GENETIC DISSECTION OF APHID RESISTANCE IN WILD SOYBEAN GLYCINE SOJA ACCESSION 85-32

By

Shichen Zhang

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

2017

ABSTRACT

GENETIC DISSECTION OF APHID RESISTANCE IN WILD SOYBEAN GLYCINE SOJA ACCESSION 85-32

By

Shichen Zhang

The soybean aphid, an invasive species, has been posing a substantial threat on soybean production in North America since its first discovery in 2000. Two novel aphid-resistance quantitative loci (QTLs) were previously revealed as controlling aphid resistance in the wild soybean, *Glycine soja* 85-32. Therefore, the first objective was to validate these QTLs under different genetic backgrounds. Using single nucleotide polymorphism (SNP) markers, discovered from whole genome resequencing data or mined from the SoySNP50K iSelect BeadChip, two aphid-resistance QTLs were successfully validated and designated *Rag6* and *Rag3c,* respectively. The second objective was to fine map these two loci and identify structural variants within the candidate genes. *Rag6* was refined to a 49-kb interval with four candidate genes, including three clustered nucleotide-binding site leucine-rich repeat (NBS-LRR) genes and an amine oxidase encoding gene. *Rag3c* was refined to a 150-kb interval with eleven candidate genes, two of which are a LRR gene and a lipase gene. By aligning the sequencing reads from the whole genome exome-capture of the resistant source to the soybean reference genome (aphid-susceptible), structural variants (including frame shifts, deletions and non-synonymous coding changes) were identified within the candidate genes of *Rag6* and *Rag3c*, and new SNPs and insertion/deletions were discovered in the exon regions. The variability and dynamics of aphid population limits the effectiveness of host-resistance gene(s). Therefore, the third objective was to develop and evaluate soybean advanced breeding lines integrated with different aphid-resistance genes. Based on the responses from the indicator lines, Biotype 3 was

determined as a major component of aphid populations collected in Michigan during 2015 - 2016. The different performance of *Rag*-'Jackson' and *Rag1*-'Dowling' along with the break-down of resistance in plant introductions (PIs) 567301B and 567324 may be explained by Biotype 3 or an unknown virulent biotype establishing in Michigan. Lines with *rag1c*, *Rag3d*, *Rag6*, *Rag3c+Rag6*, *rag1b+rag3*, *rag1c+rag4*, *rag1c+rag3+rag4*, *rag1c+Rag2+rag3+rag4* and *rag1b+rag1c+rag3+rag4* demonstrated strong and consistent resistance in five trials across 2015 - 2017. Due to the variability of virulent aphid populations, different combinations of Rag genes may perform differently across geographies. However, advanced breeding lines pyramided with three or four Rag genes will likely provide broader and more durable resistance to diverse and dynamic aphid populations across many geographic regions.

This dissertation is dedicated to my father (Qinggen Zhang), mother (Meirong Cheng), sister (Shihou Zhang) and my fiancé (Sean F. Biehn)

ACKNOWLEDGEMENTS

Throughout my journey, up to this point, I have considered myself blessed with many wonderful teachers, mentors and friends. I am grateful for the open-minded atmosphere I had during my undergraduate years, where I had opportunities to join different laboratories and explore the agricultural sciences. My undergraduate research apprenticeship with Dr. Zeng and Dr. Zhou was invaluable. They gave me lots of training in agronomy and plant pathology. A special thanks to Dr. Chen for leading me to the fantastic realm of plant breeding and genetics.

I am thankful to be a part of the PBGB program at MSU, where I met a lot of great professors and fellow students. It is a strong community where I gained knowledge and skills from different areas and witnessed collaborations among scientists. Thank you Drs Russell Freed, Dave Douches, Ning Jiang, Eric Olson, Cholani Weebadde, Tylor Johnston, Robin Buell, Robert Last and Sheng-Yang He for your guidance and training. I greatly appreciated all the discussions, advice and support from my committee members, Drs. Amy Iezzoni, Chris DiFonzo and James Kelly. My sincere gratitude to my advisor, Dr. Dechun Wang, for being the best mentor I could hope for. Thank you for training me to be independent and a critical thinker in research, and giving me opportunities to explore and implement my research ideas. I am also grateful for having my wonderful lab mates, Kate Zhang, Carmille Bales, Desmi Chandrasena, Zixiang Wen, Ruijuan Tan, Paul Collins, Feng Lin, John Boyse, Randy Laurenz and Cherry Gu, who all helped and supported my research projects.

My appreciation also goes to PBGB, Graduate School, PSM and CANR for their financial support of my Ph.D. fellowship; and to United Soybean Board and Michigan Soybean Promotion

Committee for my research funding and assistantship. I would also like to extend my appreciation to the PSM staff and to my colleagues from other labs. Thank you for providing your generous help when needed.

I would also like to acknowledge my friends, who have made my Ph.D. study a great adventure. A special thanks to Tiffany Liu for being studious with me, discussing scientific and philosophical questions about life, and pushing each other through tough times. The most important thing I learned from Tiffany was that the meaning of life is not only about enriching your own life but others as well. I would also like to thank Qianwei Jiang for always being there, supporting me, and giving me great advice.

Last, but not least, I am grateful for all the love, support and tremendous advice from my family and the Biehn family. I am fortunate to have parents and a sister, who continue to support and remind me to pursue my dreams. I am also thankful for my grandparents, uncles, aunts and cousins for their love and encouragement. A special thanks to Sean Biehn, for your unconditional love and support, helping me to get through all the challenging times.

Thank you again, to everyone, for all your support along my journey. I cannot wait to explore new adventures in science and do my best to enrich people's lives.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

LITERATURE REVIEW

Part of the work presented in this chapter has been accepted in the Book Chapter:

Advances in pest-resistant varieties of soybean.
Shichen Zhang, Dechun Wang (2017).
In H.T. Nguyen (Ed.)
Achieving sustainable cultivation of soybeans.
Sawston, Cambridge, UK: Burleigh Dodds Science Publishing

Soybean and its economic importance

The cultivated soybean, *Glycine max* (L.) Merr., is a legume species native to East Asia, and it has been believed domesticated from the wild ancestor, *Glycine soja* (Singh 2006). Both *G. max* and *G. soja* are annual dicots and belong to the subgenus Soja of the family Leguminosae (Singh 2006). The soybean genome had undergone two rounds of whole genome duplication and a process of diploidization; thus, soybean has been considered as a paleopolyploidy (2n=40) (Shoemaker et al. 1996; Roulin et al. 2013).

Soybean was first introduced to North America by a sailor, Samuel Bowen, in 1765 (Hymowitz et al. 1983). Later, soybean has become one of the major crops in the United States as it has multiple uses including animal feed, cooking oil, biofuel, human protein source, etc. In 2016, 83.4 million acres were planted with soybeans in the U.S., ranking the first in world soybean production (4.31 billion bushels) with a total value of $40.94 billion, and 2.03 billion bushels were exported (SoyStats 2016).

The soybean aphid and its impacts

The soybean aphid, *Aphis glycines* Matsumura, is an invasive species that originated from China (Wu et al. 2004). Since its first discovery in southern Wisconsin during the summer of 2000 (Alleman et al. 2002), the soybean aphid has aggressively spread to all the major soybean production area in the United States and Canada, and become an economically important pest of soybean (Ragsdale et al. 2011).

The soybean aphid is recognized by its black cornicles and pale cauda. It has a heteroecious life cycle with different physical forms and sexual stages (Wu et al. 2004). Female and male aphids

mate in the fall and produce eggs to overwinter on the primary host, buckthorn (*Frangula alnus*). The eggs hatch and develop into wingless fundatrices in the early spring. Later near the blooming stage of buckthorn, these fundatrices asexually reproduce winged alatae that migrate to the secondary host, soybean, during the spring (Wu et al. 2004). After the migration to soybean plants, aphid population build rapidly because of the parthenogenesis and the deformed paedogenesis (Zhang 1988). Due to the rapid unsexual propagation, there are about fifteen generations living on soybean plants during the summer (Wu et al. 2004). They thrive best with temperatures from 22 to 27 ºC during June to August in Michigan. Soybean aphid population densities usually peak during soybean growth stage R3 (the beginning of pod formation) to R5 (the full-size pod) (Ragsdale et al. 2007).

The aphid's stylet feeding removes nutrients and water from soybean plants, resulting in leaf curling, plant wilting and plant death under heavy infestations (Wu et al. 2004); the soybean yield loss caused by aphids' direct feeding was estimated up to 40% (Ragsdale et al. 2007). Additionally, soybean aphids cause secondary yield loss through transmitting viruses (e.g., *Soybean dwarf virus*, *Soybean mosaic virus*, *Potato virus*, *Alfalfa mosaic virus,* and *Tobacco ring spot virus*), which impair soybean growth and yield by causing plant stunting, leaf deformation and reduced pod filling (Hill et al. 2001; Clark &Perry. 2002; Davis et al. 2005). Furthermore, aphids consistently produce honeydew that can cause the growth of sooty mold; excessive sooty mold block soybean plant photosynthesis, leading to additional yield losses (Malumphy, 1997; Lemos Filho and Paiva, 2006). Overall, the economic loss caused by soybean aphids was estimated as $3.6 to $4.9 billion annually in North America (Kim et al. 2008).

Soybean aphid management

During the growing season of 2003, over 42 million acres of soybean in the north-central U.S. suffered from an outbreak of soybean aphids (Ragsdale et al. 2004). Since then, significant efforts have been made to find solutions to combat soybean aphids; different tactics including chemical, biological and cultural control of aphids have been recommended for protecting soybean yield. Among these tactics, insecticide spray has gained the most popularity in controlling soybean aphids, especially during high outbreak situations. The most commonly applied insecticides in controlling aphids include pyrethroids, neonicotinoids and organophosphates; they are available in the forms of seed treatments or foliar sprays (DiFonzo 2005; Ohnesorg et al. 2009; McCarville and O'Neal 2013). The economic threshold for controlling aphids with insecticides was estimated as 273 aphids per plant to provide a lead-time of seven days before aphid populations reach the economic injury level of 674 aphids per plant (Ragsdale et al. 2007).

Due to the rapid establishment of soybean aphids in the U.S., insecticide application on soybean increased from 0.03 million pounds in 2000 to 4.7 million pounds in 2008 (Fernandez-Cornejo et al. 2014). Although insecticide application is effective in protecting soybean yield, this control method increases agricultural input, causes environmental contamination and jeopardizes beneficial insects such as natural enemies and pollinators (Ohnesorg et al. 2009; Lundin et al. 2015). Sometimes, the extensive use of insecticides lead to the appearance of pesticide-resistant insect populations, resurgence of primary pests and secondary pests.

Natural enemies such as predators and parasitoids were identified as biological control agents to protect soybean from aphids. Common predators of soybean aphids include the multicolored

Asian lady beetle (*Harmonia axyridis*), insidious flower bug (*Orius insidiosus*), the lacewing larvae (*Chrysopidae sp.*) and damsel bugs (*Nabidae sp.*) (Rutledge et al. 2004). The most common parasitoid of soybean aphids is the parasitic wasp, *Aphelinus albipodus*, that lays eggs inside soybean aphids, leading to the development of aphid 'mummies' (Ragsdale et al. 2011). However, the environmental conditions such as humidity and temperature greatly influence the population size of these natural enemies, leading to inconsistent control of soybean aphids.

Host-plant resistance

An alternative way of controlling soybean aphids is to employ natural host-plant resistance (HPR) that can provide soybean with economically and environmentally-friendly season-long protection. Thus, HPR is an important component of integrated pest management (IPM), which aims at limiting the usage of pesticides to protect the ecosystem. HPR coupled with biological control is of most interest because they are favorable and compatible in IPM; natural enemies can effectively keep aphid populations in check when the population size is under control by HPR in the early season (Hodgson et al. 2012).

There are three categories of HPR, including antibiosis, antixenosis and tolerance (Painter 1958). Antibiosis inhibits the insect biological and/or reproductive process through toxic plant secondary metabolites. Antixenosis morphologically or biochemically deters pests based on a "non-preference" behavior. Tolerance refers to the ability of the plant to maintain its yield under a moderate amount of damage from the pest. As for host-plant resistance against soybean aphids, only antibiosis and antixenosis have been discovered to date and studied in soybean germplasm.

Host-plant resistance against soybean aphids have been identified and applied in developing aphid-resistant soybean cultivars (McCarville, 2012). However, one major drawback to employing HPR is the potential lack of its durability, which is usually caused by the emerging virulent aphid biotypes (Kim et al. 2008; Hill et al. 2010; Alt and Ryan-Mahmutagic 2013). In addition to seeking new resistance sources, several methods of sustaining HPR have been practiced, including resistance gene pyramiding, variety mixture and resistance gene deployment based on biotype distribution.

Biochemical mechanisms of host-plant resistance against aphids

Allelochemicals are secondary plant metabolites that are not essential for plant growth and development; allelochemicals with negative allelopathic effects are known as important for plant defense against herbivory (Stamp 2003). Phytoalexins are antibiotic metabolites that are produced or enriched under biotic stress (Hart et al. 1983). Some flavonoids are important phytoalexins with anti-herbivory effects in *Glycine* spp. (Burden and Norris 1992). For example, the inducible resistance against Mexican bean beetles (*Epilachna varivestis*) in PI 227687 was due to increased phenylpropanoid metabolism (total phenolic content) when this accession was challenged by Mexican bean beetles (Chiang et al. 1987). A similar result was reported by Hart et al. (1983) that glyceolin (a type of flavonoid) was functioning as a deterrent to Mexican bean beetles and some other Coleopterans. Coumestrol, another isoflavonoid, was suggested as contributing to antixenosis resistance against Mexican bean beetles in the cultivar "Davis" (Caviness and Walters 1966; Burdern and Norris 1992). Genistein and daidzin can cause greater insect mortality, lower initial larval and pupal weight, reduced growth and elongated larval cycle to southern green stink bugs (*Nezara viridula*), cabbage loopers (*Trichoplusia ni*), and

velvetbean caterpillars (*Anticarsia gemmatalis*) (Sharma and Norris 1991; Hoffmann-Campo et al. 2001; Piubelli et al. 2003, 2005). A Chinese soybean cultivar, 'Zhongdou 27', has a high isoflavone concentration that help protect soybean from soybean aphid attacks as the major aphid-resistance QTLs found in 'Zhongdou 27' were highly associated with high isoflavone content (Meng et al. 2011).

Proteinase inhibitors are known as being involved in plant defense to herbivory injury; they are peptides or proteins that inhibit activities of digestive enzymes in the insect gut, thus adversely affecting protein digestion and impeding the growth of insects (Ryan 1990). Broadway and Duffey (1986) reported both soybean trypsin inhibitor and potato proteinase inhibitor significantly reduced the growth of larval beet armyworm (*Spodoptera exigua*), and larval corn earworm (*Helicoverpa zea*). The soybean trypsin-chymotrypsin inhibitors induced significant mortality and growth inhibition of the pea aphid (*Acyrthosiphon pisum*) and potato aphid (*Macrosiphum euphorbiae*) (Rahbé et al. 2003; Azzouz et al. 2005). Potato proteinase inhibitors I and II increased mortality among late instar aphids and reduced production of nymphs in feeding trials of cereal aphid species (*Diuraphis noxia, Schizaphis graminum, Rhopalosiphon padi*) (Tran et al. 1997). The insecticidal effects of protein inhibitors were demonstrated in several transgenic plants, enhancing host-plant resistance against lepidopteran and coleopteran pests (Hilder et al. 1987; Johnson et al. 1989; De Leo et al. 2001; Falco and Silva-Filho 2003; Lecardonnel et al. 1999; Alfonso-Rubí et al. 2003).

Identification of aphid-resistance genes in soybean germplasm

Li et al. (2004) first reported three soybean genotypes, 'Dowling' (PI 548663), 'Jackson' (PI 548657) and PI 200538, confer antibiosis resistance against the Illinois soybean aphid isolate

(Hartman et al. 2001). Antibiosis resistance from 'Dowling' and 'Jackson' were determined to be controlled by a single dominant gene (*Rag1*/*Rag*) that was later mapped to the same location between markers Satt435 and Satt463 on LG M/Chr. 7 (Hill et al. 2006a, b; Li et al. 2007). A subsequent genetic allelism test was conducted among 1,000 F_2 plants from 'Dowling' × 'Jackson' and no susceptible plant was observed, suggesting the genes were allelic (Hill et al. 2012). Later, with additional single nucleotide polymorphism (SNP) markers developed, *Rag1* was refined to a 115-kb interval and two nucleotide-binding site leucine-rich repeat (NBS-LRR) genes were proposed as the candidates for *Rag1* (Kim et al. 2010a) based on the Williams 82 reference genome annotation on SoyBase (Grant et al. 2010). No yield drag was observed with *Rag1* (Kim and Diers 2009), and cultivars with *Rag1* were commercially released to growers (McCarville, 2012). However, the resistance conferred by *Rag1*/*Rag* to the Illinois soybean aphid isolate was reported overcome by an Ohio isolate in 2006 (Kim et al. 2008). Thus, the Illinois isolate was referred to as Biotype 1 and the Ohio isolate was referred to as Biotype 2 (Kim et al. 2008).

