype of the recombinant lines based on soybean aphid evaluation of progenies. **Seg** – segregating phenotype, **R** – resistant

[b] SNPs from SoySNP50 iSelect BeadChips converted to TaqMan/KASPar SNP assays. **R** – allele from resistant parent, **S** – allele from susceptible parent, **H** – heterozygous genotype

[c] Physical position (in Mbp) of SNP markers in SoySNP50 based on Glyma1 assembly of soybean Williams 82 (www.phytozome.net/soybean).

*Fine mapping of rag3*

Results from the Infinium whole genome SNP genotyping revealed SNPs that were polymorphic within the region of interest and the density delimited the position of *rag3* into a 231-kb interval (Fig 3.4 and Table 3.3).

The leftmost (towards the telomere) border of the *rag3* region was first defined by the line 04-2-653. Marker test showed association of the trait with segregation of the marker (Table 3.4) and thus, *rag3* was positioned at the right side of MSUSNP16-26 (Gm16_605831_T_C). The rightmost (towards the centromere) border of the *rag3* region was positioned by lines 04-2-471, 04-2-742, and 04-2-466 with a breakpoint between MSUSNP16-48 (Gm16_6680549_G_A) and MSUSNP16-49 (Gm16_6713173_T_G). Marker association with the trait revealed that all $F_3$ lines were resistant and this positioned *rag3* at the left side of MSUSNP16-49 (Gm16_6713173_T_G). The lines 04-2-229, 04-2-471 and 04-742 further narrowed the position of *rag3* on the left border. Marker association at the recombination breakpoint between MSUSNP16-28 (Gm16_6079769_A_G) and MSUSNP16-35 (Gm16_61339859_A_G) proved that *rag3* was located at the right side of MSUSNP16-28 (Gm16_6079769_A_G). Lines 68-1rem-39 and 68-5-146 narrowed the position of *rag3* at the right border. These lines were segregating for the trait and marker association test showed that there was significant association on the phenotype and segregation of the marker. This demonstrates that *rag3* may be located at the left side of MSUSNP16-46 (Gm16_6469551_A_C). The line 04-2-466 had a breakpoint between MSUSNP16-36 (Gm16_6179363_T_G) and MSUSNP16-37 (Gm16_6184915_A_G). The marker test did not reveal association of the trait with the segregating marker since the line had a resistant phenotype, which supported that *rag3* was located on the right side of MSUSNP16-36 (Gm16_6179363_T_G). The lines, 68-1rem-168 and 68-5-146, narrowed the

**Figure 3.4** Fine mapping the location of *rag3* on chromosome 16. **A** Genetic map of chromosome 16 with SNP AND SSR markers. Genetic positions are in centimorgans (cM). **B** Recombination breakpoints of recombinant lines from SNP genotyping by Infinium, TaqMan and KASPar assays. Genomic physical positions are in (Mbp) on the Williams 82 genome assembly, Glyma1 (Schmutz et al., 2010). **White bars** represent homozygote genotype for the allele from the resistant parent; **gray bars** represent heterozygote genotype; **black bars** represent homozygote genotype for the allele from the susceptible parent; **cross-hatched bar** represent the position of *rag3*; **dotted horizontal lines** separate recombination bins. Phenotype of recombinants as confirmed by progeny test: **Seg**, segregating phenotype; **R**, resistant phenotype.