Mian et al. (2008a, b) discovered three PIs (243540, 567301B and 567324) exhibiting resistance against Biotypes 1 and 2, and mapped a single dominant gene, *Rag2*, controlling antibiosis resistance in PI 243540 to LG F/Chr.13 between markers Satt334 and Sct_033. In the biotype study by Kim et al. (2008), PI 200538 remained strong antibiosis resistance to Biotype 2; later, a mapping study by Hill et al. (2009) revealed the underlying gene is the same gene as *Rag2* (Mian et al. 2008b) since it resides at the same genomic location and confers identical resistance reactions to different biotypes. Fine mapping of *Rag2* in PI 200538 delimited it to a 54-kb region with one NBS-LRR gene as the candidate (Kim et al. 2010b). Fox et al. (2014) found that *Rag2*

was significantly associated with resistance in 20 of the 21 F_2 populations derived from 21 newly identified aphid-resistant PIs, suggesting *Rag2* may be a major aphid resistance source in the USDA soybean germplasm collection. A different QTL conferring antixenosis resistance in PI 567301B was mapped near *Rag2* on LG F/Chr. 13. This QTL referred to a different gene (*Rag5*) in that detached leaves of PI 567301B did not maintain resistance towards aphids while PI 243540 (source of *Rag2*) did (Michel et al. 2010; Jun et al. 2012). Additionally, a minor QTL providing antixenosis resistance on LG A2/Chr. 8 near marker BARC-063283-18296 was identified in PI 567301B (Jun et al. 2012). Jun et al. (2013) discovered three aphid-resistance QTLs in PI 567324; two major ones (QTL_13_1 and QTL_13_2) were located close to the previously reported *Rag2* locus (Mian et al. 2008a; Kim et al. 2010b) and *rag4* locus (Zhang et al. 2009) respectively, and a minor one (QTL_6_1) was detected on chromosome 6, where no aphid-resistant gene has been previously reported. The oligogenic resistance from PI 567324 is expected to provide broader and more durable resistance against aphids compared to cultivars with monogenic resistance (Jun et al. 2013).

Hill et al. (2010) reported the discovery of Biotype 3 in Indiana, which readily colonized plants with *Rag2*. Mensah et al. (2005) screened 2,147 soybean accessions from MG 0 to III and identified four MG III accessions (PI 567543C, PI 567597C, PI 567541B and PI 567598B) resistant to mixed aphid biotypes in Michigan. PI 567541B and PI 567598B possess antibiosis resistance while PI 567543C and PI 567597C possess antixenosis resistance. These four PIs were included in the biotype studies by Kim et al. (2008) and Hill et al. (2010), and found resistant to multiple biotypes. Mensah et al. (2008) reported that aphid-resistance in PI 567541B and PI 567598B were both controlled by two major recessive genes. In PI 567541B, one recessive gene

was mapped to the *Rag1* region and was named *rag1c* whereas the other recessive gene was mapped to a different region (Satt649-Satt348) from *Rag2* on LG F/Chr.13 and was named *rag4* (Zhang et al. 2009). The broad antixenosis resistance provided by PI 567543C and PI 567597C were reported controlled by a single partially dominant gene, *Rag3* and *Rag3e* (mapped in close proximity on LG J/Chr.16), respectively (Zhang et al. 2010; Du 2016). Bales et al. (2013) reported the antibiosis resistance in PI 567598B were contributed by two recessive genes (*rag1b* and *rag3*) that were mapped to the previously identified *Rag1* (Li et al. 2007; Kim et al. 2010a) and *Rag3* (Zhang et al. 2010) regions.

Interestingly, in addition to *Rag3*, *Rag3e* and *rag3*, more aphid-resistance genes were mapped to this shared genomic region. Zhang et al. (2013) detected a single dominant gene, *Rag3b,* conferring antibiosis resistance against mixed biotypes of aphids in PI 567537. Du (2016) fine mapped a partially dominant gene, *Rag3d* (Liu 2010), in PI 567585A, which confers antibiosis resistance against multiple biotypes. According to the Williams 82 reference genome annotation on SoyBase (Grant et al. 2010), this shared region is enriched with NBS-LRR genes. Fine mapping studies have been conducted and delimiting these Rag genes (*Rag3*, *Rag3b*, *Rag3d*, *Rag3e* and *rag3*) to similar or different genomic locations (Bales 2013; Du 2016; unpublished data); it is possible that some of these Rag genes are different NBS-LRR genes while others are different alleles of the same NBS-LRR gene(s).

Isoflavone may play a role in plant defense against soybean aphids. Meng et al. (2011) mapped two aphid-resistance QTLs to the same locations of QTLs previously reported as associated with high isoflavone content in a Chinese soybean cultivar, ‘Zhongdou 27’. One of the QTLs (qRa_1)

was mapped to Satt470 on LG A2/Chr. 8 and explained a large portion of phenotypic variances. The second QTL (qRa_2) was mapped to Satt144 on LG F/Chr.13. The authors suggested that greater isoflavone concentration could help protect soybean from aphid attacks as some members of isoflavones have been reported as having antibiosis effects on some soybean pests. For example, genistein and daidzin are toxic to southern green stink bug, cabbage looper and velvetbean caterpillar (Sharma and Norris 1991; Hoffmann-Campo et al. 2001; Piubelli et al. 2003, 2005). A Brazilian soybean cultivar, IAC-100, has been reported as having high isoflavone (daidzin and genistin) concentrations (Carrao-Panizzil and Kitamura 1995) and antibiotic resistance against stink bugs (Rosseto 1989).

In addition to qRa_1 (Meng et al. 2011) and the minor QTL from PI 567301B (Jun et al. 2012), the antixenosis-resistance QTL, *[Rag6]_P203*, was delimited to a 192-kb interval between SSR_08_75 and SSR_08_88 on LG A2/Chr. 8 in a Chinese soybean line P203 (Xiao et al. 2013). According to the soybean reference genome (Glyma.Wm82.a1v1), five genes are present in this 192-kb interval, of which a gene encoding Ser/Thr protein kinase was proposed as the strongest candidate for *[Rag6]_P203* (Xiao et al. 2013). Compared to the NBS-LRR genes of *Rag1* and *Rag2*, this Ser/Thr protein kinase gene was believed conferring a broad resistance to different aphid populations as it recognizes conserved pathogen-associated molecular patterns, playing a major role in the first layer of the plant immune system (Jones and Dangl 2006)

Recently, Alt and Ryan-Mahmutagic (2013) reported that a new aphid biotype, Biotype 4, discovered in Wisconsin readily colonized the previously known resistant soybean genotypes, including ‘Dowling’, PI 243540, PI 200538, *Rag1*/*Rag2* pyramided material, PI 567541B and PI

567598B. Among the tested genotypes, PI 567543C and PI 567597C remained resistance to Biotype 4 (Alt and Ryan-Mahmutagic 2013). The four reported aphid biotypes (Table 1.1) and unknown virulent biotypes challenge breeders to continually seek new resistance sources. Kim et al. (2014) detected a possible new allele or gene at the *Rag1* region in PI 587732, conferring antibiosis resistance against Biotype 2. Nurden et al. (2010) reported that the antixenosis resistance against Biotype 2 in PI 71506 was controlled by a single dominant gene that is distinct from *Rag1*. This unknown gene need further study to understand its location in the soybean genome and to provide additional resistance. A single dominant gene controlling antibiosis resistance was mapped to an interval different from *Rag2*, *rag4*, *Rag5* on LG F/Chr. 13 and therefore referred to as a new aphid-resistance gene, R_P746 (Xiao et al. 2014). All the identified aphid-resistance QTLs are summarized in Table 1.2.

Table 1.1 Soybean aphid biotypes and their virulence to soybean resistance genes

Biotype	Aphid-resistance gene and the resistance source									Reference
	Rag Jackson	*Rag1 Dowling*	*Rag2 PIs 243540, 200538*	*Rag1 + Rag2 Pyramid*	*Rag3* PI 567543C	*rag1b, rag3* PI 567598B	*rag1c, rag4* PI 567541B	*Rag3e* PI 567597C	*Rag5* PI 567301B	
1	-	-	-	-	-	-	-	-	-	Hill et al. 2004
2	+	+	-	-	-	-	-	-	-	Kim et al. 2008
3	NA	-/+*	+	-	-	-	-/+¶	-	NA	Hill et al. 2010
4	NA	+	+	+	-	+	+	-	NA	Alt and Ryan-Mahmutagic 2013

+ Implies the aphid biotype readily colonize on the soybean plants with the Rag gene
– Implies the aphid biotype cannot colonize on the soybean plants with the Rag gene
NA represents 'Not Available' in the literature
* Soybeans with *Rag1 were* resistant to Biotype 3 in no-choice tests but susceptible to Biotype 3 in choice tests
¶ PI 567541B was moderate resistant to Biotype 3 in choice tests but susceptible to Biotype 3 in no-choice tes

Table 1.2 Reported aphid-resistance QTLs in soybean germplasm

Gene	Source(s)	LG/Chr.	Flanking markers	Physical Position(bp)[a]	R^{2b}	Resistance modality	Reference(s)
Rag / Rag1	'Jackson' / 'Dowling'	M/7	SNPKS9-3 -- SNPKS5	5,608,084 - 5,492,694	NA[c]	Primarily antibiosis	Hill et al. 2006a,b; Li et al. 2007; Kim et al. 2010a
rag1b	PI 567598B	M/7	Satt567 -- Satt435	4,510,477 - NA	14.0-35.5%	Antibiosis	Bales et al. 2013; Mensah 2008
rag1c	PI 567541B	M/7	Satt299 -- Satt435	NA	44.7-87.7%	Antibiosis	Mensah 2008; Zhang et al. 2009
Rag2	PIs 200538, 243540	F/13	SNP46169.7 -- SNP21A	29,212,318 - 29,266,469	NA	Antibiosis	Hill et al. 2009; Kim et al. 2010b; Mian et al. 2008a, b
Rag3	PI 567543C	J/16	Sat_339 -- Satt414	NA	74.3-90.4%	Antixenosis	Mensah et al. 2005; Zhang et al. 2010
rag3	PI 567598B	J/16	Satt285 -- Satt414	2,802,418 - NA	28.4-45.8%	Antibiosis	Bales et al. 2013; Mensah 2008
Rag3b	PI 567537	J/16	Satt654 -- Sat_399	9,145,723 - 7,799,265	78.9-87.4%	Antibiosis	Zhang et al. 2013
Rag3d	PI 567585A	J/16	MSUSNP16-44 -- MSUSNP16-124	6,438,676 - 6,484,276	93.1%	Antibiosis	Liu 2010; Du 2016
Rag3e	PI 567597C	J/16	MSUSNP16-13 -- MSUSNP16-124	6,424,067 - 6,484,676	90%	Antixenosis	Du 2016
rag4	PI 567541B	F/13	Satt348 -- Satt649	5,491,250 - 12,953,321	0.9-9.2%	Antibiosis	Mensah 2008; Zhang et al. 2009
Rag5	PI 567301B	F/13	BARCSOYSSR_13_1131 -- BARCSOYSSR_13_1148	29,036,526 - 29,548,875	75-91%	Antixenosis	Jun et al. 2012; Mian et al. 2008a
[Rag6]_P203	P203	A2/8	SSR_08_75--SSR_08_88	39,218,719 - 39,410,489	NA	Antixenosis	Xiao et al. 2013
R_P746	P746	F/13	BARCSOYSSR_13_1278 -- BARCSOYSSR_13_1363	31,803,199 - 33,448,866	NA	Antibiosis	Xiao et al. 2014
qRa_1	Zhongdou 27	A2/8	Satt470	35,187,929	25-35%	Antibiosis	Meng et al. 2011
qRa_2	Zhongdou 27	F/13	Satt144	36,462,969	7-11%	Antibiosis	Meng et al. 2011
QTL_6_1	PI 567324	C2/6	BARCSOYSSR_06_0998	18,713,522	4.4-11.6%	Antixenosis	Jun et al. 2013
QTL_13_1	PI 567324	F/13	BARCSOYSSR_13_1139	29,274,967	42.7-70.6	Antixenosis	Jun et al. 2013
QTL_13_2	PI 567324	F/13	Satt649	12,953,321	2.1-13.1%	Antixenosis	Jun et al. 2013

[a] Physical position is according to Glyma.Wm82.a1 (Schmutz et al. 2010)
[b] R^2 represents the percentage of phenotypic variation explained by a QTL
[c] NA represents 'Not Available' in the literature

Genetic approaches to sustain host-plant resistance against soybean aphids

To date, four soybean aphid biotypes have been discovered (Kim et al. 2008; Hill et al. 2010; Alt and Ryan-Mahmutagic 2013) (Table 1.1), and there are likely more unknown virulent biotypes not yet reported. In addition to the continuing discovery of new resistance sources, pyramiding of different *Rag* genes (Table 1.2) with the assistance of flanking markers could provide cultivars with broader and more durable resistance. Wiarda et al. (2012) investigated aphid development on soybeans with *Rag1* alone, *Rag2* alone and both genes combined, and discovered soybeans with both genes were more resistant to aphids. A similar investigation also confirmed that pyramiding *Rag1* and *Rag2* provides yield protection from aphids in North America (McCarville et al. 2014). Therefore, pyramiding different *Rag* genes, especially with different resistance modalities, has potential to combat diverse and dynamic aphid populations.

Gene deployment based on biotype distribution is another effective method to combat different aphid populations geographically. Therefore, knowledge regarding distribution of the different biotypes is important. Cooper et al. (2015) studied the geographic distribution of aphid biotypes at ten locations between 2008 and 2010 and developed a panel of host differentials (indicator lines) to characterize aphid biotypes. According to this study, aphid populations had been diverse and dynamic across the U.S. and Canada. Additionally, PI 567598B and PI 567541B were identified as the most resistant and durable genotypes against aphid populations. The authors inferred the high level of resistance was due to the natural pyramids of two recessive genes in these two accessions. Deploying different *Rag* gene(s) according to the soybean aphid biotype distribution could avoid genetic vulnerability of a certain resistant cultivar over a large geographic area.

Candidate genes for aphid resistance in soybean

The candidate genes of *Rag1* and *Rag2* were identified as NBS-LRR genes (Kim et al. 2010a, b), which are known to play critical roles in host-plant defense against insects or diseases (Marone et al. 2013). NBS-LRR genes were also predicted as candidate genes for aphid-resistance in other crops, including the *Mi* gene in tomato that confers resistance to potato aphid (*Macrosiphum euphorbiae*) (Rossi et al. 1998; Kaloshian et al. 2000; Cooper et al. 2004), the *Vat* gene in melons underlying resistance to the melon aphid (*Aphis gossypii*) (Villada et al. 2009), the AKR gene against the blue-green aphid (*Acyrthosiphon kondoi* Shinji) in *Medicago truncatula* (Klingler et al. 2005), the TTR gene against the spotted alfalfa aphid (*Therioaphis trifolii*) (Klingler et al. 2007) and the *RAP1* gene resistant to the pea aphid (*Acyrthosiphon pisum*) (Stewart et al. 2009).

In addition to NBS-LRR genes as candidates for aphid-resistance genes, a serine/threonine protein kinase encoding gene was predicted as the candidate gene of *[Rag6]_P203* (Xiao et al. 2013). As the serine/threonine protein kinase belongs to the family of transmembrane pattern recognition receptors that recognize conserved pathogen-associated molecular patterns, it was believed to play an important role in the first layer in plant immune system that provided broad resistance against different aphid isolates in P203 (Xiao et al. 2013).

Studham and MacIntosh (2013) reported that the soybean aphid colonization leads to a decrease of poly-unsaturated fatty acids, which are used by soybean plants for Jasmonic acid (JA) biosynthesis. JA signaling triggered by aphid infestation is known to play a critical role in regulating plant defense (Thompson et al. 2006). Li et al. (2008) also discovered that the direct feeding from soybean aphids partially activates JA-regulated signaling pathways in soybean

defense. Additionally, hundreds of transcripts induced by soybean aphids in the susceptible plants were related to hormone signaling pathways, including abscisic acid and ethylene pathways during plant defense (Studham and MacIntosh 2013).

Next-gen sequencing technologies applied to soybean genomics study

Recently, massively parallel sequencing platforms have become wildly available, which lead to the dramatic reduction in the cost. In 2010, a high-quality soybean reference genome was built by sequencing the cultivar 'Willams 82' with whole-genome shutgun sequencing approach (Schmutz et al. 2010). With the availability of the soybean reference genome, SNPs and insertion/deletions have been efficiently identified by aligning the sequencing reads from diverse soybean genotypes to the reference genome. Song et al. (2013) sequenced reduced representation libraries from six cultivated and two wild soybean (*G. soja* Sieb. et Zucc.) genotypes; a total of 52,041 SNPs identified from this reduced representation sequencing were used to produce the SoySNP50K iSelect BeadChip. A mapping population consisting of 246 recombinant inbred lines were sequenced at an average of 0.19x depth and 109,273 SNPs were identified and used to construct a linkage map; three QTLs were identified as resistant to southern root-knot nematode (Xu et al. 2013). Li et al. (2014) conducted the *de novo* assembly of seven phylogenetically and geographically representative *G. soja* accessions and discovered a broader range of NBS LRR-gene domain architectures present in the in the *G. soja* genome than in the *G. max* genome.