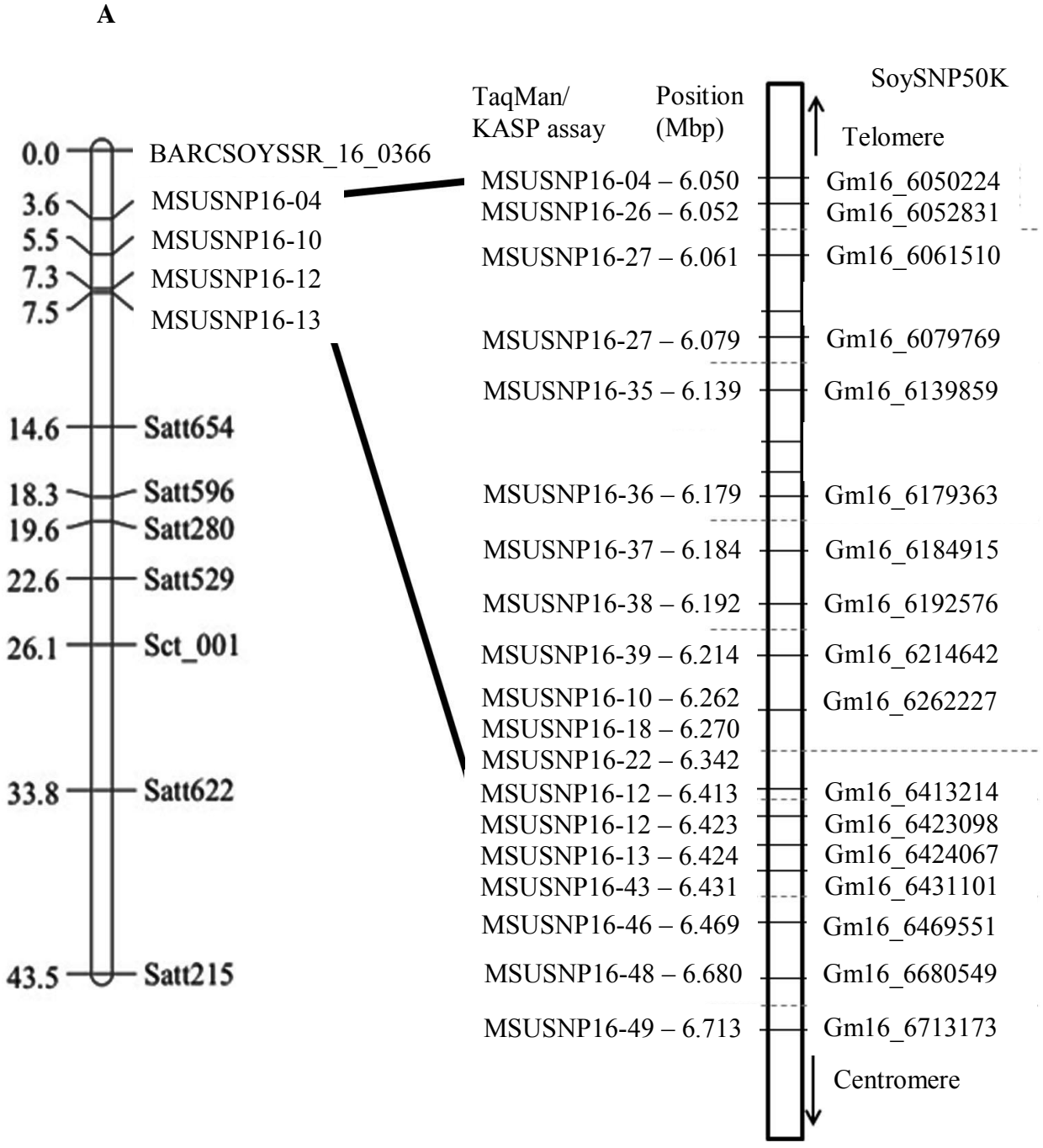


**Figure 3.4 (cont'd)**

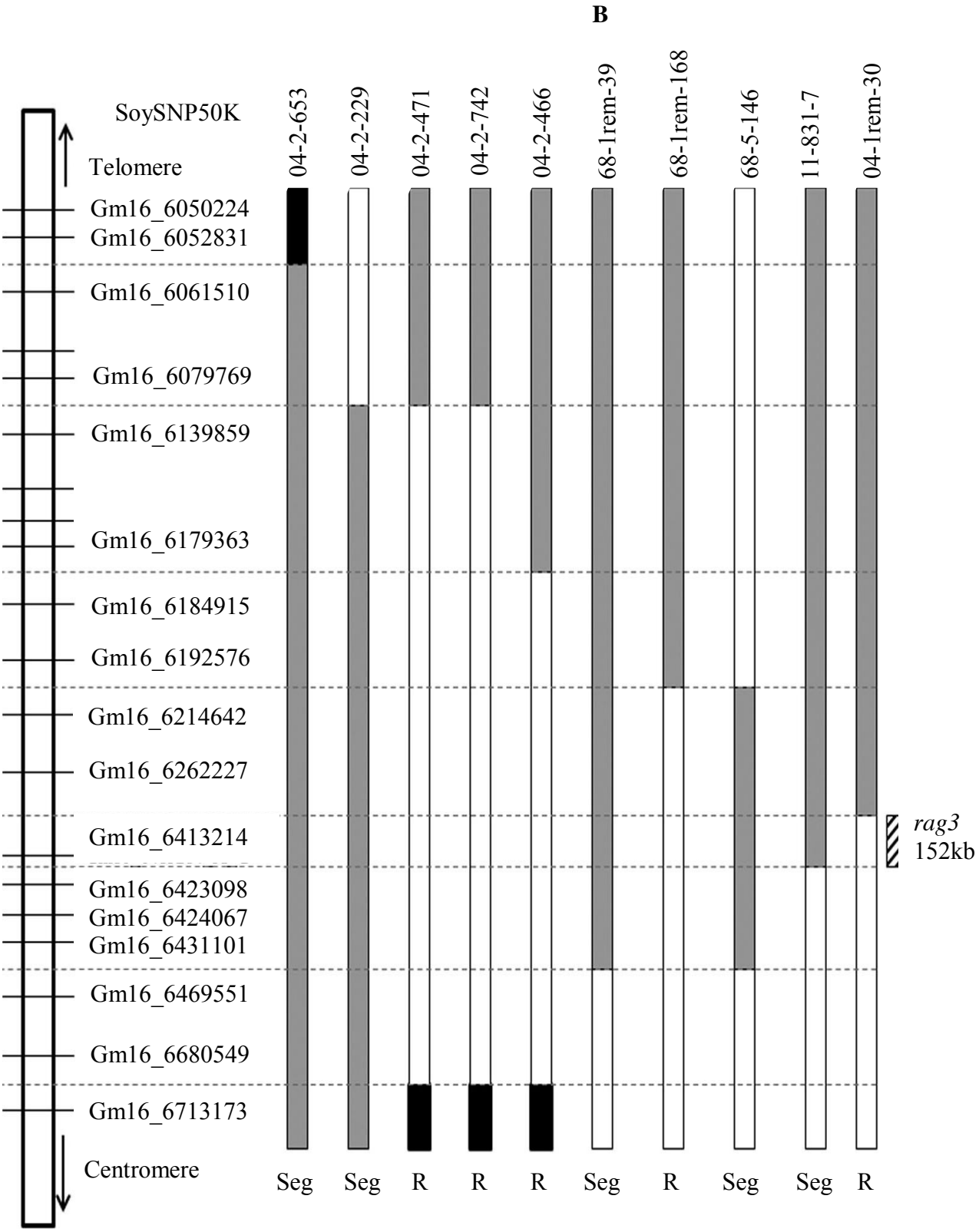


**Table 3.4** Phenotype of the progenies of recombinants and their association with the marker tested.

| Line ID | Generation | Phenotype [a] | No. of progenies | Markers tested (MSUSNP 16- ) | Correlation of markers tested | |
|---|---|---|---|---|---|---|
| | | | | | $R^{2}$ [b] | Pr(>F) [c] |
| 04-2-653 | $F_{2:3}$ | Segregating | 14 | SNP16-27 | 0.74 | 0.003** |
| 04-2-229 | $F_{2:3}$ | Segregating | 13 | SNP16-35 | 0.74 | 0.003** |
| 04-2-471 | $F_{2:3}$ | Resistant | 13 | SNP16-28 | 0.26 | 0.387 |
| 04-2-742 | $F_{2:3}$ | Resistant | 12 | SNP16-28 | 0.32 | 0.303 |
| 04-2-466 | $F_{2:3}$ | Resistant | 12 | SNP16-36 | 0.20 | 0.545 |
| 68-1rem-39 | $F_{2:3}$ | Segregating | 10 | SNP16-43 | 0.76 | 0.009** |
| 68-1rem-168 | $F_{2:3}$ | Resistant | 14 | SNP16-38 | 0.14 | 0.632 |
| 68-5-146 | $F_{2:3}$ | Segregating | 15 | SNP16-39 | 0.87 | <0.0001*** |
| 11-831-7 | $F_{8:9}$ | Segregating | 16 | SNP16-11 | 0.79 | 0.0002** |
| 04-1rem-30 | $F_{5:6}$ | Resistant | 14 | SNP16-10 | 0.10 | 0.782 |

[a] Phenotype of the recombinant lines based on soybean aphid evaluation of progenies

[b] $R^{2}$ value of the marker association

[c] Level of significance of the marker association

SNP marker border at the left side of *rag3* gene. Marker association of both lines showed that the gene was on the right side of MSUSNP16-38 (Gm16_6192576_T_G). The line 11-831-7, had a breakpoint towards the centromere side that defined the right border of *rag3*, which was MSUSNP16-12 (Gm16_6423098_A_C). This delimited *rag3* at an interval of 231 kbp between MSUSNP16-38 (Gm16_6192576_T_G) and MSUSNP16-12 (Gm16_6423098_A_C).

Lam et al. (2010) reported that cultivated soybeans have an average linkage disequilibrium of ~150kbp. Since it cannot be ruled out that recombination may still exist within a 231-kbp interval, the interval of *rag3* needs to be narrowed down to less than 200 kbp and will need more SNP markers to resolve breakpoints. MSUSNP16-18 (6,270,557 bp) and MSUSNP16-22 (6,342,000 bp) were developed using SNP discovery bioinformatics method by re-sequencing the whole genome of PI 567598B and aligning the reads to the reference genome, Williams 82 genome assembly Glyma1 (Schmutz et al., 2010). Using these two SNPs, a breakpoint from the line 04-1rem-30 delimited *rag3* down to 152 kbp. Among 2,214 individual lines examined in the 090004 population (Table 3.1), there was only one recombinant within a 231-kbp interval, thus recombination frequency can be calculated as (1/2,214)*100 = 0.045%. With that chance, recombination was close to 0% within the152-kbp interval that contains *rag3*. The tightly-linked markers that flank the final interval of *rag3* are MSUSNP16-18 (6,270,557 bp) and MSUSNP16-12 (Gm16_6423098_A_C).

Bales et al. (2013) reported that significant interaction existed between *rag3* and *rag1b* loci. To eliminate the confounding effects of two epistatic loci in a segregating population, a population developed from four selfed heterozygous lines forming $F_{10}$ families was used. These were lines selected for heterozygosity in *rag3* region and homozygous for the susceptible parent's allele in *rag1b*. Validation using residual heterozygous lines derived from E10011

population show that *rag3* was mapped close to the region as the $F_2$ fine mapping results (Fig 3.5). Without confounding effects from *rag1b* locus on chromosome 7, genetic segregation of the phenotype for soybean aphid ratings shown on a histogram (Fig 3.3) follow a 1:2:1 ratio similar to the expected segregation controlled by a co-dominant gene (Chi-square test data not shown). This finding supports that *rag3* alone was sufficient to control resistance against SBA in PI 567598B. An inheritance study of PI 567598B conducted by Mensah et al. (2008) categorized $F_2$ lines with a rating greater than 1.5 as susceptible that fit a 15:1 ratio for a two recessive gene model; however, we found that lines with heterozygote *rag3* genotypes were rated to be moderately resistant (~2-2.5). Using a large population without the confounding effect of *rag1b* locus, we found that the genetic control of *rag3* was partially dominant in nature.