In addition to DNA sequencing, transcriptome sequencing also has been applied in soybean genomics studies. Severin et al. (2010) sequenced the transcriptomes of fourteen diverse tissues; the transcripts discovered from this study greatly helped evaluate gene model annotations for the soybean reference genome. Lee et al. (2017) investigated the transcriptome profiles of soybean

near-isogenic lines either with the resistant *Rag5* allele or the susceptible *rag5* allele before and after the infestation with soybean aphid Biotype 2, and discovered three differentially expressed genes near the *Rag5* locus as strong candidate genes.

Aims of dissertation research

G. soja 85-32 was identified as possessing strong resistance against aphids by Yang et al. (2004). Later, the resistance in *G. soja* 85-32 was initially discovered as being controlled by two QTLs within a bi-parental population (070020) consisting of 140 $F_{3:4}$ lines (Zhang 2012). The first objective of the present study was to validate these two QTLs in different genetic backgrounds provided by populations 110193 and 110201, and investigate the inheritance manner of these two QTLs with using the F_3-derived lines from population 070020. The second objective was to fine map these two QTLs using SNPs discovered from whole genome re-sequencing of the resistant parent (E12901) and the high throughput genome-wide genotyping technology realized by the Illumina Infinium SoySNP50K/8K iSelect BeadChip. The third objective was to identify structural variants in the exons of candidate genes using the whole-genome exome capture sequencing approach. The fourth objective was to develop and evaluate soybean advanced breeding lines pyramided with different aphid-resistance allelic combinations to combat the dynamic aphid populations in Michigan. Discoveries from the present dissertation will be significant resources for improving soybeans with aphid resistance and for future functional genetics studies in unraveling the interaction between host-plant resistance and soybean aphid biotypes.

REFERENCES

REFERENCES

Alfonso-Rubí J, Ortego F, Castañera P, Carbonero P, Díaz I (2003) Transgenic expression of trypsin inhibitor CMe from barley in indica and japonica rice, confers resistance to the rice weevil Sitophilus oryzae. Transgenic research 12(1):23-31

Alleman RJ, Grau CR, Hogg DB (2002) Soybean aphid host range and virus transmission efficiency. In: Proceedings of Wisconsin Fertilizer, Aglime, and Pest Management Conference. Madison, WI. http://www. soils. wisc. edu/extension/FAPM/fertaglime02. htm

Alt J, Ryan-Mahmutagic M (2013) Soybean aphid biotype 4 identified. Crop Sci 53(4):1491-1495

Azzouz H, Cherqui A, Campan ED, Rahbe Y, Duport G, Jouanin L, Kaiser L, Giordanengo P (2005) Effects of plant protease inhibitors, oryzacystatin I and soybean Bowman-Birk Inhibitor, on the aphid *Macrosiphum euphorbiae* (Homoptera, Aphididae) and its parasitoid *Aphelinus abdominalis* (Hymenoptera, Aphelinidae). J. Insect Physiol 51(1):75-86

Bales C (2013) Fine mapping of aphid resistance genes in soybean plant introduction (PI) 567598B. Dissertation, Michigan State University, East Lansing, MI

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. Theor Appl Genet 126:2081-2091

Broadway RM, Duffey SS (1986) Plant proteinase inhibitors: mechanism of action and effect on the growth and digestive physiology of larval *Heliothis zea* and *Spodoptera exiqua*. J. Insect Physiol 32(10):827-833

Burden BJ, Norris DM (1992) Role of the isoflavenoid coumestrol in the constitutive antixenosic properties of “Davis” soybeans against and oligophagous insect, the Mexican bean beetle. J. Chem. Ecol. 18(7):1069-1081

Carrao-Panizzil MC, Kitamura K (1995) Isoflavone content in Brazilian soybean cultivars. Japanese Journal of Breeding 45:295-300

Caviness CE and Walters HJ (1966) Registration of 'Davis' Soybean. Crop Sci 6(5): 502

Chiang HS, Norris DM, Ciepiela A, Shapiro P, Oosterwyk A (1987) Inducible versus constitutive PI 227687 soybean resistance to Mexican bean beetle, *Epilachna varivestis*. J Chem Ecol 13(4):741-749

Clark AJ, Perry KL (2002) Transmissibility of field isolates of soybean viruses by *Aphis glycines*. Plant Dis 86(11):1219-1222

Cooper SG, Concibido V, Estes R, Hunt D, Jiang GL, Krupke C, McCornack B, Mian R, O'Neal M, Poysa V, Prischmann-Voldseth D (2015) Geographic distribution of soybean aphid biotypes in the United States and Canada during 2008–2010. Crop Sci 55 (6):2598-2608

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

Davis JA, Radcliffe EB, Ragsdale DW (2005) Soybean aphid, *Aphis glycines* Matsumura, a new vector of potato virus Y in potato. Am J Potato Res 82(3):197-201

De Leo F, Bonadé-Bottino M, Ceci LR, Gallerani R, Jouanin L (2001) Effects of a mustardtrypsin inhibitor expressed in different plants on three lepidopteran pests. Insect Biochem Mol Biol 31(6):593-602

DiFonzo C (2005) Soybean aphid control using insecticides (Michigan State University: Michigan). http://fieldcrop.msu.edu/uploads/documents/Soy%20Aphid%20Insecticides.pdf

Du W (2016) Map and fine map aphid resistance genes in soybean plant introduction (PI) 567597C, 567585A and 567537. Dissertation, Michigan State University, East Lansing, MI

Falco MC, Silva-Filho MC (2003) Expression of soybean proteinase inhibitors in transgenic sugarcane plants: effects on natural defense against *Diatraea saccharalis*. Plant Physiol Biochem 41(8):761-766

Fernandez-Cornejo J, Nehring R, Osteen C, Wechsler S, Martin A, Vialou A (2014) Pesticide Use in U.S. Agriculture: 21 Selected Crops, 1960-2008. Washington DC: US Department of Agriculture. Economic Research Service, Economic Information Bulletin. 124

Fox CM, Kim KS, Cregan PB, Hill CB, Hartman GL, Diers BW (2014) Inheritance of soybean aphid resistance in 21 soybean plant introductions. Theor Appl Genet 127(1):43-50

Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl Acids Res 38:D843-D846. doi:10.1093/nar/gkp798

Hart SV, Kogan M, Paxton JD (1983) Effect of soybean phytoalexins on the herbivorous insects Mexican bean beetle and soybean looper. J Chem Ecol 9(6):657-672

Hartman GL, Domier LL, Wax LM, Helm CG, Onstad DW, Shaw JT, Solter LF, Voegtlin DJ, d'Arcy CJ, Gray ME, Steffey KL (2001) Occurrence and distribution of *Aphis glycines* on soybeans in Illinois in 2000 and its potential control. Plant Health Progr doi:10.1094/PHP-2001-0205-01-HN

Hilder VA, Gatehouse AM, Sheerman SE, Barker RF, Boulter D (1987) A novel mechanism of insect resistance engineered into tobacco. Nature, 330:160-163

Hill JH, Alleman R, Hogg DB, Grau CR (2001) First report of transmission of Soybean mosaic virus and Alfalfa mosaic virus by *Aphis glycines* in the New World. Plant Dis 85(5):561-561

Hill CB, Chirumamilla A, Hartman GL (2012) Resistance and virulence in the soybean-*Aphis glycines* interaction. *Euphytica*, 186(3):635-646

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

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, Li Y, Hartman GL (2006a) A single dominant gene for resistance to the soybean aphid in the soybean cultivar Dowling. Crop Sci 46(4):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(4):1606-1608

Hodgson EW, McCornack BP, Tilmon K, Knodel JJ (2012) Management recommendations for soybean aphid (Hemiptera: Aphididae) in the United States. Journal of Integrated Pest Management 3(1):E1-E10

Hoffmann-Campo CB, Harborne JB, McCaffery AR (2001) Pre-ingestive and post-ingestive effects of soya bean extracts and rutin on Trichoplusia ni growth. Entomologia Experimentalis et Applicata 98(2):181-194

Hymowitz T, Harlan JR (1983) Introduction of soybean to North America by Samuel Bowen in 1765. Economic Botany 37(4):371-379

Johnson R, Narvaez J, An G, Ryan C (1989) Expression of proteinase inhibitors I and II in transgenic tobacco plants: effects on natural defense against *Manduca sexta* larvae. Proc Natl Acad Sci 86(24):9871-9875

Jones JD, Dangl JL (2006) The plant immune system. Nature, 444(7117):323-329

Jun TH, Mian MAR, Michel AP (2012) Genetic mapping revealed two loci for soybean aphid resistance in PI 567301B. Theor Appl Genet 124:13-22

Jun TH, Mian MAR, Michel AP (2013) Genetic mapping of three quantitative trait loci for soybean aphid resistance in PI 567324. Heredity 111:16-22

Kaloshian I, Kinsey MG, Williamson VM, Ullman DE (2000) *Mi*-mediated resistance against the potato aphid *Macrosiphum euphorbiae* (Hemiptera: *Aphididae*) limits sieve element ingestion. Environ Entomol 29: 690-695

Kim CS, Schaible G, Garrett L, Lubowski R, Lee D (2008) Economic impacts of the US soybean aphid infestation: a multi-regional competitive dynamic analysis. Agricultural and resource economics review 37(2):227-242

Kim KS, Diers BW (2009) The associated effects of the soybean aphid resistance locus on soybean yield and other agronomic traits. Crop Sci 49(5):1726-1732

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

Kim KS, Chirumamilla A, Hill CB, Hartman GL, Diers BW (2014) Identification and molecular mapping of two soybean aphid resistance genes in soybean PI 587732. Theor Appl Genet 127(5):1251-1259

Kim KS, Hill CB, Hartman GL, Hyten DL, Hudson ME, and Diers BW (2010b) Fine mapping of the soybean aphid-resistance gene *Rag2* in soybean PI 200538. Theor Appl Genet 121:599-610

Kim KS, Hill CB, Hartman GL, Mian MA, Diers BW (2008) Discovery of soybean aphid biotypes. Crop Sci 48(3):923-928

Klingler J, Creasy R, Gao L, Nair RM, Calix AS, Jacob HS, Edwards OR, Singh KB (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

Klingler JP, Edwards OR, Singh KB (2007) Independent action and contrasting phenotypes of resistance genes against spotted alfalfa aphid and bluegreen aphid in Medicago truncatula. New Phytologist 173(3):630-640

Lecardonnel A, Chauvin L, Jouanin L, Beaujean A, Prévost G, Sangwan-Norreel B (1999) Effects of rice cystatin I expression in transgenic potato on Colorado potato beetle larvae. Plant Science 140(1):71-79

Lee S, Cassone BJ, Wijeratne A, Jun TH, Michel AP, Mian MR (2017) Transcriptomic dynamics in soybean near-isogenic lines differing in alleles for an aphid resistance gene, following infestation by soybean aphid biotype 2. BMC genomics 18(1):472

Lemos Filho JP, Paiva ÉA (2006) The effects of sooty mold on photosynthesis and mesophyll structure of mahogany (Swietenia macrophylla King., Meliaceae). Bragantia. 65(1):11-17

Li Y, Zou J, Li M, Bilgin DD, Vodkin LO, Hartman GL, Clough SJ (2008) Soybean defense responses to the soybean aphid. New Phytol 179(1):185-195

Li YH, Zhou G, Ma J, Jiang W, Jin LG, Zhang Z, Guo Y, Zhang J, Sui Y, Zheng L, Zhang SS (2014) *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nature Biotechnol 32(10):1045-1052

Li Y, Hill CB, Hartman GL (2004) Effect of three resistant soybean genotypes on the fecundity, mortality, and maturation of soybean aphid (Homoptera: Aphididae). J Econ Entomol 97(3):1106-1111

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 Breeding 19:25-34

Liu M (2010) Inheritance of resistance and molecular mapping of soybean aphid resistance genes in soybean PI 567585A; identification of aphid resistance genes in soybean using modified nested association mapping (MNAM). Ph.D. dissertation, Michigan State University

Lundin O, Rundlöf M, Smith HG, Fries I, Bommarco R (2015) Neonicotinoid insecticides and their impacts on bees: a systematic review of research approaches and identification of knowledge gaps. PLoS One 10(8):e0136928

Malumphy CP (1997) Morphology and anatomy of honeydew eliminating organs. In: Ben-Dov and Hodgson CJ (eds) World Crop Pests. Elsevier Science B.V., Amsterdam, The Netherlands. 7:269-274

Marone D, Russo MA, Laidò G, De Leonardis AM, Mastrangelo AM. Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses (2013) Int J Mol Sci 14(4):7302-7326

McCarville M, Hodgson E, O'Neal M (2012) Soybean aphid- resistant soybean varieties for Iowa. Iowa State University Extension and Outreach PM 3023, Ames, IA, USA

McCarville MT, O'Neal ME (2013) Soybean aphid (Aphididae: *Hemiptera*) population growth as affected by host plant resistance and an insecticidal seed treatment. J Econ Entomol 106:1302-1309

McCarville MT, O'Neal ME, Potter BD, Tilmon KJ, Cullen EM, McCornack BP, Tooker JF, Prischmann-Voldseth DA (2014) One gene versus two: a regional study on the efficacy of single gene versus pyramided resistance for soybean aphid management. J Econ Entomol 107:1680-1687

Meng F, Han Y, Teng W, Li Y, 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 123:1459-1465

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

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

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 PI243540. Theor Appl Genet 117:955-962

Michel AP, Mian MR, Davila-Olivas NH, Cañas LA (2010) Detached leaf and whole plant assays for soybean aphid resistance: differential responses among resistance sources and biotypes. J Econ Entomol 103(3):949-957

Van Nurden A, Scott RA, Hesler L, Tilmon K, Glover K, Carter C (2010) Inheritance of soybean aphid resistance from PI 71506. Journal of Crop Improvement 24:400-416

Ohnesorg WJ, Johnson KD, O'Neal ME (2009) Impact of reduced risk insecticides on soybean aphid and their natural enemies. J Econ Entomol 102:1816-1826

Painter RH (1958) Resistance of Plants to Insects. Annu Rev Entomol 3:267-290

Piubelli GC, Hoffmann-Campo CB, De Arruda IC, Franchini JC, Lara FM (2003) Flavonoid increase in soybean as a response to *Nezara viridula* injury and its effect on insect-feeding preference. J Chem Ecol 29(5):1223-1233

Piubelli GC, Hoffmann-Campo CB, Moscardi F, Miyakubo SH, Neves De Oliveira MC (2005). Are chemical compounds important for soybean resistance to *Anticarsia gemmatalis*. J Chem Ecol 31(7):1509-1525

Ragsdale DW, Voegtlin DJ, O'Neil RJ (2004) Soybean aphid biology in North America. Ann Entomol Soc Am 97(2):204-208

Ragsdale DW, McCornack BP, Venette RC, Potter BD, MacRae IV, Hodgson EW, O'Neal ME, Johnson KD, O'Neil RJ, DiFonzo CD, Hunt TE (2007) Economic threshold for soybean aphid (Hemiptera: Aphididae). J Econ Entomol 100:1258-1267

Ragsdale DW, Landis DA, Brodeur J, Heimpel GE, Desneux N (2011) Ecology and management of the soybean aphid in North America. Annu Rev Entomol 56(1):375-399

Rahbé Y, Ferrasson E, Rabesona H, Quillien L (2003) Toxicity to the pea aphid Acyrthosiphon pisum of anti-chymotrypsin isoforms and fragments of Bowman–Birk protease inhibitors from pea seeds. Insect Biochem Mol Biol 33(3):299-306

Rosseto CJ (1989) Breeding for resistance to stink bugs, In A. J. Pascale [ed.], Proceedings of the World Soybean Research Conference IV, Buenos Aires, Argentina. Assoc. Argentina de la Soja Press, Buenos Aires, Argentina. 2046-2060

Rossi M, Goggin FL, Milligan SB, Kaloshian I, Ullman DE, Williamson VM (1998) The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc Natl Acad Sci 95:9750-9754

Roulin A, Auer PL, Libault M, Schlueter J, Farmer A, May G, Stacey G, Doerge RW, Jackson SA (2013) The fate of duplicated genes in a polyploid plant genome. Plant J 73:143-153

Rutledge CE, O'Neil RJ, Fox TB, Landis DA (2004) Soybean aphid predators and their use in integrated pest management. Ann Entomol Soc Am 97(2):240-248

Ryan CA (1990) Protease inhibitors in plants: genes for improving defenses against insects and pathogens. Annu Rev Phytopathol 28(1):425-449