*Fine mapping of rag1b*

Recombinants from *rag3* screening were used to investigate fine mapping for *rag1b* (Table 3.5). Markers flanking the *rag1b* region from the previous study by Bales et al. (2013) cannot be used for the 090004 and 090068 population. New polymorphic markers from the SoySNP50K array had to be developed for the 090004 population within the *rag1b* region. Interestingly, none of the polymorphic markers from 090004 population are shared by the 090068 population.

When investigating for recombinants, none of the *rag3* recombinants had a recombination within the *rag1b* interval, which was not surprising. Heterozygous $F_2$ lines were further selfed into the next generation and investigated for recombination among progenies. It was found that the presence of *rag3* could confound the analysis of *rag1b* recombinants. For

**Figure 3.5** Genetic linkage map and QTL position of *rag3* using composite interval mapping of 1,682 $F_{10:11}$ lines derived from four residual heterozygous families of the cross IA2064 and E06906. The LOD threshold is 1.60.

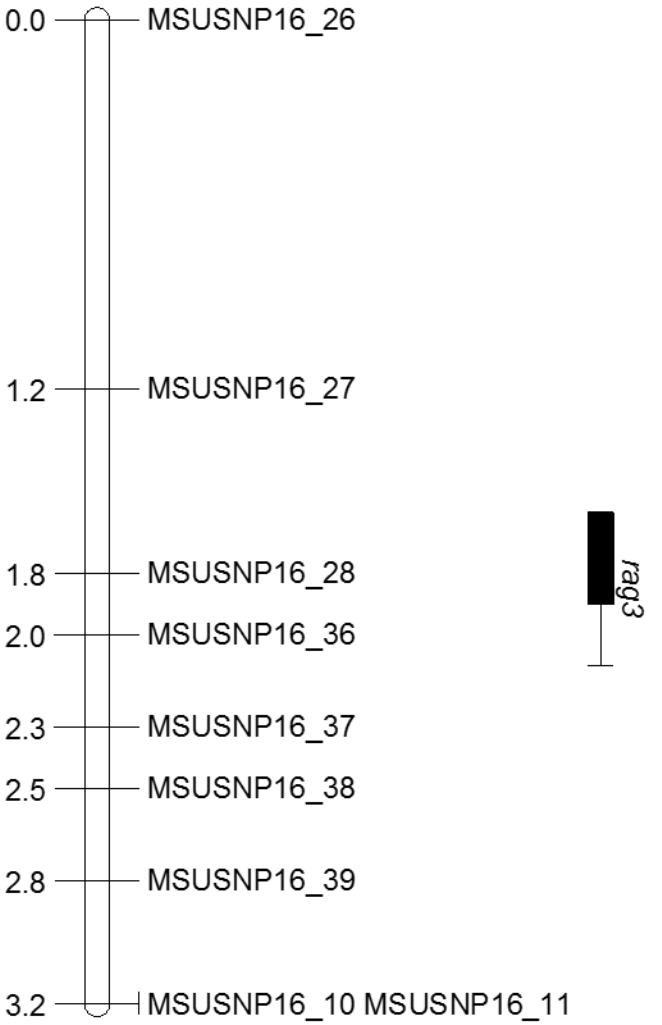


**Table 3.5** Recombination breakpoints among identified recombinants that mapped the position of *rag1b* on Chromosome 07.

| Line ID | Gen | Pheno[a] | KASPar assay[b] (MSUSNP07- ) Soy50K marker names[c] and physical positions[d] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | **07-1**[b] | | | | | **07-2** | | |
| | | | Gm07_ 5484977_ T_G[c] | Gm07_ 5519521_ G_A | Gm07_ 5636973_ T_C | Gm07_ 5763368_ A_G | Gm07_ 5863012_ C_A | Gm07_ 5944283_ A_G | Gm07_ 5961174_ C_T | Gm07_ 6016358_ A_G |
| | | | 5.484[d] | 5.519 | 5.636 | 5.763 | 5.863 | 5.944 | 5.961 | 6.016 |
| 04-2-653 | $F_2$ | Seg* | H | H | H | H | H | H | H | H |
| 04-2-229 | $F_2$ | Seg* | H | H | H | H | H | H | H | H |
| 04-2-471 | $F_2$ | R* | R | R | R | R | R | R | R | R |
| 04-2-742 | $F_2$ | R* | R | R | R | R | R | R | R | R |
| 04-2-466 | $F_2$ | $R^{NS}$ | S | S | S | S | S | S | S | S |
| 68-1rem-39 | $F_2$ | $Seg^{ND}$ | - | - | - | - | - | - | - | - |
| 68-1rem-168 | $F_2$ | $R^{ND}$ | - | - | - | - | - | - | - | - |
| 68-5-146 | $F_2$ | $Seg^{ND}$ | - | - | - | - | - | - | - | - |
| 11-831-7 | $F_{7:8}$ | $Seg^{NS}$ | S | S | S | S | - | S | - | - |
| 04-1rem-30 | $F_{4:5}$ | R* | R | R | R | R | R | R | R | R |

[a] Phenotype of the recombinant lines based on soybean aphid evaluation of progenies (Seg – segregating, Res – resistant)

[b] SNPs from SoySNP50 iSelect BeadChips converted to KASPar SNP assays

[c] Marker name from SoySNP50K genotyping array (Song et al., 2013)

[c] Physical position (in Mbp) of SNP markers in SoySNP50 based on Glyma1 assembly of soybean Williams 82 (www.phytozome.net/soybean)

- Monomorphic SNP between parents

**Table 3.5 (cont'd)**

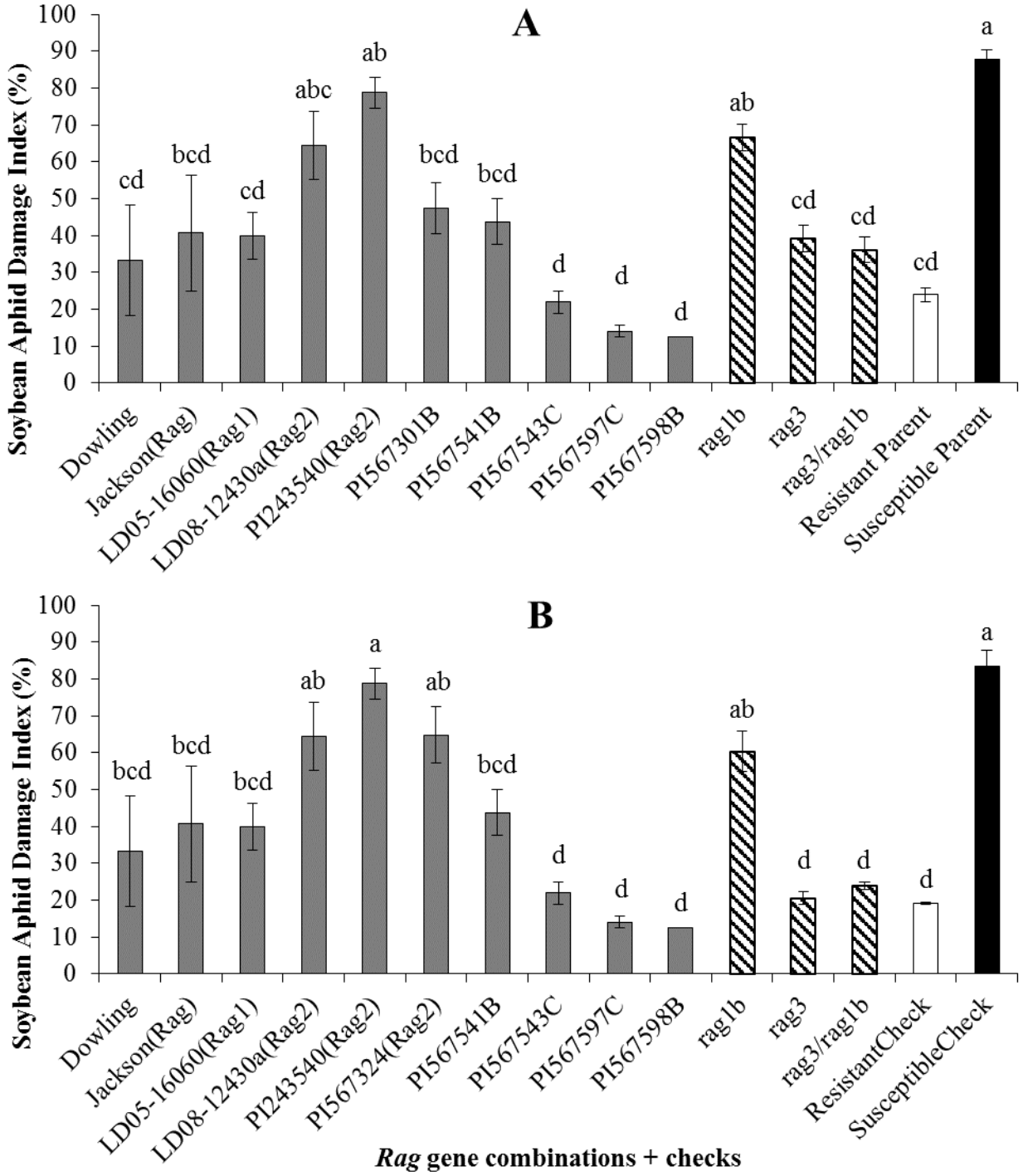
* Phenotype and genotype are associated
NS Phenotype and genotype are not associated
ND No data for genotype