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178-183

Severin AJ, Woody JL, Bolon YT, Joseph B, Diers BW, Farmer AD, Muehlbauer GJ, Nelson RT, Grant D, Specht JE, Graham MA (2010) RNA-Seq Atlas of Glycine max: a guide to the soybean transcriptome. BMC Plant Biol 10(1):160

Sharma HC, Norris DM (1991) Chemical basis of resistance in soy bean to cabbage looper, *Trichoplusia ni*. J Sci Food Agr 55:353-364

Shoemaker RC, Polzin K, Labate J, Specht J, Brummer EC, Olson T, Young N, Concibido V, Wilcox J, Tamulonis JP, Kochert G (1996) Genome duplication in soybean (Glycine subgenus soja). Genetics. 144(1):329-38

Singh RJ (ed) (2006) Genetic resources, chromosome engineering, and crop improvement: vegetable crops (Vol. 3). CRC press, Boca Raton, FL

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. doi:10.1371/journal.pone.0054985

SoyStats. http://soystats.com/

Stamp N (2003) Out of the quagmire of plant defense hypotheses. The Quarterly Review of Biology, 78(1):23-55

Stewart SA, Hodge S, Ismail N, Mansfield JW, Feys BJ, Prospéri JM, Huguet T, Ben C, Gentzbittel L, Powell G (2009) The RAP1 gene confers effective, race-specific resistance to the pea aphid in *Medicago truncatula* independent of the hypersensitive reaction. Mol Plant Microbe Interact 22(12):1645-1655

Studham ME, MacIntosh GC (2013) Multiple phytohormone signals control the transcriptional response to soybean aphid infestation in susceptible and resistant soybean plants. Mol Plant Microbe Interact 26(1):116-129. doi: 10.1094/MPMI-05-12-0124-FI

Thompson GA, Goggin FL (2006) Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. J Exp Bot 57(4):755-766. doi: 10.1093/jxb/erj135

Tran P, Cheesbrough TM, Keickhefer RW (1997) Plant proteinase inhibitors are potential anticereal aphid compounds. J Econ Entomol 90(6):1672-1677

Villada ES, González EG, López-Sesé AI, Castiel AF, Gómez-Guillamón ML (2009) Hypersensitive response to *Aphis gossypii* Glover in melon genotypes carrying the Vat gene. J Exp Bot 60:3269-3277

Wiarda SL, Fehr WR, O'Neal ME (2012) Soybean aphid (Hemiptera: Aphididae) development on soybean with *Rag1* alone, *Rag2* alone, and both genes combined. J Econ Entomol 105(1):252-258

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(2):209-218

Xiao L, Hu Y, Wang B, 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 126:2279-2287

Xiao L, Zhong YP, Wang B, Wu TL (2014) Mapping an aphid resistance gene in soybean [*Glycine max* (L.) Merr.] P746. Genetics and Molecular Research 13(4):9152-9160

Xu X, Zeng L, Tao Y, Vuong T, Wan J, Boerma R, Noe J, Li Z, Finnerty S, Pathan SM, Shannon JG (2013) Pinpointing genes underlying the quantitative trait loci for root-knot nematode resistance in palaeopolyploid soybean by whole genome resequencing. Proc Natl Acad Sci 110:13469-13474

Yang Z, Honda K, Wang S, Ma X (2004) Re-evaluation of Glycine soja germplasm from north eastern China for aphid resistance. Journal of Jilin Agricultural Sciences 29(5):3-6

Zhang Z (2012) QTL mapping of Phytophthora root rot resistance and aphid resistance in soybean. Dissertation, Michigan State University, East Lansing, MI

Zhang G, Gu C, Wang D (2009) Molecular mapping of soybean aphid resistance genes 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

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

Zhang XR (1988) Study on the instars of soybean aphid, *Aphis glycines* Matsumura, in Jilin province. Translation from Journal of Jilin Agricultural University (Acta Agriculturae Universitatis Jilinensis). 10:15-17

CHAPTER 2

VALIDATION OF QTLS CONTROLLING APHID RESISTANCE IN *GLYCINE SOJA* 85-32

The work presented in this chapter is part of the final publication:

Zhang S, Zhang Z, Bales C, Gu C, DiFonzo C, Li M, Song Q, Cregan P, Yang Z, Wang D (2017) Mapping novel aphid resistance QTL from wild soybean, *Glycine soja* 85-32. Theor Appl Genet. doi:10.1007/s00122-017-2935-z

Abstract

The soybean aphid is a major pest of soybean. E08934, derived from the wild soybean *Glycine soja* 85-32, has shown strong and consistent resistance to soybean aphids. Two major quantitative trait loci (QTLs) were previously detected as significantly associated with aphid resistance in a mapping population (070020) derived from the cross E08934 x E00003 (aphid-susceptible). With using indicator lines, E08934 was demonstrated resistant to all known aphid biotypes. E12901, derived from E08934, was used as the resistant parent to construct two validation populations. The BC_1F_2 population, 110193, was comprised of 262 individuals derived from E12901 x E00003. The F_2 population, 110201, was comprised of 396 individuals derived from the cross between E12901 x E09014 (aphid-susceptible). Both populations were evaluated for aphid resistance at three weeks and four weeks after the initial infestation in the greenhouse trial during Fall, 2012. With SNPs discovered from whole genome resequencing of E12901 or mined from the SoySNP50K iSelect BeadChip, the two aphid-resistance QTLs were successfully validated in populations 110193 and 110201; the QTL confirmed between MSUSNP08-40 and MSUSNP08-4 on chromosome 8 was designated *Rag6* whereas the QTL confirmed between MSUSNP16-10 and MSUSNP16-15 on chromosome 16 was designated *Rag3c*. No significant interaction between *Rag6* and *Rag3c* was detected. A total of 75 F_3-derived lines from the mapping population, 070020, were used to determine the gene action of *Rag6* or *Rag3c*. Both QTLs were demonstrated as additive; *Rag6* is partially dominant while *Rag3c* was not determined due to the small sample size of *Rag3c*-heterozygotes in the analysis. The new aphid-resistance gene(s) from the wild soybean *G. soja* 85-32 are valuable in breeding soybeans for aphid resistance.

For a full text of this work please go to https://doi.org/10.1007/s00122-017-2935-z.

CHAPTER 3

FINE MAPPING OF THE SOYBEAN APHID RESISTANCE GENES *RAG6* AND *RAG3C* FROM *GLYCINE SOJA* 85-32

The work presented in this chapter is under review:

Zhang S, Zhang Z, Wen Z, Gu C, An Y, Bales C, DiFonzo C, Song Q, Wang D (2017). Fine mapping of the aphid resistance genes *Rag6* and *Rag3c* from *Glycine soja* 85-32. Theor Appl Genet (Under Review)

Abstract

The soybean aphid, an invasive species, has significantly threatened soybean production in North America since 2001. Host-plant resistance is known as an ideal management strategy for aphids. Two novel aphid-resistance loci, *Rag6* and *Rag3c*, from *Glycine soja* 85-32, were previously detected in a 10.5 centiMorgan (cM)-interval on chromosome 8 and a 7.5 cM-interval on chromosome 16, respectively. Defining the exact genomic position of these two genes is critical for improving the effectiveness of marker-assisted selection for aphid resistance and for identification of the functional genes. To pinpoint the locations of *Rag6* and *Rag3c*, four populations segregating for *Rag6* and *Rag3c* were used to fine map these two genes. The availability of the Illumina Infinium SoySNP50K/8K iSelect BeadChip, combined with single nucleotide polymorphism (SNP) markers discovered through the whole genome re-sequencing of E12901, facilitated the fine mapping process. *Rag6* was refined to a 49-kb interval on chromosome 8 with four candidate genes, including three clustered nucleotide-binding site leucine-rich repeat (NBS-LRR) genes and an amine oxidase encoding gene. *Rag3c* was refined to a 150-kb interval on chromosome 16 with eleven candidate genes, two of which are a NBS-LRR gene and a lipase gene. Moreover, by sequencing the whole genome exome-capture of the resistant source (E12901), structural variants were identified in the exons of the candidate genes of *Rag6* and *Rag3c.* The closely linked SNP markers and the candidate gene information presented in this study will be significant resources for integrating *Rag6* and *Rag3c* into elite cultivars and for future functional genetics studies.

Introduction

The cultivated soybean, *Glycine max* (L.) Merr., is widely grown for multiple uses including livestock feed, cooking oil, protein source and biodiesel. However, over the past decade, North American soybean production has been threatened by an invasive species, the soybean aphid (*Aphis glycines* Matsumura), that originated from Asia (Wu et al. 2004). Since its discovery in the Great Lakes region in 2000 (Hartman et al. 2001), soybean aphid has aggressively spread to all the major soybean production areas in the United States and Canada (Ragsdale et al. 2011). Aphid feeding can lead to up to 40 % yield loss (Ragsdale et al. 2007) by the removal of nutrients and water from soybean plants. This results in stunted plants, reduction in yield components (such as seed number and seed weight), lowered oil production (Beckendorf et al. 2008; Ragsdale et al. 2011) and plant death under heavy infestations (Wu et al. 2004). Virus transmission by aphids also impairs soybean growth and yield by causing plant stunting, leaf deformation and reduced pod-fill (Hill et al. 2001; Clark and Perry 2002). Furthermore, aphids excrete sticky honeydew that can lead to the growth of sooty mold, which may block soybean plant photosynthesis and cause additional yield loss (Malumphy 1997; Lemos Filho and Paiva 2006).

After the rapid establishment of soybean aphids in North America, insecticide applications on soybeans increased from 0.03 million pounds in 2000 to 4.7 million pounds in 2008 (Fernandez-Cornejo et al. 2014). The growth of insecticide use not only increased costs of production, but also impacted the ecosystem by removing beneficial insects such as natural enemies and pollinators (Ohnesorg et al. 2009; Lundin et al. 2015). An alternative strategy of controlling aphids is to use the native host-plant resistance existing in soybean germplasm. Host-plant resistance can provide plants with economical, environmentally-friendly, and season-long

protection against insects or disease. To date, over thirty *G. max* plant introductions (PIs) and cultivars have been reported with antibiosis resistance (affecting insect growth, survival, or reproduction) or antixenosis resistance (affecting insect behavior) against aphids in North America (Hill et al. 2004a; Li et al. 2004; Mensah et al. 2005; Hesler et al. 2007; Mian et al. 2008a; Fox et al. 2014; Kim et al. 2014). Usually, choice and no-choice tests are used to distinguish between these two different types of host resistance (Mensah et al. 2005).

The aphid-resistance QTLs identified in North America were designated as *Rag* (Resistance to *Aphis glycines*) genes. *Rag* and *Rag1* were revealed as a single dominant QTL between Satt463 and Satt435 on chromosome 7 that controls antibiosis resistance to aphids in 'Dowling' (PI 548663) and 'Jackson' (PI 548657), respectively (Hill et al. 2006a, b; Li et al. 2007). The dominant antibiosis resistance gene *Rag2* was detected between Satt334 and Sct_033 on chromosome 13 in PI 200538 and PI 243540 (Kang et al. 2008; Mian et al. 2008b; Hill et al. 2009). *Rag5,* from PI 567301B, was mapped to the same interval as *Rag2*, but it confers antixenosis resistance (Jun et al. 2012). Zhang et al. (2010) reported a dominant QTL, *Rag3*, between Sat_339 and Satt414 on chromosome 16, delivering antixenosis resistance in PI 567543C. Another single dominant QTL that controls antibiosis resistance in PI 567537 was later assigned to the same region as *Rag3*, and therefore designated *Rag3b* (Zhang et al. 2013). The antibiosis aphid-resistance in PI 567541B was controlled by two recessive QTLs, *rag1c* and *rag4* (Mensah et al. 2008; Zhang et al. 2009). The recessive *rag1c* was located between Satt299 and Satt435 on chromosome 7 (Zhang et al. 2009), which is in close proximity with *Rag*/*Rag1* (Li et al. 2007). The recessive *rag4* was assigned to an interval between Satt649 and Satt348 on chromosome 13 (Zhang et al. 2009), which is different from the location of *Rag2* and *Rag5* (Kang et al. 2008; Mian et al. 2008b; Hill et al. 2009; Jun et al. 2012). Bales et al. (2013) mapped

two recessive QTLs conferring antibiosis resistance in PI 567598B to the same regions as *Rag*/*Rag1* (Li et al. 2007) and *Rag3* (Zhang et al. 2010), and designated these QTLs as *rag1b* and *rag3*, respectively.

Some of the *Rag* QTLs clearly overlap physically or are in close proximity. Fine mapping studies or allelism tests of these QTLs are needed to unravel their relationships. Among these *Rag* QTLs, only *Rag1* and *Rag2*-PI 200538 have been fine mapped thus far (Kim et al. 2010a, b). *Rag1* was refined to a 115-kb interval on chromosome 7 with two NBS-LRR genes proposed as the best candidates (Kim et al. 2010a). The fine mapping of *Rag2* from PI 200538 delimited it to a 54-kb region on chromosome 13, and again, a NBS-LRR gene was the strongest candidate (Kim et al. 2010b). Kim et al. (2014) detected QTLs controlling antibiosis resistance against aphids in the fine-mapped *Rag1* and *Rag2*-PI 200538 regions (Kim 2010a, b) from PI 587732 that might provide different genes or alleles from *Rag1* and *Rag2*. Additional fine mapping studies or allelism tests are needed to understand the relationships among these *Rag* QTLs that were mapped to similar genomic regions.

It is believed that cultivated soybean (*G. max*) was domesticated from *Glycine soja*, a wild annual species native to Asia (Singh 2006). *G. soja* has been reported as resistant to a wide range of diseases and insects, including soybean aphid (Hill et al. 2004b; Yang et al. 2004). A broader range of NBS R-gene domain architectures were discovered present in the *G. soja* genome than in the *G. max* genome after the *de novo* assembly of seven phylogenetically and geographically representative *G. soja* accessions (Li et al. 2014). *Glycine soja* 85-32 was reported as resistant to aphids by Yang et al. (2004). This resistance was later identified as antibiosis delivered by two partially dominant QTLs, *Rag6* and *Rag3c* (Zhang et al. 2017). *Rag6* was located at a 1.7 Mb-interval (40.9 – 42.6 Mb, Glyma.Wm82.a2 hereafter) with the mapping population and a 6.0 Mb-

interval (40.0 – 46.0 Mb) with the validation populations on chromosome 8 (Zhang et al. 2017); *Rag6* is likely a novel gene as it was mapped to a different interval from any other aphid-resistance QTL previously identified on chromosome 8 (Meng et al. 2011; Jun et al. 2012; Xiao et al. 2013). *Rag3c* was located at a 0.9 Mb interval (6.3 - 7.2 Mb) with the mapping population and a 1.9-Mb interval (6.3 - 8.2 Mb) with the validation populations on chromosome 16 (Zhang et al. 2017); it is within the region of *Rag3*, *Rag3b*, and *rag3* on chromosome 16 (Zhang et al. 2010, 2013; Bales et al. 2013). Despite good insects/disease resistance genes, *G. soja* carries undesirable agronomic traits that restrict its direct application in commercial breeding programs. Therefore, fine mapping is needed to develop markers closely linked to *Rag6* and *Rag3c* to assist efficient introgression of aphid-resistance from *G. soja* 85-32 to cultivated soybeans with minimum negative linkage drags.

The objectives of this study were to: (1) fine map *Rag6* and *Rag3c* to identify closely linked markers that could be useful in marker-assisted selection, (2) assess the robustness of these markers in assisting selections for these two genes in breeding populations and (3) identify structural variations within the candidate genes of *Rag6* and *Rag3c* by aligning the whole genome exome-capture sequencing reads of the resistant source (E12901) to the reference genome, Glyma.Wm82.a1 (Schmutz et al. 2010).

Materials and Methods

Plant materials

E12901 is an advanced breeding line derived from *G. soja* 85-32, and has the same aphid resistance phenotype and genotype (*Rag6* + *Rag3c*) (Zhang et al. 2017). As the resistant parent, E12901 was crossed with three aphid-susceptible parents (E00003, E09014 and E09088) to

construct four independent populations (110193, 110201, 110202-1, and 110202-2). All four parents are fully homozygous inbred lines. None of the susceptible parents carries any known aphid-resistance gene. The present fine mapping study started with a total of 1161 BC_1F_2 and F_2 plants from these four fine mapping populations (Table 3.1).