example, the line 04-2-466 containing a homozygous susceptible genotype in *rag1b* but homozygous resistant in *rag3* conferred resistance against soybean aphid (Table 3.5). This further supports that *rag3* was sufficient to control resistance against SBA especially in the field (Bales et al., 2013). The *rag3* heterozygous line from E10011 population, 11-831-7, had a susceptible genotype in *rag1b* but its progenies are clearly segregating for soybean aphid damage in association with *rag3* genotype (Fig 3.3).

One of the reasons that *rag1b* alone may not be effective can be due to the genetic background of the susceptible line. This has been shown on findings from Russian wheat aphid and spotted alfalfa aphid (Randolph et al., 2005; Kamphuis et al., 2013). Van der Westhuizen et al (1998) found that the resistance performance of *Dn1* was dependent on whether it was bred into 'Tugela', 'Betta', or 'Molopo' wheat cultivars. Similar results were found by Randolph et al (2005) in evaluating backcross introgressions of *Dn4*. To look at this possibility, a choice test was set up in the summer field 2013 using lines with different combinations of *rag3* and *rag1b* from different genetic backgrounds. The lines were selected from different crosses with PI 567598B (Table 3.1). There was also the possibility of population variation of SBA biotypes present in the field during progeny testing; thus, replication was conducted across three different field cages to account for differences of soybean aphid populations infested in each cage. In each cage, a differential of known resistant sources was included.

Figure 3.6 shows the mean soybean aphid damage index (%) for selected lines from three different populations having alternate alleles within the intervals of *rag1b* and *rag3*. All populations were evaluated in the field 2013 (Fig 3.6A, B, Appendix Fig 4.2B). The third population, 090039, was also previously evaluated in the field 2011 (Appendix Fig 4.2A). Among all the resistant checks, *Rag2* sources were consistently overcome by Michigan aphids,

**Figure 3.6** Mean soybean aphid damage index (%) for selected lines having alternate alleles within the intervals of *rag1b* and *rag3*. **(A)** fine mapping population E00003 x E09902 in the summer field 2013, **(B)** population from IA2070 x E06902 in the summer field 2013. **Solid bar** represent susceptible parents, **open bar** represent resistant parents, **gray** bars represent known resistant sources and **diagonally-hatched** bars represent the selected lines with *rag* gene combinations. Bars with the same letter are not significantly different according to Fisher's Protected LSD ($P = 0.05$).



which could categorize the SBA used in field 2013 as biotype 3. However in 2011 (Appendix Fig 4.2A), Dowling and Jackson had considerably high soybean aphid damage (>60%), which provides evidence that a mix of SBA biotypes 3 and 4 are present in the Michigan fields.

Across genetic backgrounds, the presence of *rag1b* alone was not sufficient to control SBA, since it was not significantly different than the susceptible parent. When *rag1b* was combined with *rag3*, SBA damage was as low as the resistant parent and lines having *rag3* alone. It was observed that different genetic backgrounds may affect resistance activity of *rag3* but not *rag1b* alone. Lines having *rag3* and *rag1b*/*rag3* from the IA2070 background (070063 population) had significantly lower aphid damage index similar to PI 567598B than lines of the same allele combination from E00003 background (090004 population). This finding was important in trying to decide effective ways of pyramiding soybean aphid resistance genes and combining ability into different genetic backgrounds.

*Candidate genes*

The genomic region 6,270,557bp – 6,423,098bp of chromosome 16 has 17 annotated genes (Table 3.6) based on the Williams 82 Glyma1 v1.1 annotation (www.phytozome.net/soybean). Out of the 17 annotated genes, 11 were identified to be candidate genes that may contribute to soybean aphid resistance based on literature search. Of the 11 candidate genes, eight were annotated to encode serine-threonine protein kinase and/or NBS-LRR tandem repeat genes. The eight tandem kinase genes that were annotated were named Glyma16g06940 to Glyma16g07100. These genes span an interval within the recombination bin that contains the molecular markers MSUSNP16-22 (6,342,000 bp) and MSUSNP16-11

**Table 3.6** List of annotated gene models upstream and within the interval of *rag3* (from 6,270,557 bp to 6,423,098 bp) and their functional annotations (Glyma v1.1 gene annotations released 2013).

| Locus Name | Genomic physical position (bp) | Functional annotation and Pfam domains when available |
|---|---|---|
| Candidate genes upstream but outside of *rag3* interval | | |
| Glyma16g06880 | 6202028 - 6203554 | Pfam:00646 F-box domain |
| Glyma16g06890 | 6212705 - 6216258 | Pfam:00646 F-box domain |
| Genes within *rag3* interval | | |
| Glyma16g06940 | 6270027 - 6273436 | Pfam:08263 Leucine rich repeat N-terminal domain<br>Pfam:00560 Leucine Rich Repeat<br>Pfam:00069 Protein kinase domain |
| Glyma16g06950 | 6281613 - 6284165 | Pfam:00069 Protein kinase domain<br>Pfam:00560 Leucine Rich Repeat |
| Glyma16g06965 | 6299244 - 6299587 | None |
| Glyma16g06980 | 6303607 - 6307365 | Pfam:08263 Leucine rich repeat N-terminal domain<br>Pfam:00560 Leucine Rich Repeat<br>Pfam:00069 Protein kinase domain |
| Glyma16g07010 | 6324041 - 6325605 | Pfam:00560 Leucine Rich Repeat<br>Pfam:00069 Protein kinase domain<br>Pfam:07714 Protein tyrosine kinase |
| Glyma16g07021 | 6330803 - 6331982 | Cystinosin |
| Glyma16g07031 | 6333101 - 6334422 | Pfam:00069 Protein kinase domain |
| Glyma16g07041 | 6334599 - 6336986 | Pfam:08263 Leucine rich repeat N-terminal domain<br>Pfam:00560 Leucine Rich Repeat |
| Glyma16g07051 | 6356966 - 6360928 | Pfam:00069 Protein kinase domain<br>Pfam:00560 Leucine Rich Repeat<br>Pfam:08263 Leucine rich repeat N-terminal domain |

**Table 3.6 (cont'd)**

| | | |
|---|---|---|
| Glyma16g07060 | 6374105 - 6377864 | Pfam:00560 Leucine Rich Repeat |
| | | Pfam:08263 Leucine rich repeat N-terminal domain |
| | | Pfam:00069 Protein kinase domain |
| | | Pfam:07714 Protein tyrosine kinase |
| Glyma16g07071 | 6378808 – 6379293 | Pfam:05758 Ycf1 |
| Glyma16g07081 | 6379430 – 6380140 | None |
| Glyma16g07090 | 6385249 – 6387943 | Pfam:04193 PQ loop repeat |
| Glyma16g07100 | 6389306 – 6393052 | Pfam:00560 Leucine Rich Repeat |
| | | Pfam:08263 Leucine rich repeat N-terminal domain |
| | | Pfam:00069 Protein kinase domain |
| Glyma16g07110 | 6407889 - 6412105 | Pfam:04193 PQ loop repeat |
| Glyma16g07125 | 6416720 – 6419217 | None |

(Gm16_6413214_A_G) that can be used for marker-assisted breeding and stacking of *rag3* with other soybean aphid resistance genes.

This is the first report to identify a tightly linked cluster of eight NBS-LRR genes conferring soybean aphid resistance. A study by Liu et al. (2005) have determined that Russian wheat aphid resistance genes *Dn1*, *Dn2*, *DN5*, *DN6* and *DnX* mapped from different sources are tightly linked to the same marker and may be located in the same gene cluster. PI 567543C and PI 567537 SBA resistance QTLs are also mapped in the same region as *rag3* (Zhang et al., 2010; Zhang et al., 2013). PI 567598B was reported to be a durable source of resistance against SBA even with the combination of different biotypes (Cooper, 2012; Alt and Ryan-Mahmutagic, 2013). It is also important to note that two genes outside the *rag3* interval, annotated as CPR30 (F-box domain), can be potential candidate genes or may indirectly be involved in aphid resistance regulation. CPR30 was reported to negatively regulate defense-related genes in Arabidopsis susceptible to the bacterial pathogen, *Pseudomonas syringae* (Gou et al., 2009).

The presence of several or alternative forms of genes within the *rag3* locus that interact with the soybean aphid can explain the durability of resistance in PI 567598B. It is possible that all genes work together to defeat the soybean aphid or that one or different combinations of the R genes can counteract a specific biotype. Molecular cloning and characterization of the different soybean aphid candidate genes will be necessary to understand the molecular basis of aphid resistance in soybean PI 567598B.

# APPENDIX

# APPENDIX

## EVALUATION FOR SOYBEAN APHID RESISTANCE OF SOYBEAN LINES PYRAMIDED WITH RAG1B, RAG1C, RAG3 AND RAG4

The objectives of this experiment include: 1) evaluating for soybean aphid damage of soybean lines containing *rag1b, rag1c, rag3* and *rag4* individually or in different combinations and, 2) to identify genetic interaction of the *rag* gene loci in the greenhouse and field conditions. It is hypothesized that lines with all combinations of *rag* genes will confer the most resistance against soybean aphid, while some *rag* genes will not be as effective when deployed individually.

*Plant materials*

Progenies derived from a cross between E08907 and E09907 were evaluated for the study. E08907 is an advance breeding line derived from a cross of Plant Introduction (PI) 567598B and a susceptible line, Titan RR. E09907 is derived from a cross between PI 567541B and Skylla. Both parental lines were screened to contain the *rag* genes derived from the original PI parents: E08907 contains *rag1b* and *rag3* and E09907 contains *rag1c* and *rag4*.

*Aphid resistance evaluation*

An $F_2$ population of 727 individuals was screened in the greenhouse in the fall of 2010 for soybean aphid damage. A single seed from individual $F_2$ lines were planted in the greenhouse of spring 2011. Each of the 633 $F_3$ plant was individually rated for soybean aphid damage and genotyped for the markers that are closely linked to the mapped *rag* genes. All seeds were

harvested from each $F_3$ plant to form 633 $F_{3:4}$ families that were evaluated for soybean aphid damage in the summer field 2011. Ten seeds for each $F_{3:4}$ family were planted in a 2-ft plot. In the summer of 2013, a total of 190 $F_{3:5}$ lines were selected that contained different combinations of *rag* genes inherited from the parents (based on marker genotypes). These lines were evaluated in the field cage to confirm their phenotype. The different *rag* gene combinations are presented in Table 4.1.

Greenhouse and field conditions were the same as previously described (Zhang et al., 2010; Bales et al., 2013). Soybean aphid damage was rated using the rating scale as used by Mensah et al. (2005). Single $F_2$ and $F_3$ plants were rated in the greenhouse while a damage index (%) was used for the field trials. In all trials, the parental lines were replicated three times. Other known SBA resistance sources were evaluated in the field 2011 and 2013 to compare differences of soybean aphid isolates present in the field.

*Genotyping for selection of rag gene combinations and linkage map analysis*

Tissues were collected from single $F_3$ plants in the greenhouse trial of spring 2011. The CTAB extraction protocol was used to extract the genomic DNA as described by Kisha et al (1997) and concentration was measured using the ND-1000 Spectrophotometer.

The SNP markers used for selection of the $F_3$ lines that contain the *rag* genes of interest were based in fine mapping results of Bales (personal communication, 2013; Chapter 3 results) and Yuan (personal communication, 2013). Custom KASP assays were designed for each SNP

marker. PCR reactions and fluorescent acquisition were ran as previously described by Bales et al (Chapter 3, methods).

To determine whether the known QTLs explain significant phenotypic variation in the $F_3$ population, linkage map analysis was performed as previously described by Bales et al. (2013; Chapter 2, methods). Multiple interval mapping method was performed to determine significant QTL x QTL interactions.

**Results and Discussion**

This study investigates the relationship of genes controlling antibiosis resistance against soybean aphids in PI 56798B and PI 567541B. Mensah et al. (2008) reported two recessive genes controlling SBA resistance for each plant introduction. The two genes in PI 56741B were mapped on chromosomes 7 and 13, named *rag1c* and *rag4*, respectively (Zhang et al. 2009). The two genes in PI 567598B were mapped on chromosomes 7 (*rag1b*) and 16 (*rag3*) (Bales et al. 2013). Both plant introductions have a QTL common in chromosome 7 -- *rag1b* and *rag1c*. Zhang et al. (2009) reported that *rag1c* conferred resistance in all trials while Bales et al. (2013) reported that *rag1b* was defeated in the field. It is not yet clear whether *rag1b* and *rag1c* are the same genes or maybe allelic, but they show different response to soybean aphid infestation and in combination with the other gene mapped in the same source.

Figure 4.1 shows the mean soybean aphid damage index (%) for 190 selected lines having different combination of alleles within the intervals of *rag1*, *rag3* and *rag4*. All $F_3$-derived lines were evaluated in the greenhouse of spring 2011, summer field 2011 and field 2013. The presence of *rag1b* alone is not sufficient to control SBA, but lower SBA damage is observed

when in combination with *rag3*. Lines containing *rag1c* alone have lower SBA damage than *rag1b* alone (Fig 4.1); however, significant SBA damage is found in the field 2011 (Fig 4.2A).

Among gene combinations, any other gene combined with *rag3* and *rag1c* is effective in controlling SBA while *rag4+rag1b* combination is the least effective of all combinations. This finding is supported by previous QTL mapping results done on PI 567541B and PI 56598B. Zhang et al (2009) have reported that *rag4* is not as effected than *rag1c* locus in the PI 567541B source and explains a lower effect on the phenotypic variation. Bales et al (2013) found that *rag1b* is overcome by SBA in the field.