Table 3.1 Fine mapping populations derived from *G. soja* 85-32 that were used for screening of recombinants to delimit the locations of *Rag6* and *Rag3c*

Population	Female Parent	Male Parent	Starting Generation	Number of Lines
110193	E00003¶	E12901*	BC_1F_2	262
110201	E09014	E12901	F_2	396
110202-1	E09088	E12901	F_2	321
110202-2	E12901	E09088	F_2	182

¶ E00003 was the recurrent parent for population 110193
* E12901 is the resistant parent that was derived from E00003 X (G. soja 85-32 X Jiyu71)

Soybean aphid resistance bioassay

Greenhouse trials were performed in the fall of 2012, 2013, 2014 and the spring of 2013 and 2015 in the Plant Science Greenhouse of Michigan State University (MSU) in East Lansing, Michigan. The greenhouse was maintained at 26/15 °C day/night. Sodium vapor lights were applied to supplement light intensity during the day for 14 hours. In the greenhouse trials, eight seeds per line were planted in a plastic pot that was 105 mm in diameter and 125 mm deep. Field trials were conducted on the Agronomy Farm of MSU in East Lansing, Michigan in the summers of 2013, 2014, 2015 and 2016. Fifteen seeds of each line were planted in a single row that was 30 cm long with a row spacing of 60 cm inside a 12.8 x 19.5 m aphid / predator-proof polypropylene cage (Redwood Empire Awning Co., Santa Rosa, CA, USA). Soybean lines from the fine mapping populations were randomly arranged in the greenhouse and field trials without replication.

In the greenhouse and field-cage trials, each plant was inoculated at the V_2 stage (Fehr and Caviness 1977) with two wingless aphids. As shown in the initial mapping study by Zhang et al. (2017), *G. soja* 85-32 possesses strong and broad resistance to mixed Michigan biotypes that have overcome the resistance provided by 'Dowling', 'Jackson', PI 200538 and PI 243540. Therefore, the aphids used to infest plants in the present fine mapping study were mixed biotypes collected from the same locations as in the initial mapping study (Zhang et al. 2017) during the early summer of each testing year. Indicator lines ('Dowling' and PI 243540) were included as checks in the field-cage trials and they were readily colonized by mixed Michigan aphids, indicating Michigan aphid populations were primarily comprised of Biotype 3 along with possible Biotype 4 and/or unknown biotype(s). Aphid resistance was visually rated for each plant with a scale of 0 to 4 (with increments of 0.5) when the susceptible checks reached a rating of 3.0 (usually three weeks after the initial infestation); the higher score indicates heavier infestation (Mensah et al. 2005). The fine mapping analysis used an aphid damage index (DI) for each line, calculated as DI (%) = $\sum$ (rating value x no. of plants in the category) / (4 x total no. of plants) x 100 (Mensah et al. 2005). The DI (%) ranged between 0 for no aphid infestation, and 100 for the most severe aphid damage (Mensah et al. 2005, 2008).

High-throughput DNA isolation with 96-well plates

Before aphid infestation, tissue samples were collected from a non-expanded trifoliolate leaf of each plant, and placed in 1.0 mL individual wells of 96-well plates (USA Scientific, Irvine, CA). Freeze-dried tissue samples were ground with four 4 mm glass beads (Fisher Scientific, Pittsburgh, PA) per well on a modified paint shaker made by Radia (Plymouth, MN) for 2 minutes. To allow ground tissue settle to the bottom of the wells, plates were centrifuged at 3400 rpm for 15 minutes before opening the plate caps (Greiner Bio-One, Kremsmünster, Austria). A

CTAB based DNA extraction buffer (buffer A + buffer B, 200 µL) (Kisha et al. 1997) was added to each well and plates were vortexed for 30 seconds to mix the tissue sample with the buffer. After vortexing, plates were placed in a water-bath set at 65°C for 15 minutes. Plates were centrifuged at 3400 rpm for 1 minute before adding 200 µl chloroform : isoamyl alcohol (24:1) to each well at room temperature. After vortexing the plates slightly to mix the solution, plates were centrifuged at 3400 rpm for 15 minutes. 100 µL of supernatant from each well was transferred to a new 0.5 mL 96-well plate (USA Scientific, scientific, Irvine, CA). 200 µL of chilled (-20 °C) ethanol (95%) was added to each well to precipitate DNA. After the centrifugation at 3400 rpm for 15 minutes, plates were quickly drained and 100 µL of room-temperature 70% ethanol was added to each well to wash the DNA pellets. After centrifuging the plates at 3400 rpm for 5 minutes, plates were quickly drained and air-dried for 30-45 minutes under the fume hood. DNA pellets were re-hydrated overnight with 100 µL of ddH_2O in a 4 °C refrigerator. DNA concentration of each sample was determined with a ND-1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA) and was normalized to 20-150 ng/ µL for Taqman® or KASPTM SNP genotyping reactions.

Screening of recombinants and marker development

To be conservative, the BC_1F_2 and F_2 recombinants-screening started with markers flanking the larger intervals of *Rag6* (Gm08-3, Gm08-28) and *Rag3c* (Gm16-1, Gm16-9) suggested by the validation populations in the initial mapping study (Zhang et al. 2017). The Taqman® and $KASP_{TM}$ SNP genotyping reactions were performed as described by Zhang et al. (2014) and Zhang et al. (2017). Although no interaction has been detected between *Rag6* and *Rag3c* (Zhang et al. 2017), it was found that the presence of one could confound the recombinant analysis of the

other. Therefore, F_2 individuals with recombination events in the *Rag6* region but with the susceptible genotype of *Rag3c* were selected as *Rag6*-recombinants; vice versa, F_2 individuals with recombination events in the *Rag3c* region but with the susceptible genotype of *Rag6* were selected as *Rag3c*-recombinants. F_2 individuals with the heterozygous genotype of *Rag6* or *Rag3c*, but with the susceptible genotype of the other *Rag* QTL, were also selected to be screened for new recombination events of interest in their progenies. A total of 116 F_2 individuals (including recombinants and heterozygotes of interest) were selected and their progenies (2,479 F_3 recombinant plants) were genotyped and phenotyped in the 2013 spring greenhouse trial (Table 3.4a). Additionally, genomic DNA of nine critical *Rag6*-recombinants were isolated with a modified CTAB protocol (Kisha et al. 1997), and genotyped with the Illumina Infinium SoySNP50K iSelect BeadChip (Illumina, San Diego, USA) (Song et al. 2013). Starting with the $F_{2:3}$ generation, the screening for recombinants of *Rag6* or *Rag3c* at each generation were conducted separately (Table 3.4a).

By comparing phenotypic data with genotypic data of each recombinant line, the genomic intervals of *Rag6* and *Rag3c* were gradually delimited over generations. At each generation, lines were screened with markers flanking the newly refined regions of *Rag6* and *Rag3c* (Table 3.4a). Recombinant and heterozygous lines of interest were then tested with additional SNP markers (Table 3.8a) that were within the newly refined regions of interest. These SNPs were discovered from whole genome re-sequencing data of E12901 (Bales 2013) or mined from the Illumina Infinium SoySNP50K iSelect BeadChip. Flanking sequences of these SNPs were obtained from the soybean reference genome, Glyma.Wm82.a1(Schmutz et al. 2010). Progenies of each line were assayed for aphid resistance in the next season's greenhouse or field-cage trial. At the F_8 generation, genomic DNA of twelve critical *Rag6*-recombinant lines were isolated with the

modified CTAB protocol (Kisha et al. 1997), and genotyped with the customized Illumina Infinium SoySNP8K iSelect BeadChip that is a subset of ~ 7000 SNPs from SoySNP50K with additional SNPs discovered from the whole genome-resequencing of aphid-resistant soybean genotypes, including E12901 (Bales 2013). In the delimitation analyses of *Rag6* and *Rag3c*, genetic associations between a segregating genetic marker and the aphid-resistance phenotypic data of recombinant lines were analyzed using the PROC GLM function in SAS9.4 (SAS Institute, Cary NC).

Rag6 validation with a $F_{10:12}$ residual heterozygous family

A F_{10} residual heterozygous line of *Rag6* was identified in the 2015 field-cage trial and developed into a $F_{10:12}$ residual heterozygous family. Genomic DNA of each line in this family was isolated with the 96-well plate CTAB protocol described earlier. The whole family was genotyped with eleven SNP markers (Table 3.8a) that cover ± 1Mb of the fine-mapped region of *Rag6* in spring 2016. A total of 201 $F_{10:12}$ lines from this residual heterozygous family were evaluated for aphid resistance (DI, %) in the 2016 field-cage trial. Phenotypic and genotypic data of each line from this family were input into the QTL Cartographer V2.5 (Wang et al. 2012) using physical positions obtained from Glyma.Wm82.a2 (Grant et al. 2010) as the map positions of each SNP marker; the forward and backward regression method was used in the composite interval mapping analysis. The LOD threshold was statistically determined from 1000 permutations at a significance level of 0.05. The physical map and the LOD plot were drawn with MapChart 2.2 (Voorrips. 2002).

Assessment of flanking markers of Rag6 and Rag3c in breeding populations

To assess the effectiveness of flanking markers in assisting selections for soybean lines integrated with *Rag6* and/or *Rag3c*, two breeding populations (130103 and 130170) with a shared genetic background that is different from that of any fine mapping population were evaluated for aphid resistance in the 2016 field-cage trial and genotyped with two markers flanking the fine-mapped *Rag6* or *Rag3c.* Breeding population 130103, consisting of 156 $F_{2:3}$ individuals, was derived from a cross between E13918 (carrying *Rag6* and *Rag3c*) and E07051 (possessing soybean cyst nematode resistance, phytophthora root and stem rot resistance, high yielding traits). Breeding population 130170, consisting of 502 $F_{2:3}$ individuals, was derived from a cross between E14902 (carrying *Rag6* and *Rag3c*) and E07051. E13918 and E14902 are two homozygous sibling lines from the cross between E09088 and E12901. E07051 is an advanced homozygous breeding line derived from IA3017 (Bilyeu et al. 2006) x Loda (Nickell et al. 2001). The phenotyping and genotyping process of these two breeding populations were the same as described earlier. Individuals in each breeding population were classified into different genotypic groups based on the alleles of the flanking markers. Distinct genotypes were defined by the presence or absence of the allele from E12901 for flanking markers of *Rag6/Rag3c.* Genotypes with corresponding aphid resistance phenotype data were analyzed with one-way analysis of variance (ANOVA) and paired-wise comparisons using the PROC GLM function in SAS9.4 (SAS Institute, Cary NC).

Structural variants identification through whole genome exome-capture sequencing of E12901

Leaf tissue was collected from young soybean seedlings of E12901, and DNA isolation was performed with the modified CTAB method (Kisha et al. 1997). Fragmented DNA with a peak of 150 to 200 bp long was used to prepare a DNA library with an Illumina TruSeq kit. The library was then hybridized with the NimbleGen SeqCap oligo pool (Roche NimbleGen, Madison, WI) designed to capture and enrich targeted DNA fragments, according to Glyma.Wm82.a1v1 (Schmutz et al. 2010). The enriched DNA fragments were amplified and sequenced on Illumina HiSeq for 2 x 100bp paired end reads. Sequence reads for each sample were quality-trimmed and then mapped to the reference genome Glyma.Wm82.a1 (Schmutz et al. 2010). After sequence reads were mapped to the reference genome, SNPs and insertion/deletions (INDELs) were called using the probabilistic model with CLC Genomics Workbench 7.02, then further filtered with a minimum frequency of 20 %, a minimum average base quality 30. For heterozygous SNP/INDEL, a minimum coverage was 6. The rest had a minimum coverage of 3. Variant annotation was performed using snpEff v3.3 (Cingolani et al. 2012).

Results

Fine mapping delimited Rag6 to a 49-kb interval

At the $F_{2:3}$ generation, nine critical *Rag6* recombinant lines were identified with KASP^{TM} and/or Taqman® single SNP genotyping assays and further genotyped with the Illumina Infinium SoySNP50K iSelect BeadChip (Figure 3.1A). Progeny tests of these nine recombinant lines were conducted in the 2013 field-cage trial and some of them were retested in the following seasons (Figure 3.1A). As shown in Figure 3.1A, the associations between the susceptible phenotype and

the susceptible genotype of five recombinant lines, R6-1 to R6-5, defined the bottommost border of *Rag6* as marker Gm08-9. This conclusion was supported by four aphid-resistant lines, R6-6 to R6-9, with resistant genotype above marker Gm08-9 (Figure 3.1A).

Additionally, nineteen recombinant lines (R6-10 to R6-28) were identified with twelve SNP markers at the *Rag6* region (Table 3.2). The strong associations between marker Gm08-16 and the segregating phenotypes in the progenies of lines R6-13 and R6-14 suggested *Rag6* was to the right of marker Gm08-12 (Table 3.2). In addition, no significant association was observed between the segregating marker Gm08-12 and the susceptible phenotype of line R6-12 (Table 3.2), supporting that the left border of *Rag6* was at marker Gm08-12. The significant association between the segregation of aphid resistance and the tested segregating marker in each of the six lines (R6-19, R6-22, R6-23, R6-24, R6-27 and R6-28) defined the right border of *Rag6* as marker Gm08-19 (Table 3.2). Therefore, *Rag6* was delimited to a 390-kb interval between markers Gm08-12 and Gm08-19 (41,948,645 - 42,338,179 bp, Glyma.Wm82.a2, hereafter, unless otherwise stated) on chromosome 8. This conclusion was supported by ten additional fixed recombinant lines shown with arrows in Table 3.2.

To further refine *Rag6*, a total of twelve high-generation (F_8) critical recombinant lines were genotyped with the customized Illumina Infinium SoySNP8K iSelect BeadChip in fall, 2014. Phenotype-genotype associations of five recombinant lines, R6-29 to R6-33, suggested *Rag6* was between markers Gm08-6 and Gm08-20 (41,402,338 - 42,448,802 bp), which verified the delimited 390-kb interval of *Rag6* (41,948,645 - 42,338,179 bp) (Figure 3.1B). The susceptible phenotype and the susceptible genotype of line R6-34 suggested *Rag6* was below marker Gm08-18 (Figure 3.1B). Additionally, four lines (R6-35 to R6-38) showed consistent resistance against aphids across 2014 and 2015, and had resistant genotypes below marker Gm08-17. Moreover,

resistant lines R6-39 and R6-40 had the resistant genotypes above marker Gm08-14 (Figure 3.1B). Therefore, *Rag6* was further defined to a 100-kb interval between markers Gm08-14 and Gm08-17 (42,095,417 – 42,195,720 bp) on chromosome 8.

This refined 100-kb interval of *Rag6* was supported by the associations between the genotype and phenotype of eight additional recombinant lines, R6-41 to R6-48, listed in Table 3.2. Out of 1295 *Rag6*-recombiant and -heterozygous lines, only one line (R6-49) was found as having a recombination event in this 100-kb region. Its progenies were heavily colonized by aphids and there was no association between the susceptible phenotype and the segregating marker Gm08-15 ($P = 0.65$, $R^2 = 0.03$), indicating *Rag6* is on the right side of marker Gm08-15 (Table 3.2). Therefore, *Rag6* was delimited to a 49-kb interval between makers Gm08-15 and Gm08-17 (42,146,252 - 42,195,720 bp) on chromosome 8.