To understand genetic interactions of the different *rag* genes in the population, a multiple interval mapping analysis was conducted. Only the *rag1c* allele from the PI 567541B and *rag3* from the PI 567598B were found to significantly explain the highest percentage of the phenotypic variation in the population (Table 4.2). The *rag1c* and *rag3* loci detected are mapped at the same location as previously reported (Fig 4.3), thus confirming that the markers used for selection were tightly linked to the loci. A significant additive x additive interaction (2-24%) between the two loci also contributes to controlling SBA resistance. In total, *rag1c*, *rag3* loci and their interaction explains 33-59% of the phenotypic variation. This implies that selecting for *rag1c* and *rag3* loci in pyramiding *rag* genes is enough to provide resistance against SBA. This also validates the findings that *rag4* and *rag1b* does not significantly contribute to SBA resistance of PI 567541B and PI 56598B, respectively. *Rag4*, however, can still be used for SBA control but needs to be combined with either *rag1c* or *rag3*.

It has been reported that different soybean aphid biotypes exist in the fields of Michigan based on the feeding behavior of SBA on Dowling (Mensah, 2007). In the field trials of summer 2011 and 2013, *Rag2* sources are consistently overcome by Michigan aphids, which categorizes

the SBA used in field 2013 as biotype 3 (Fig 4.2 A, B). In the field 2011 (Fig 4.2A), Dowling and Jackson have considerably high soybean aphid damage (>60%), which provides evidence that a mix of SBA biotypes 3 and 4 are present in the Michigan fields.

The findings in this study will provide useful information for breeders to develop varieties for sustainable management of soybean aphids and the different biotypes present in the field. Gene pyramiding is an important approach to pest management, but careful evaluation for efficacy must be done in the selection for gene loci to combine.

**Table 4.1** Genotypic groups of 190 F3-derived lines from the mapping population E08907 x E09907 with different combinations of *rag* genes.

| *rag* genotypes * | *n* |
|---|---|
| *rag1b* | 5 |
| *rag1c* | 7 |
| *rag1b/c* | 12 |
| *rag3+rag1b* | 87 |
| *rag3+rag1c* | 7 |
| *rag3+rag1b/c* | 15 |
| *rag4+rag1b* | 3 |
| *rag4+rag1c* | 8 |
| *rag4+rag1b/c* | 17 |
| *rag3+rag4+rag1b* | 7 |
| *rag3+rag4+rag1c* | 10 |
| *rag3+rag4+rag1b/c* | 12 |
| Total lines evaluated | 190 |

*All *rag* alleles are homozygous except for *rag1b/c* which denotes for heterozygote alleles from *rag1b* and *rag1c*.

**Table 4.2** Summary of QTLs for soybean aphid resistance detected in the 633 $F_3$-derived lines from E08907 x E09907 population using the multiple interval mapping method

| Trials | Chr/*rag*[a] | Peak pos.[b] | Flanking markers[c] | Genetic effect | | |
|---|---|---|---|---|---|---|
| | | | | LOD[d] | $R^{2e}$ | a[f] |
| Greenhouse 2011 | | | | | | |
| 3-week rating | 07/*rag1c* | 3.7 | SNP07-5rhL2 ~ SNP07-7rhR2 | 27.1 | 14.9 | 8.3 |
| | 16/*rag3* | 3.2 | SNP16-18 ~ SNP16-22 | 46.9 | 20.9 | -11.4 |
| | Additive x additive interaction (*rag1c* x *rag3*) | | | 46.7 | 23.9 | -15.1 |
| | | | Total | | 59.7 | |
| Field 2011 | | | | | | |
| 3-week rating | 07/*rag1c* | 1.71 | SNP07-05rhL2 ~ SNP07-07rhR2 | 16.6 | 4.3 | 11.3 |
| | 16/*rag3* | 3.22 | SNP16-18 ~ SNP16-22 | 40.5 | 26.7 | -18.8 |
| | Additive x additive interaction (*rag1c* x *rag3*) | | | 2.6 | 2.6 | -6.4 |
| | | | Total | | 33.6 | |

[a]Chromosome/*rag* gene. Chromosome numbers according to the SoyBase (Grant et al. 2010) and *rag* gene designation according to Zhang et al. (2010) and Bales et al. (2013)

[b]QTL peak position is expressed in cM

[c]Markers flanking the peak position

[d]LOD thresholds are 2.12 and 2.08 for greenhouse 2011 and field 2011, respectively

[e]$R^2$, percentage of phenotypic variation explained by a QTL

[f]Additive effect. The negative value implies that the PI 567598B (*rag3* and *rag1b*) allele contributes to the phenotypic value. The positive value implies that the PI 567541B (*rag4* and *rag1c*) allele contributes to the phenotypic value.

**Figure 4.1** Mean soybean aphid damage index (%) of 190 selected lines from E08907 x E09907 population with corresponding *rag* gene combinations across different trials. Combinations (regardless of trials) with the same letter are not significantly different according to Fisher's Protected LSD ($P = 0.05$)

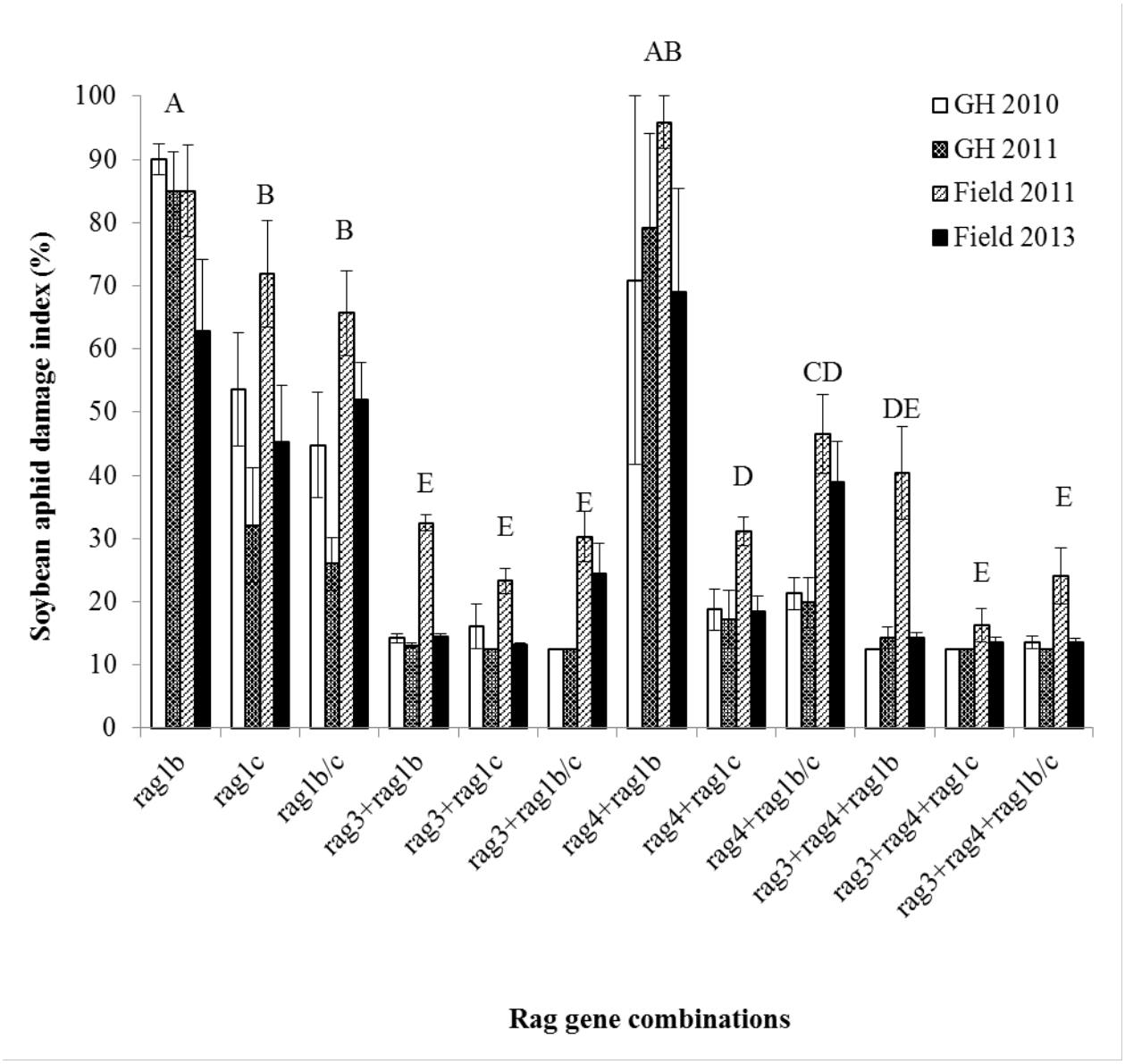


**Figure 4.2** Mean soybean aphid damage index (%) for 190 selected lines having alternate alleles within the intervals of *rag1b* and *rag3*. **(A)** Lines from the cross E08907 x E09907 evaluated in summer field 2011 and **(B)** summer field 2013. **Solid bar** represent susceptible check, **gray** bars represent known resistant sources and **diagonally-hatched** bars represent the selected lines with *rag* gene combinations.

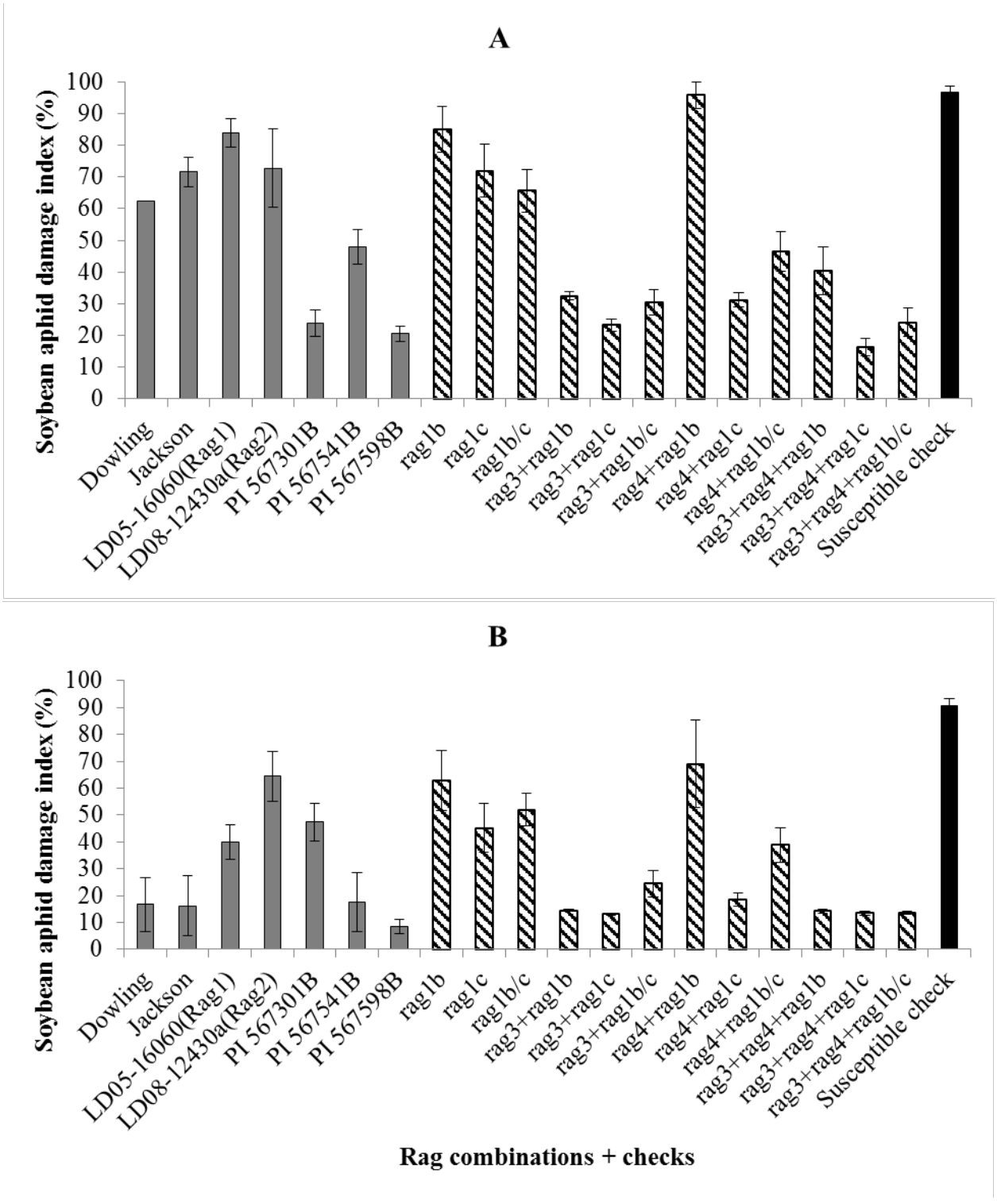
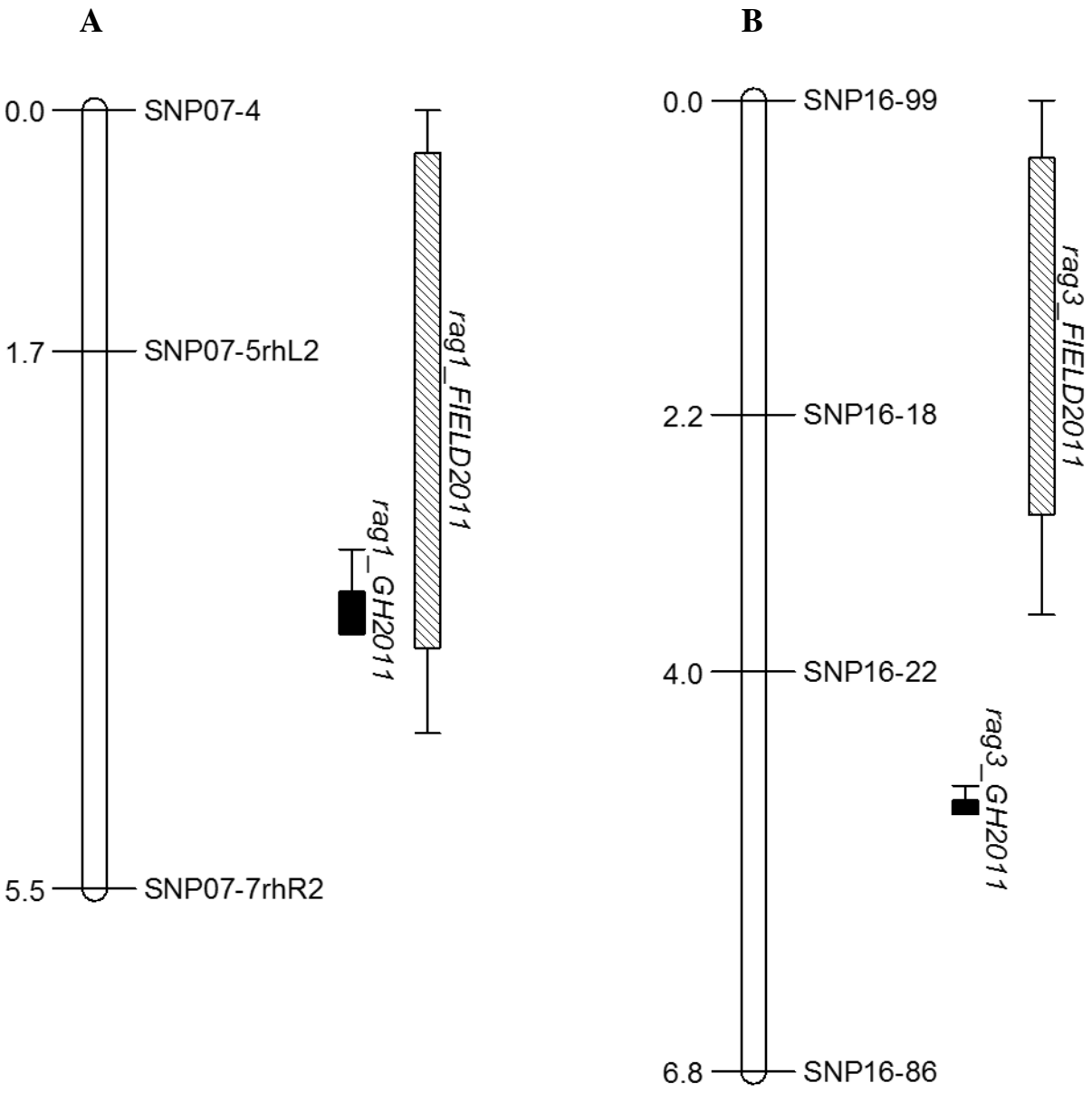