Figure 3.1 Graphical representation of chromosome 8 genotypes of critical *Rag6* recombinant lines without *Rag3c*. **A** Genotypes tested on the Illumina Infinium SoySNP50K iSelect BeadChip array in summer 2013. **B** Genotypes tested on the Illumina Infinium SoySNP8K iSelect BeadChip array in fall 2014. Colors in bars denote either homozygous genomic regions inherited from the resistant (black) or susceptible (white) parent, or heterozygous regions (gray). Black hatching indicates the previously mapped region of *Rag6* (40,047,323 bp - 46,037,031 bp) (Zhang et al. 2017) in **A** or the narrowed-down region of *Rag6* (41,948,645 - 42,338,179 bp) in **B**. Delimitation analyses are shown with gray lines and black arrows. **S** represents susceptible phenotype with average DI (%) ranging from 75-100% in the progeny test; MR represents moderate-resistant phenotype with average DI (%) ranging from 37.5-75% in the progeny test; **R** represents resistant phenotype with average DI (%) ranging from 0-37.5% in the progeny test

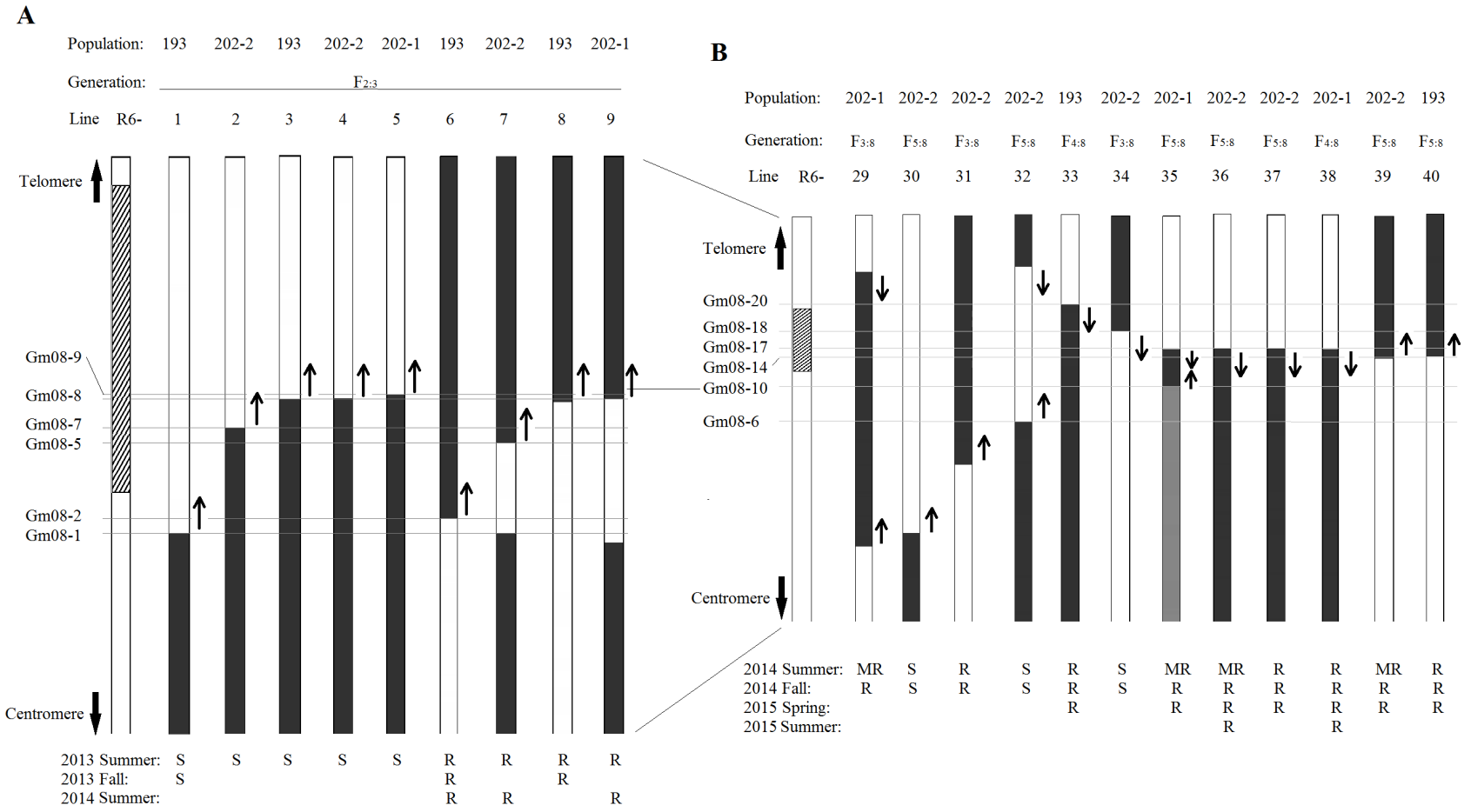


Table 3.2 Progeny test of *Rag6*-recombinant lines tested with KASP™ SNP markers

			KASP™ SNP markers (Gm-) and physical positions (Mb)[a]												Progeny test						
Pop	Line	Gen	08-11	08-12	08-14	08-15	08-16	08-17	08-19	08-21	08-24	08-25	08-26	08-27	2014	2015	2015	2015	Marker[h]	Pr > F[i]	R^{2j}
110-	number		41.716	41.949	42.095	42.146	42.160	42.195	42.338	42.522	43.391	44.129	45.398	45.913	summer	fall	spring	summer			
202-2	R6-10	$F_{5:6}$	r[b]→	s[c]	s	s	s	s	s	s	s	s	s	s	S[e]						
202-1	R6-11	$F_{3:4}$	s	s	s	s	s	s	s	s	s	s	s	←r	S						
202-1	**R6-12**	$F_{5:6}$	h[d]	h→	s	s	s	s	s	s	s	s	s	s	S				08-12	0.1570	0.25
193	**R6-13**	$F_{5:6}$	r	r→	-	-	h	h	h	h	h	h	h	←r	Seg[f]				08-16	<0.0001	0.85
202-2	**R6-14**	$F_{5:6}$	s	s→	-	-	h	h	h	h	h	h	h	←r	Seg				08-16	0.0013	0.7
193	R6-15	$F_{3:4}$	r→	s	s	s	s	s	s	s	s	s	←r	r	S						
202-1	R6-16	$F_{4:5}$	r	r	r	r	r	r	r	r	r	r	←s	s	R[g]	R					
202-2	R6-17	$F_{3:4}$	s	s	s	s	s	s	s	s	s	←r	r	r	S						
202-2	R6-18	$F_{3:4}$	r	r	r	r	r	r	r	r	r	←s	s	s	R	R					
202-2	**R6-19**	$F_{5:6}$	s	s→	-	-	h	h	h	h	h	←r	r	r	Seg				08-21	<0.0001	0.9
202-2	R6-20	$F_{4:5}$	s	s	s	s	s	s	s	s	←r	r	r	r	S	S					
193	R6-21	$F_{4:5}$	r	r	r	r	r	r	r	r	←s	s	s	s	R	R					
202-1	**R6-22**	$F_{5:6}$	h	h	h	h	h	h	h	h	←r	r	r	r	Seg				08-21	0.0003	0.80
202-2	**R6-23**	$F_{5:6}$	h	h	h	h	h	h	h	h	←r	r	r	r	Seg				08-21	0.0072	0.91
202-1	**R6-24**	$F_{5:6}$	h	h	h	h	h	h	h	h	←s	s	s	s	Seg				08-21	<0.0001	0.91
202-2	R6-25	$F_{5:6}$	s	s	s	s	s	s	s	s	←r	r	r	r	S	S					
202-2	R6-26	$F_{5:6}$	r	r	r	r	r	r	←s	s	s	s	s	s	R	R	R	R			
202-1	**R6-27**	$F_{5:6}$	h	h	h	h	h	-	←r	r	r	r	r	r	Seg				08-16	<0.0001	0.82
202-1	**R6-28**	$F_{4:5}$	h	h	h	h	h	-	←s	s	s	s	s	s	Seg				08-16	<0.0001	0.97
193	R6-41	$F_{8:9}$	-	s	s→	r	r	r	r	-	-	-	-	-		R		R			
193	R6-42	$F_{7:8}$	-	s	s→	r	r	r	r	-	-	-	-	-		R	R	R			
202-2	R6-43	$F_{6:7}$	r	s	s→	r	r	r	r	-	-	-	-	-		R	R	R			
202-1	R6-44	$F_{7:8}$	r	r	r→	s	s	s	s	-	-	-	-	-		S	S	S			
202-2	R6-45	$F_{7:8}$	r	r	r→	s	s	s	s	-	-	-	-	-			S	S			
202-1	R6-46	$F_{7:8}$	-	s	s	s	s	←r	r	-	-	-	-	-		S	S	S			
202-1	R6-47	$F_{5:6}$	r	r	r	r	r	←s	s	-	-	-	-	-		R	R	R			
202-2	R6-48	$F_{5:6}$	r	r	r	r	r	←s	s	-	-	-	-	-		R	R	R			
202-1	R6-49	$F_{4:10}$	h	h	h	h→	s	s	s	s	s	s	s	s	S	S	S		08-15	0.65	0.03
193	R6-50	$F_{8:10}$	s	s	s→	h	h	h	←s	s	r	r	r	r			Seg		08-17	0.0003	0.84

[a] Physical positions of markers according to Glyma.Wm82.a2 on SoyBase (Grant et al. 2010)
[b] r represents alleles from the resistant parent
[c] s represents alleles from the susceptible parent
[d] h represents heterozygous alleles

Table 3.2 (cont'd)

[e] S represents susceptible phenotype with average DI (%) ranging from 75-100%

[f] Seg represents segregating phenotypes

[g] R represents resistant phenotype with average DI (%) ranging from 0-37.5%

[h] Marker used in F-test

[i] Significance level of the marker-trait association

[j] R^2 value of the marker-trait association

Fine mapping delimited Rag3c to a 150-kb interval

At the $F_{2:3}$ generation, three lines (R3c-1 to R3c-3) were identified as having recombinant events in the *Rag3c* region (Table 3.3). A strong association between the segregation of aphid resistance and marker Gm16-7 was observed for each of these three lines (Table 3.3), indicating *Rag3c* resides to the left of marker Gm16-8. According to the associations between the susceptible genotype and susceptible phenotype of three recombinant lines (R3c-4 to R3c-6), *Rag3c* was further delimited to the left of marker Gm16-6. This was supported by resistant lines R3c-7 and R3c-8 with the resistant genotype to the left of marker Gm16-6. Furthermore, recombinant line R3c-9 delimited *Rag3c* to the left of marker Gm16-5 in that a strong association was observed between marker Gm16-4 and the phenotypes of the progenies ($P = 0.0020$, $R^2 = 0.75$). The association between the resistant phenotype and the resistant genotype in each of the lines (R3c-10 to R3c-15) suggested the same right border, marker Gm16-5, for *Rag3c*. Additionally, six resistant lines, R3c-16 to R3c-21, suggested *Rag3c* resides to the right of marker Gm16-2. The left border of *Rag3c* was further pushed to marker Gm16-3 by the phenotype-genotype association of line R3c-22. Therefore, *Rag3c* was refined to a 150-kb interval between markers Gm16-3 and Gm16-5 (6,621,540 - 6,771,675 bp) on chromosome 16.

Table 3.3 Progeny test of recombinant lines of *Rag3*c tested with KASP™ SNP markers

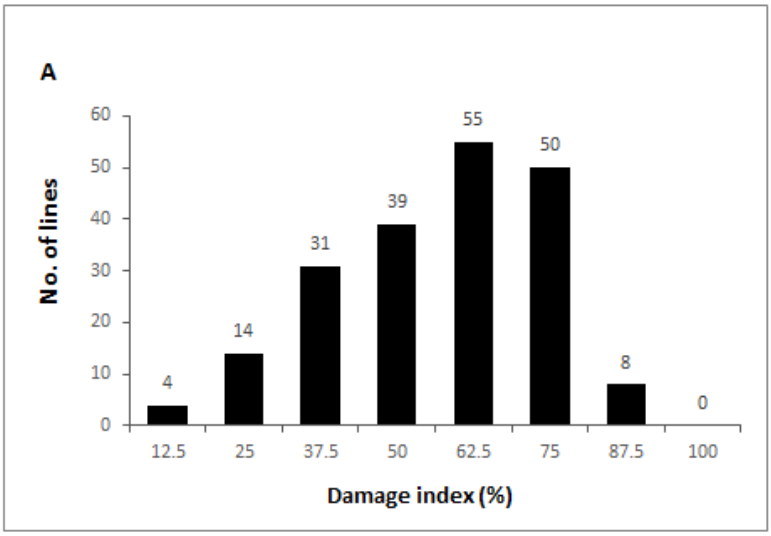
			KASP™ SNP markers (Gm-) and physical positions (Mb)[a]									Progeny test					
Pop	Line	Gen	16-1	16-2	16-3	16-4	16-5	16-6	16-7	16-8	16-9	2013	2013	2014	2014	2015	2015
110-	number		6.314	6.618	6.622	6.657	6.772	6.871	6.884	7.229	8.208	summer	fall	summer	fall	spring	summer
193	R3c-1	$F_{2:3}$	h[b]	h	h	h	h	h	h	←s[c]	s	Seg[e] (P<0.0001, R^2=0.96, 16-7)*					
193	R3c-2	$F_{2:3}$	h	h	h	h	h	h	h	←s	r[d]	Seg (P=0.0022, R^2=0.74, 16-7)*					
193	R3c-3	$F_{2:3}$	h	h	h	h	h	h	h	←s	s	Seg (P=0.0027, R^2=0.91, 16-7)*					
193	R3c-4	$F_{3:4}$	s	s	s	s	s	←r	r	r	r		S[f]	S			
193	R3c-5	$F_{3:4}$	s	s	s	s	s	←r	r	r	-		S	S			
193	R3c-6	$F_{3:4}$	s	s	s	s	s	←r	r	r	r		S	S			
202-1	R3c-7	$F_{7:8}$	r	r	r	r	r	←s	s	s	s			R[g]	R	R	R
202-1	R3c-8	$F_{6:7}$	r	r	r	r	r	←s	s	s	r			R	R	MR[h]	R
193	R3c-9	$F_{2:3}$	h	h	h	h	←s	s	s	s	s	Seg (P=0.0020, R^2=0.75, 16-4)*					
193	R3c-10	$F_{2:3}$	r	r	r	r	←s	s	s	s	s	R	R	MR			
193	R3c-11	$F_{3:4}$	r	r	r	r	←s	s	s	s	-		MR	R	R		
202-1	R3c-12	$F_{5:6}$	r	r	r	r	←s	s	s	s	s			MR	R	MR	R
202-1	R3c-13	$F_{6:7}$	r	r	r	r	←s	s	s	s	s			MR	R	MR	R
202-1	R3c-14	$F_{4:5}$	r	r	r	r	←s	s	s	s	s			MR	R	R	R
202-1	R3c-15	$F_{4:5}$	r	r	r	r	←s	s	s	s	s			MR	R	MR	R
202-1	R3c-16	$F_{3:4}$	s	s→	-	r	r	r	r	r	-		MR	R	R		
202-1	R3c-17	$F_{5:6}$	s	s→	-	r	r	r	r	r	r			MR	R	R	R
202-1	R3c-18	$F_{6:7}$	s	s→	-	r	r	r	r	r	r			MR	R	R	R
202-1	R3c-19	$F_{6:7}$	s	s→	-	r	r	r	r	r	r			MR	R	R	R
202-1	R3c-20	$F_{6:7}$	s	s→	-	r	r	r	r	r	r			R	R	R	R
202-1	R3c-21	$F_{6:7}$	s	s→	-	r	r	r	r	r	r			MR	MR	MR	R
202-2	R3c-22	$F_{5:6}$	s	s	s→	r	r	r	r	r	r			MR	R	R	R

[a] Physical positions of markers according to Glyma.Wm82.a2 on SoyBase (Grant et al. 2010)
[b] h represents heterozygous alleles
[c] s represents alleles from the susceptible parent
[d] r represents alleles from the resistant parent
[e] Seg represents segregating phenotypes
[f] S represents susceptible phenotype with DI (%) ranging from 75-100%
[g] R represents resistant phenotype with DI (%) ranging from 0-37.5%
[h] MR represents moderate resistant phenotype with DI (%) ranging from 37.5-75%
*The association between marker Gm16-7/16-4 and the segregation of aphid resistance in the progenies was significant

Refined Rag6 region was validated with a $F_{10:12}$ residual heterozygous family

A F_{10} line, R6-50, with a 243-kb heterozygous interval (42,095,417 - 42,338,179 bp) was identified in the 2015 field-cage trial (Table 3.2) and developed into a $F_{10:11}$ residual heterozygous family in spring, 2016. A total of 201 $F_{10:12}$ lines of this family were tested with aphids in the 2016 field-cage trial. Aphid-resistance phenotype of the family distributed normally (Figure 3.2A). As suggested in the initial mapping study, *Rag6* has an additive effect (Zhang et al. 2017). When grouping lines with a damage index of 0 - 37.5% as resistant, 37.5 - 75% as moderate resistant, and 75 - 100% as susceptible, these three categories followed a 1:2:1 segregation ratio according to the chi-square test, X^2 (2, N = 201) = 1.481, P = 0.4768 (Figure 3.2B). The entire family was genotyped with eleven SNP markers covering ± 1 Mb of the fine-mapped *Rag6* region. As shown in Figure 3.2C, a significant peak (LOD = 33.4 while LOD threshold was 1.7 from 1000 permutations at a significance level of 0.05) was detected between markers Gm08-15 and Gm08-19 (42,146,252 - 42,338,179 bp), which validated the fine-mapped region of *Rag6* (42,146,252 - 42,195,720 bp). The phenotypic variance explained by this QTL peak was 67.6 %. The additive effect provided by *Rag6* was -18.6, indicating the *Rag6* allele helps reduce the aphid damage index (%) by 18.6 %.