**Figure 4. 3** QTL locations of *rag* genes using multiple interval mapping analysis of 633 $F_3$-derived lines from E08907 x E09907 population evaluated in the greenhouse 2011 (GH2011) and field 2011 (FIELD2011). **A** Linkage map and *rag1* on chromosome 7, **B** linkage map and *rag3* detected on chromosome 16. **Solid bar** represent QTL detected in the greenhouse 2011. **Diagonally-hatched bar** represent QTL detected in the field 2011.



# REFERENCES

## REFERENCES


Alt, J. and Ryan-Mahmutagic, M. (2013). Soybean aphid biotype 4 identified. Crop Science. **53**:1491.

Applied Biosystems by Life Technologies. Available at: http://www.appliedbiosystems.com/.

Bales, C., Zhang, G., Liu, M., Mensah, C., Gu, C., Song, Q., Hyten, D., Cregan, P. and Wang, D. (2013). Mapping soybean aphid resistance genes in PI 567598B. Theoretical and Applied Genetics.1–11.

Chen, J.-Q., Martin B, Rahbe Y and Fereres A (1997). Early intracellular punctures by two aphid species on near-isogenic melon lines with and without the virus aphid transmission (*Vat*) resistance gene. European Journal of Plant Pathology. **103**:521–536.

Cooper, S.G. (2012). Soybean aphid biotypes: understanding geographic distribution (2008–2012). In Conference presentation for Soybean Breeders' Workshop, St. Louis, MO, USA: Monsanto Company.

Cooper, W.C., Jia, L. and Goggin, F.L. (2004). Acquired and R-gene-mediated resistance against the potato aphid in tomato. J Chem Ecol. **30**:2527–2542.

Der Westhuizen, A.J. Van, Qian, X.-M. and Botha, A.-M. (1998). β -1,3-Glucanases in wheat and resistance to the Russian wheat aphid. Physiologia Plantarum. **103**:125–131.

Hartman, G.L., Domier, L.L., Wax, L.M., Helm, C.G. and Onstad, D.W. (2001). Occurrence and distribution of *Aphis glycines* on soybeans in Illinois in 2000 and its potential control. Available at: http://www.plantmanagementnetwork.org/pub/php/brief/aphisglycines/.

Hill, J.H., Alleman, R., Hogg, D.B. and Grau, C.R. (2001). First report of transmission of soybean mosaic virus and alfalfa mosaic virus by *Aphis glycines* in the new world. Plant Disease. **85**:561–561

Kaloshian, I., Kinsey, M.G., Williamson, V.M. and Ullman, D.E. (2000). Mi-mediated resistance against the potato aphid *Macrosiphum euphorbiae* (Hemiptera: *Aphididae*) limits sieve element ingestion. Environmental Entomology. **29**:690–695.

Kamphuis, L.G., Lichtenzveig, J., Peng, K., Guo, S.-M., Klingler, J.P., Siddique, K.H.M., Gao, L.-L. and Singh, K.B. (2013). Characterization and genetic dissection of resistance to spotted alfalfa aphid (*Therioaphis trifolii*) in *Medicago truncatula*. J. Exp. Bot. **64**: 5157-72.

Kim, K.-S., Bellendir, S., Hudson, K.A., Hill, C.B., Hartman, G.L., Hyten, D.L., Hudson, M.E. and Diers, B.W. (2009). Fine mapping the soybean aphid resistance gene *Rag1* in soybean. Theor Appl Genet. **120**:1063–1071.

Kim, K.-S., Hill, C.B., Hartman, G.L., Hyten, D.L., Hudson, M.E. and Diers, B.W. (2010). Fine mapping of the soybean aphid-resistance gene *Rag2* in soybean PI 200538. Theor Appl Genet. **121**:599–610.

Kisha, T.J., Sneller, C.H. and Diers, B.W. (1997). Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci. **37**:1317–1325.

Klingler, J., Creasy, R., Gao, L., Nair, R.M., Calix, A.S., Jacob, H.S., Edwards, O.R. and Singh, K.B. (2005). Aphid resistance in *Medicago truncatula* involves antixenosis and phloem-specific, inducible antibiosis, and maps to a single locus flanked by NBS-LRR resistance gene analogs. Plant Physiol. **137**:1445–1455.

Lam, H.-M. et al. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat Genet. **42**:1053–1059.

Li, Y., Zou, J., Li, M., Bilgin, D.D., Vodkin, L.O., Hartman, G.L. and Clough, S.J. (2008). Soybean defense responses to the soybean aphid. New Phytol. **179**:185–195.

Liu, X.M., Smith, C.M., Friebe, B.R. and Gill, B.S. (2005). Molecular mapping and allelic relationships of russian wheat aphid–resistance genes. Crop Science. **45**:2273.

Mensah, C., DiFonzo, C., Nelson, R.L. and Wang, D. (2005). Resistance to soybean aphid in early maturing soybean germplasm. Crop Science. **45**:2228.

Mensah, C., DiFonzo, C. and Wang, D. (2008). Inheritance of soybean aphid resistance in PI 567541B and PI 567598B. Crop Science. **48**:1759.

R Development Core Team (2011). R: A language and environment for statistical computing, R foundation for statistical computing, Vienna, Austria.

Randolph, T.L., Peairs, F.B., Koch, M., Walker, C.B., Stubbs, J.R., Quick, J.S. and Haley, S.D. (2005). Yield response and categories of resistance to russian wheat aphid in four *Dn4* hard red winter wheat cultivars. Journal of Economic Entomology. **98**:588–594.

Roche Applied Science Roche Applied Science: LightCycler® 480 System. Available at: https://www.roche-applied-science.com.

Rossi, M., Goggin, F.L., Milligan, S.B., Kaloshian, I., Ullman, D.E. and Williamson, V.M. (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proceedings of the National Academy of Sciences. **95**:9750 –9754.

Schmutz, J. et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature. **463**:178–183.

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

Villada, E.S., González, E.G., López-Sesé, A.I., Castiel, A.F. and Gómez-Guillamón, M.L. (2009). Hypersensitive response to *Aphis gossypii* Glover in melon genotypes carrying the *Vat* gene. Journal of Experimental Botany. **60**:3269 –3277.

Wang, D., Diers, B.W. and Boyse, J. (2006). Registration of "Skylla" Soybean. Crop Science. **46**:974.

www.soystats.com (2012). Soy Stats 2012. Available at: http://www.soystats.com/2012

Xiao, L., Hu, Y., Wang, B. and Wu, T. (2013). Genetic mapping of a novel gene for soybean aphid resistance in soybean (Glycine max [L.] Merr.) line P203 from China. Theor. Appl. Genet. **126**:2279-87.

Zhang, G., Gu, C. and Wang, D. (2010). A novel locus for soybean aphid resistance. Theor Appl Genet. **120**:1183–1191.

Zhang, G., Gu, C. and Wang, D. (2013). Mapping and validation of a gene for soybean aphid resistance in PI 567537. Mol. Breeding **32**: 131–138.