Figure 3.2 *Rag6* validation with a $F_{10:12}$ residual heterozygous family. **A** Continuously phenotypic distribution of aphid damage index (%). **B** Discrete phenotypic distribution of aphid resistance. R means lines with 0-37.5% damage index. H means lines with 37.5-75% damage index. S means lines with 75-100% damage index. **C** QTL detection of *Rag6* using composite interval mapping method. Marker positions (Mb) are physical positions according to Glyma.Wm82.a2 on SoyBase (Grant et al. 2010). Damage index (%) = ∑ (scale value X no. of plants in the category) / (4 X total no. of plants) X 100 (Mensah et al. 2005)



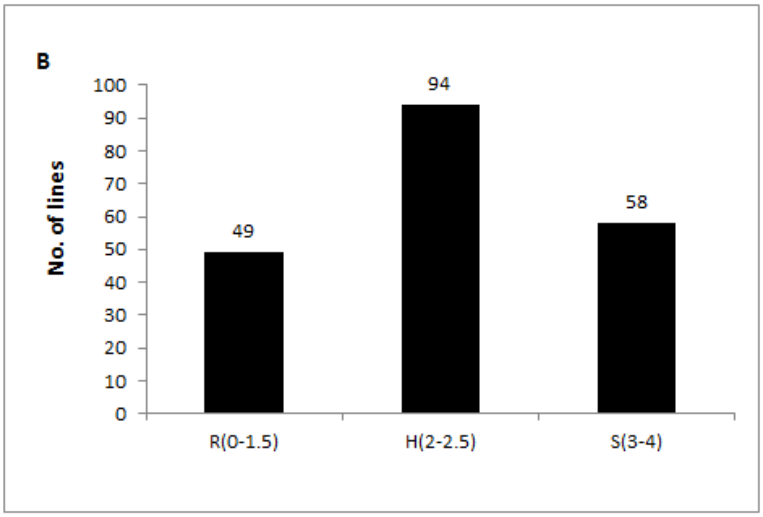
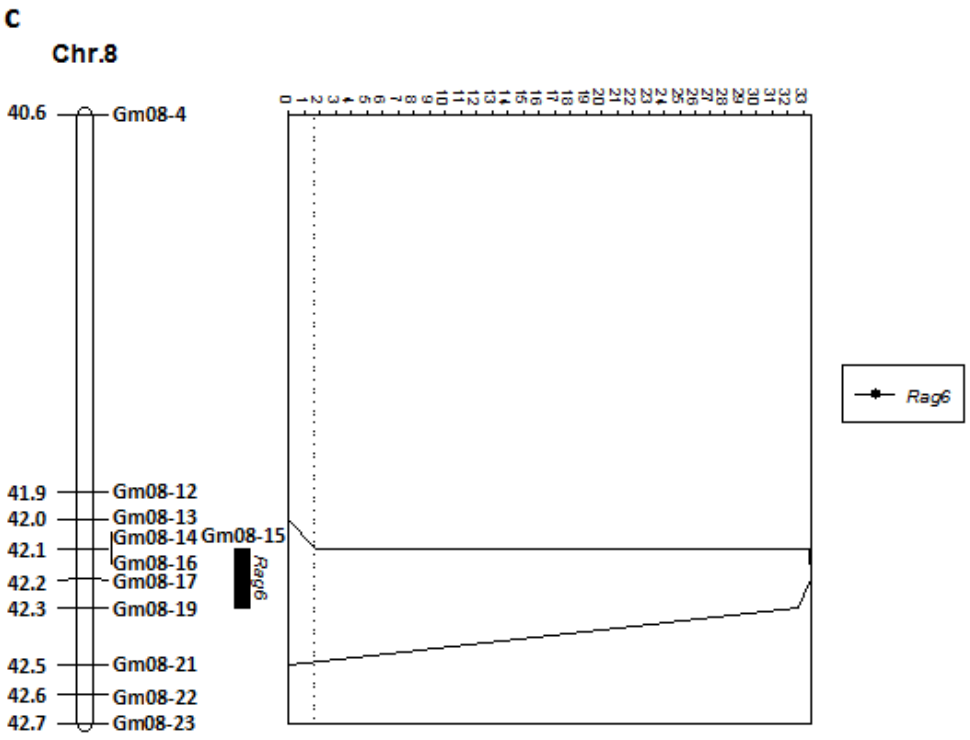


Figure 3.2 (cont'd)



Flanking markers of the refined regions were demonstrated robust in assisting selection for Rag6 and/or Rag3c in breeding populations

Breeding population 130103, consisting of 156 $F_{2:3}$ individuals, and breeding population 130170, consisting of 502 $F_{2:3}$ individuals, were genotyped with KASP™ SNP markers that flank *Rag6* and *Rag3c* fine-mapped regions, including Gm08-15 (42,146,252 bp), Gm08-17 (42,195,720 bp), Gm16-2 (6,617,689 bp) and Gm16-5 (6,771,675 bp). A total of 328 individuals were grouped into four distinct homozygous genotypes by the presence or absence of the allele from E12901 for these flanking markers (Table 3.5a). Individuals with ambiguous or missing genotype data were excluded from the analysis. From the one-way ANOVA and pair-wise comparison analyses of these four genotype groups, the LSMEANs of the rating score of each genotype group were significantly different (P <0.05), with -/- having the highest damage score

and *Rag6/Rag3c* having the lowest damage score in each breeding population (Figure 3.3A, B). Between the genotypes with only one of the *Rag* genes from *G. soja* 85-32, genotypes with *Rag6* showed significantly lower aphid damage than genotypes with *Rag3c*, which is consistent with the observation of *Rag6* conferring a stronger resistance in the previous initial mapping study (Zhang et al. 2017). The strong associations between the flanking markers (Gm08-15, Gm08-17, Gm16-2, and Gm16-5) and aphid damage indices demonstrated the robustness of these markers in assisting selections of *Rag6* and/or *Rag3c* under different genetic backgrounds.

Figure 3.3 Efficiency of marker-assisted selection for *Rag6* and/or *Rag3c* in breeding populations 130103 (**A**) and 130170 (**B**). Four distinct genotypes were defined by the presence or absence of the allele from the resistant parent (E12901) for the flanking markers Gm08-15, Gm08-17, Gm16-2, and Gm16-5; - / - represents genotypes carrying susceptible alleles of *Rag6* and *Rag3c*, - / *Rag3c* represents genotypes carrying susceptible alleles of *Rag6* but resistant alleles of *Rag3c*, *Rag6* / - represents genotypes carrying resistant alleles of *Rag6* but susceptible alleles of *Rag3c* and *Rag6* / *Rag3c* represents genotypes carrying resistant alleles of *Rag6* and *Rag3c.* Bars with different letter are significantly different at $P < 0.05$

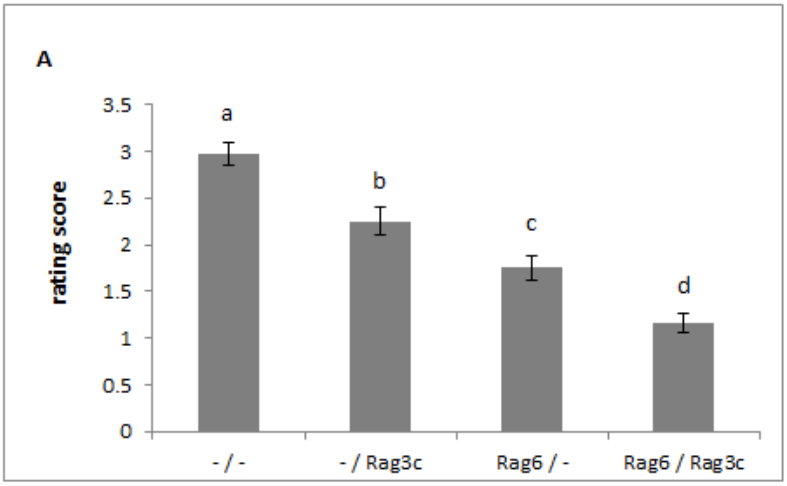
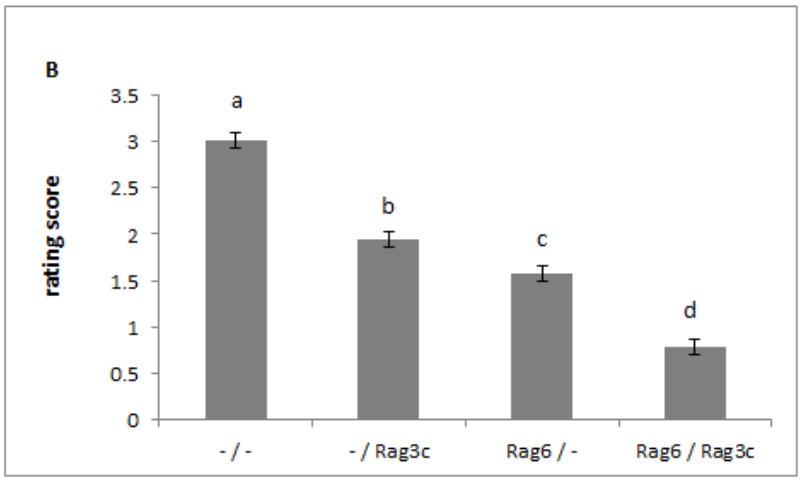


Figure 3.3 (cont'd)



Structural variant analysis uncovered high-confidence candidate genes

According to Glyma.Wm82.a2.v1 on SoyBase (Grant et al. 2010), four candidate genes are present in the 49-kb interval (42,146,252 - 42,195,720 bp) of *Rag6*, including three clustered NBS-LRR genes (Glyma.08g303500, Glyma.08g303600, and Glyma.08g303700) and one amine oxidase gene (Glyma.08g303800). DNA sequence variations were detected within these candidate genes by comparing the reads from the whole genome exome-capture sequencing of E12901 to the reference genome Glyma.Wm82.a1(Schmutz et al. 2010). Variants with moderate to high effects are listed in Table 3.6a. Two frame shifts caused by deletions, along with a codon deletion and a non-synonymous coding change, were detected in Glyma.08g303500. Due to the absence/unavailability of Glyma.08g303600 in Glyma.Wm82.a1v1, which was used to design the bait sequences for whole genome exome pull-down of soybean, the DNA sequence variants of Glyma.08g303600 are unknown. A total of six non-synonymous coding changes with

moderate effects were detected in Glyma.08g303700. Besides a few synonymous coding changes, no variations in the exons of Glyma.08g303800 were detected that can change the protein product.

As shown in Table 3.7a, eleven candidate genes are present in the 150-kb interval (6,621,540 - 6,771,675 bp) of *Rag3c* based on Glyma.Wm82.a2.v1 (Grant et al. 2010). Of these eleven candidate genes, there is only one NBS-LRR gene (Glyma.16g066800); however, no variations were detected in the exons of this gene. A total of thirteen variations, including a frame shift, were detected in the exons of Glyma.16g066900, which is annotated as encoding lipase. Additionally, a few structural variations were discovered in five other candidate genes (Glyma.16g067000, Glyma.16g067200, Glyma.16g067500, Glyma.16g067800 and Glyma.16g068100) that are unclear in defense mechanisms (Table 3.7a). A total of 185 SNPs and INDELs in *Rag6* and *Rag3c* fine-mapped regions are summarized in Table 3.9.

Discussion

In this study, two aphid-resistance genes, *Rag6* and *Rag3c*, were fine-mapped to a 49-kb interval (42,146,252 - 42,195,720 bp) and 150-kb interval (6,621,540 - 6,771,675 bp) on chromosome 8 and chromosome 16, respectively. The availability of the Illumina Infinium SoySNP50K/8K iSelect BeadChip, combined with the SNPs discovered through the whole genome re-sequencing of E12901, facilitated the fine mapping process and the development of robust SNP markers in assisting selections of *Rag6* and/or *Rag3c*. Detailed information of all the SNP markers used in this study are summarized in Table 3.8a.

The antibiosis resistance gene, *Rag6* from *G. soja* 85-32, discovered by Zhang et al. (2017) was refined to a 49-kb interval in the present study that does not overlap with any other aphid-

resistance loci identified on chromosome 8 (Meng et al. 2011; Jun et al. 2012; Xiao et al. 2013). Therefore, *Rag6* is a novel gene that can provide additional resistance against soybean aphids. Although multiple aphid-resistance genes (*Rag3*, *Rag3b*, *Rag3c*, and *rag3*) were detected in a shared region (Zhang et al. 2010, 2013; Bales et al. 2013; Zhang et al. 2017), *Rag3c* confers a lower aphid-resistance level and is located at a different fine-mapped region than the others (Unpublished dissertations by Zhang 2012, Bales 2013, and Du 2016). According to the genome annotation of Williams 82, this shared region is enriched with NBS-LRR genes; it may explain why multiple aphid-resistance genes were mapped to this region.

By mapping the reads from the whole genome exome-capture sequencing of E12901 (derived from *G. soja* 85-32) to the aphid-susceptible reference genome, Glyma.Wm82.a1.v1, DNA sequence variations were detected within the regions of interest. Multiple structural variants (two frame shifts, two deletions and a non-synonymous coding change) were detected in Glyma.08g303500 that is highly homologous with the gene AT5G45520.1 encoding LRR family protein in *Arabidopsis thaliana* (Zybailov et al. 2008). These structural variants might change the final protein product of Glyma.08g303500 and lead to the resistance phenotype. It is also possible that Glyma.08g303500 is a susceptibility gene with NBS-LRR domains as a few NBS-LRR genes in plants have been reported conferring sensitivity to pathogens (Lorang et al. 2007; Nagy and Bennetzen 2008; Faris et al. 2010); the disrupted protein produced by Glyma.08g303500 in *G. soja* 85-32 might lose the function of sensitivity to aphids. The homologous gene (AT5G46450.1) of Glyma.08g303600 has been reported as encoding disease resistance proteins (TIR-NBS-LRR family) that confer resistance against cabbage leaf curl virus and *Pseudomonas syringae pv. tomato* in Arabidopsis (Mohr et al. 2007; Ascencio-Ibáñez et al. 2008). Tan et al. (2007) also reported that the expression level of AT5G46450.1 was altered after

flagellin or salicylic acid treatment. Due to the absence/unavailability of Glyma.08g303600 in Glyma.Wm82.a1v1, which was used to design the bait sequence for exome-capture of E12901, structural variants within this gene are unknown. However, possible variants may exist in Glyma.08g303600 and result in the aphid-resistance provided by *Rag6*. The Arabidopsis gene, AT5G17680.1, is highly homologous to Glyma.08g303700 and was predicted as encoding disease resistance protein (TIR-NBS-LRR family). Interestingly, Yu et al. (2015) reported that eight out of ten NBS-LRR genes in the fine-mapped region of the Asian soybean rust resistance gene *Rpp2* are highly homologous with AT5G17680.1. This suggests Glyma.08g303700 might play a role in defense to soybean aphid and its multiple non-synonymous coding changes in *G. soja* 85-32 may lead to the resistance phenotype. The homolog of Glyma.08g303800 (which encodes amine oxidase) in *A. thaliana* is AT2G43020.1 with multiple molecular functions, including reducing reactive oxygen species production and increasing defense gene expression (Ascencio-Ibáñez et al. 2008; Sagor et al. 2016). No effective variations in the exons of Glyma.08g303800 were detected, however, structural variation(s) might occur at the promoter or UTR regions that can change the expression of this gene. Therefore, the three NBS-LRR genes (Glyma.08g303500, Glyma.08g303600 and Glyma.08g303700) and the amine oxidase gene (Glyma.08g303800) present in the *Rag6* interval are important for future investigations.

The best candidate genes for *Rag3c* are Glyma.16g066800 and Glyma.16g066900 that encode NBS-LRR protein kinase and lipase, respectively. Although no structural variants were detected in the exons of Glyma.16g066800, variation(s) may exist in the promoter or UTR region(s) and lead to the moderate resistance. Multiple structural variants, including a frame shift, were detected within Glyma.16g066900, of which the homolog gene in Arabidopsis is the lipase-encoding gene AT1G53920.1 (Oh et al. 2005). Interestingly, most members in the lipase gene

family are required for full local and systemic resistance against pathogen infection in plants (Kumar and Klessig 2003; Jakab et al. 2003; Oh et al. 2005). Jasmonic acids (JAs) are well known as lipid-based hormone signals that are critical for plant defense against herbivory (Paré and Tumlinson 1999). Studham and MacIntosh (2013) reported that soybean aphid colonization leads to a decrease of poly-unsaturated fatty acids, which is used by soybean plants for JA biosynthesis. JA signaling triggered by aphid infestation plays a critical role in regulating plant defense (Thompson et al. 2006). Li et al. (2008) also suggested that aphid-feeding partially activates JA-regulated signaling pathways in soybean defense. Thus, besides the NSB-LRR gene (Glyma.16g066800), it is possible that the moderate resistance of *Rag3c* is contributed by the lipase gene, Glyma.16g066900.

It is well-known that NBS-LRR genes are present in clusters as the result of gene duplication and recombination (Martin et al. 2003; Leister 2004). Clusters of NBS-LRR genes may confer resistance to different pathogens or to different races of the same pathogen (Leister 2004). Three NBS-LRR genes are clustered within the 49kb-interval of *Rag6* (Glyma.08g303500, Glyma.08g303600 and Glyma.08g303700) and all share the core sequence information as tandem repeats. In addition to the possibility of one or two NBS-LRR gene(s) governing *Rag6* resistance, it is possible that all three NBS-LRR genes are needed for the *Rag6*-mediated resistance. It is also possible that the durable resistance of *Rag6* is achieved by different NBS-LRR genes conquering different soybean aphid biotypes. Interestingly, a QTL resistant to *Sclerotinia sclerotiorum* was mapped in PI 391589B near marker Satt209 (42,190,808 bp) (Guo et al. 2008), which is in close proximity to the NBS-LRR cluster of *Rag6*; however, PI 391589B is susceptible to soybean aphids (Mensah et al. 2005). According to the field data, *G. soja* 85-32 is also resistant to *S. sclerotiorum* (unpublished data). This NBS-LRR gene cluster in *G. soja* 85-

32 may confer resistance to both soybean aphids and *S. sclerotiorum.* Markers closely linked to this NBS-LRR cluster (e.g. Gm08-15, Gm08-17 and Satt209) can be screened for associations with resistance against *S. sclerotiorum* in a segregating population derived from *G. soja 85-32* to test this hypothesis in the future.

Detected DNA structural variants (Table 3.6a and Table 3.7a) within the fine-mapped regions of *Rag6* and *Rag3c* may explain the resistance phenotype delivered by *G. soja* 85-32. However, candidate gene predictions for *Rag6* and *Rag3c* are based on the reference genome, Williams 82, which is susceptible to soybean aphids. Therefore, the resistance phenotype might be attributed to a different copy number of the NBS-LRR gene(s) or some unknown gene(s) exclusively existing within the fine-mapped regions of *Rag6* and *Rag3c* in the *G. soja* 85-32 genome. *De novo* assembly of longer reads from deep re-sequencing (e.g. PacBio) of *G. soja 85-32* would be helpful to validate the structure variants detected in this study and uncover possible gene(s) or copy number variance exclusively present in the regions of interest in the *G. soja 85-32* genome.

The breakdown of aphid-resistance genes by emerging new biotypes in North America (Hill et al. 2004a; Kim et al. 2008; Hill et al. 2010; Alt and Ryan-Mahmutagic 2013) increases the need to identify new resistance genes and the need of gene pyramiding in commercial soybean varieties to achieve durable resistance to aphids. The closely-linked SNP markers identified in the present study were proved robust in assisting selections for *Rag6* and/or *Rag3c* under different genetic backgrounds. They will facilitate marker-assisted selections for aphid resistance from *G. soja* 85-32 and gene pyramiding with minimum negative linkage drags. Additionally, the predicted candidate genes and the identified structural variations would help expedite future functional studies and map-based cloning efforts.

APPENDIX

APPENDIX

Table 3.4a Number of plants tested with markers to identify recombination events in the *Rag6* and *Rag3c* intervals each generation and the number of plants selected

	Rag6			*Rag3c*		
Gen	Tested plants[a]	Flanking markers	Selected plants[b]	Tested plants[a]	Flanking markers	Selected plants[b]
F_2	1161[c]	Gm08-3, Gm08-28	58/9	1161[c]	Gm16-1, Gm16-9	36/13
F_3	1533	Gm08-11, Gm08-27	124/80	946	Gm16-1, Gm16-8	18/75
F_4	1780	Gm08-12, Gm08-26	21/10	797	Gm16-2, Gm16-7	25/5
F_5	1277	Gm08-12, Gm08-25	44/159	673	Gm16-2, Gm16-6	96/67
F_6	1570	Gm08-12, Gm08-24	111/156	1140	Gm16-2, Gm16-5	89/58
F_7	2522	Gm08-14, Gm08-21	93/55	1388	Gm16-2, Gm16-5	76/9
F_8	1428	Gm08-14, Gm08-19	72/81	694	Gm16-3, Gm16-5	76/6
F_9	1209	Gm08-14, Gm08-17	115/34	708	Gm16-3, Gm16-5	92/23
F_{10}	1295	Gm08-15, Gm08-16	1	911	Gm16-3, Gm16-5	-

[a] At higher generations, among the tested plants, many of them shared pedigree because they were derived from same heterozygous recombinants. Additionally, some of the tested plants were recombinants that were carried over from the previous season to confirm their phenotype and genotype

[b] Selected plants included recombinants and heterozygotes, indicated by recombinants/heterozygotes. Some recombinant selections were carried over to the following seasons to confirm their phenotype and genotype

[c] Recombinants screening for *Rag6* and *Rag3c* started with 1161 F_2 plants. Starting with the F_3 generation, screening for recombinants of *Rag6* and *Rag3c* were separated as no interaction had been detected between *Rag6* and *Rag3c*

Table 3.5a Effectiveness of flanking markers in assisting selection for *Rag6* and *Rag3c* in breeding population 130103 and 130170

			Flanking markers[a]					
Pop	Genotype	No. of lines					Mean	Standard error
			Rag6		*Rag3c*			
			Gm08-15	Gm08-17	Gm16-2	Gm16-5		
130103	-/-	21	-	-	-	-	2.98	0.12
	-/*Rag3c*	15	-	-	+	+	2.25	0.15
	Rag6/-	18	+	+	-	-	1.75	0.13
	Rag6/Rag3c	27	+	+	+	+	1.17	0.11
130170	-/-	59	-	-	-	-	3.01	0.08
	-/*Rag3c*	68	-	-	+	+	1.94	0.07
	Rag6/-	64	+	+	-	-	1.58	0.08
	Rag6/Rag3c	56	+	+	+	+	0.79	0.08

[a] + Implies allele from E12901; - Implies allele from the susceptible parent

Table 3.6a List of annotated gene models within the interval of *Rag6* and structural variants with moderate or high effects

Gene model (a2.v1)	Physical position of gene (bp) (a2.v1)	Gene model (a1.v1)	Physical position of gene (bp) (a1.v1)	Physical position of variant (bp) (a1.v1)	Glyma.Wm82.a1	E12901	DNA sequence variations	Annotation (a2.v1)	*A.thaliana* homolog (annotation)
Glyma.08g303500	42150074 - 42151723	Glyma08g41550	41508240 - 41509192	41509151	ACGT	A	Codon deletion	LRR-containing protein	AT5G45520.1 (LRR family protein)
Glyma.08g303500	42150074 - 42151723	Glyma08g41550	41508240 - 41509192	41509156	GA	G	Frame shift*	LRR-containing protein	AT5G45520.1 (LRR family protein)
Glyma.08g303500	42150074 - 42151723	Glyma08g41550	41508240 - 41509192	41509158	TACCCAAT A	T	Deletion, frame shift*	LRR-containing protein	AT5G45520.1 (LRR family protein)
Glyma.08g303500	42150074 - 42151723	Glyma08g41550	41508240 - 41509192	41509173	G	T	Non-synonymous coding change	LRR-containing protein	AT5G45520.1 (LRR family protein)
Glyma.08g303600	42160184 - 42161884	-	-	-	-	-	-	LRR-containing protein	AT5G46450.1(disease resistance protein, TIR-NBS-LRR family)
Glyma.08g303700	42162347 - 42167295	Glyma08g41560	41520515 - 41525323	41521312	TG	GA	Non-synonymous coding change	LRR-containing protein	AT5G17680.1 (disease resistance protein, TIR-NBS-LRR family)
Glyma.08g303700	42162347 - 42167295	Glyma08g41560	41520515 - 41525323	41521322	C	T	Non-synonymous coding change	LRR-containing protein	AT5G17680.1 (disease resistance protein, TIR-NBS-LRR family)
Glyma.08g303700	42162347 - 42167295	Glyma08g41560	41520515- 41525323	41521326	C	A	Non-synonymous coding change	LRR-containing protein	AT5G17680.1 (disease resistance protein, TIR-NBS-LRR family)
Glyma.08g303700	42162347 - 42167295	Glyma08g41560	41520515 - 41525323	41521466	G	T	Non-synonymous coding change	LRR-containing protein	AT5G17680.1 (disease resistance protein, TIR-NBS-LRR family)
Glyma.08g303700	42162347 - 42167295	Glyma08g41560	41520515 - 41525323	41523569	T	A	Non-synonymous coding change	LRR-containing protein	AT5G17680.1 (disease resistance protein, TIR-NBS-LRR family)
Glyma.08g303700	42162347 - 42167295	Glyma08g41560	41520515 - 41525323	41523617	A	C	Non-synonymous coding change	LRR-containing protein	AT5G17680.1 (disease resistance protein, TIR-NBS-LRR family)
Glyma.08g303800	42174238 - 42179859	Glyma08g41570	41532406 - 41538027	-	-	-	-	Amine oxidase	AT2G43020.1 (polyamine oxidase 2)

* Mutations with high effects

Pull-down sequences /primer sequences were designed based on Glyma.Wm82.a1v1

Table 3.7a List of annotated gene models within the interval of *Rag3c* and structural variants with moderate or high effects

Gene model (a2.v1)	Physical position of gene (bp) (a2.v1)	Gene model (a1.v1)	Physical position of gene (bp) (a1.v1)	Physical position of mutation (bp) (a1.v1)	Glyma.Wm82.a1	E12901	DNA sequence variations	Annotation (a2.v1)	*A. thaliana* homolog (annotation)
Glyma.16g066700	6620330 - 6621988	Glyma16g07200	6473613 - 6475030	-	-	-	-	Ubiquitin	AT5G14360.1 (Ubiquitin-like superfamily protein)
Glyma.16g066800	6627025 - 6628243	Glyma16g07220	6480168 - 6481544	-	-	-	-	LRR protein kinase	AT1G58190.2 (LRR)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6490128	A	G	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6490135	T	G	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6490341	G	A	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6491630	A	C	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492005	G	A	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492389	G	T	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492399	T	G	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492448	A	C	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492455	G	GTT	Frame shift*	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492460	C	T	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)

Table 3.7a (cont'd)

Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492464	G	T	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492509	CTTA	C	Codon deletion	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g066900	6636803 - 6639454	Glyma16g07230	6490023 - 6492571	6492539	A	T	Non-synonymous coding change	GDSL-like Lipase/Acylhydrolase	AT1G53920.1 (GDSL-motif lipase 5)
Glyma.16g067000	6640311 - 6646267	Glyma16g07240	6493587 - 6499049	6498906	G	GGAA	Codon insertion	CCT motif	AT5G14370.1 (CCT motif family protein)
Glyma.16g067200	6658305 - 6665699	Glyma16g07260	6511457 - 6518819	6511512	T	C	Non-synonymous coding change	Kub3-prov protein	AT3G03420.1 (Ku70-binding family protein)
Glyma.16g067200	6658305 - 6665699	Glyma16g07260	6511457 - 6518819	6511797	A	T	Non-synonymous coding change	Kub3-prov protein	AT3G03420.1 (Ku70-binding family protein)
Glyma.16g067500	6679009 - 6683608	Glyma16g07280	6532136 - 6536360	6536221	T	A	Non-synonymous coding change	alpha/beta hydrolase	AT5G14390.1 (alpha/beta-Hydrolases superfamily protein)
Glyma.16g067700	6699818 - 6704992	Glyma16g07300	6552432 - 6557313	-	-	-	-	Mediator complex subunit 28	AT3G01680.1 (Mediator complex subunit)
Glyma.16g067800	6715004 - 6726073	Glyma16g07330	6574306 - 6578855	6574345	A	G	Non-synonymous coding change	Unknown protein	AT5G40600.1 (unknown protein)
Glyma.16g067900	6729704 - 6731321	-	-	-	-	-	-	Methyltransferase	AT3G01660.1 (methyltransferase)
Glyma.16g068000	6733543 - 6736390	Glyma16g07350	6586743 - 6589581	-	-	-	-	Translation factor	AT1G07930.1 (translation elongation factor)
Glyma.16g068100	6761573 - 6765861	Glyma16g07360	6614785 - 6619065	6614967	G	T	Non-synonymous coding change	Cytochrome P450	AT5G14400.1 (cytochrome P450)
Glyma.16g068100	6761573 - 6765861	Glyma16g07360	6614785 - 6619065	6616052	T	G	Non-synonymous coding change	Cytochrome P450	AT5G14400.1 (cytochrome P450)
Glyma.16g068100	6761573 - 6765861	Glyma16g07360	6614785 - 6619065	6616627	T	C	Non-synonymous coding change	Cytochrome P450	AT5G14400.1 (cytochrome P450)

*DNA sequence variations with high effects
Pull-down sequences /primer sequences were designed based on Glyma.Wm82.a1v1

Table 3.8a Information of SNPs used in the present fine mapping study

Coded Name	Original Name	Assay type	Reference	Position (bp)[a]	Position (bp)[b]	Flanking sequence (201bp or 100 bp)[c]
Gm08-1	Gm08_38277532_T_C	-	SoySNP50K	38277532	38909869	-
Gm08-2	Gm08_39004896_A_G	-	SoySNP50K	39004896	39638822	-
Gm08-3	MSUSNP08-40	KASP™	SNP discovery[d]	39410860	40047323	5'GACAAGAAGCAACGAATTCCTCAAATTCAAACATC TTAATGCAATCAATGCTTCCAATCAACCGGAGTTAAT ACACTTGATTAGGAGCGGACGATATTTA[G/T]CAAAA CAAAACTGCAGATGGGAGACAAAGTGACAGATCCCA GTAGCTGAAGATGACACAAAATTCCATACAGAAGCA TGCAAAAGTTAATCAGTCAAATG -3'
Gm08-4	MSUSNP08-44	KASP™	SNP discovery	40020445	40648794	5'ATCCGCGCGAAGCGTGCCTGAATCCTACCAAATGC GGATCAACTGAAACCTAAGGGGATCATCTACATCCC TACCCTATGACTACTGTGCTCAAGTATAA[C/T]GCAA ACGCGGATCAGCCAAATGCATACGCGGATCAACTAT CCTAAGCACTACACAAAAACCCCACAATGGAAACGC AAACGACATCGAGGGAGAAGAGTG-3'
Gm08-5	Gm08_40289000_C_T	-	SoySNP50K	40289000	40925463	-
Gm08-6	Gm08_40766548_G_T	-	SoySNP50K	40766548	41402338	-
Gm08-7	Gm08_40801297_C_T	-	SoySNP50K	40801297	41437087	-
Gm08-8	Gm08_40884892_A_G	-	SoySNP50K	40884892	41520716	-
Gm08-9	Gm08_40995409_T_C	-	SoySNP50K	40995409	41640233	-
Gm08-10	Gm08_41061846_G_A	-	SoySNP50K	41061846	41705040	-
Gm08-11	MSUSNP08-49	KASP™	SNP discovery	41072605	41715799	5'AAAGTAGGGACATTGGCCGAATAGCCTACAACACA CGATACACGAACTGGGAAAAAATAACTTTAAATTGC ACAATAATGTAATGCAGTTTTTCTAAAAT[G/C]TATTT GATTAATTTTGAAAATTAAGCTATTATTATAACCTTA ATGTGCTTACAAACTCATTGATTTGCTTGATATATAA GTTTAAGTTCGACTGATGAAA-3'
Gm08-12	MSUSNP08-50	KASP™	SNP discovery	41305451	41948645	5'AATTGCTTTCAAGTAGTTGTCGGACCAATTGTGTAA GGAACATAACCAAATAACAACATTGTCCTTAGGACA TATAATGACCATTTGCCAATGTGCATTG[C/T]ATTGGA GAATATTAAGTTATTATAATGCATACATTATTTAAAT TGATTTGTTCAGTTTTACTTACTCATTCAAGTAGGTTC TTAGGTAAACATCTCTATG-3'

Table 3.8a (cont'd)

Gm08-13	MSUSNP08-96	KASP™	SoySNP50K	41419491	42061322	5'TTGTTAGAGATGTCCATATTTAGTGTCTGACCGTGC CTCAAACAAGATTGCTTATGAAGGAGGAGAATTGAG GGAAACAAAAACAAATAAGCCTTATTTT[A/G]TTGAG CTAGTGACCTATGCATATTGTCAGCTAGCCAAACCAT ATGTTGTTGGTTGCAACCTATATACCTAACTCTTGGT CTAAGTGGCTCGTCATTGAAG-3'
Gm08-14	MSUSNP08-97	KASP™	SoySNP50K	41453586	42095417	5'ATATTTTCAAAATCTATTCATTCTTGTAATTTTTTTA AGAAATTAACCCATTTGTGTAAATTTGCCAACATTTG AAGATATAGAAGCAATTTTAAATTTT[T/C]TCTGTAGT AAATAATATTTTATGGGCTTTGACTTTTGCAGGAAGG GTAAATTAAATAAATTATGAAAAGGAAAATGTTAGT GATACATCCCATTTTGGTT-3'
Gm08-15	MSUSNP08-100	KASP™	SoySNP50K	41504420	42146252	5'TCATTTAATTACAAAAACCTCATCATTTTTTTAAAA CTTTATTTATTTATAAATAATAATTCTTTTTAAATTAA TCTACGAAAAATGGGATGTTACACCT[T/C]CACCCTT GGATTCTCCCTTCTCAACCTTGTGCTATGCCTGCCTCC CTCACTTTGCGGATCGTGAGCCACGTGTCCCTTTCTT CTAACAGAATTTCGTGCC-3'
Gm08-16	MSUSNP08-51	KASP™	SNP discovery	41518390	42160222	5'GAGGGGTAGAGGGTGTCACATGAGTGAAGTTTCAT ACCGGTTAAGTAATCACAAGAAGATTATCATTTCTGC TGGCATGAAATCTTCCTCCTCTTTGATG[T/C]ATGTAG AAGTTGTTCCCACTCAATGGATTCACATTCAAAAATT TCTTTGCCACTATTGCTGACCAGCTTCAATCCTGATG TTGTGGCATATACTGGGAAG-3'
Gm08-17	MSUSNP08-101	KASP™	SoySNP50K	41553888	42195720	5'TAGTGGTGAGCACGAGTCGGTTTGAGAAAAAAATC CTAATCCGATAAAAACCGACCACTGAAAAATATGGC CCATCACTGTCTACTGTCCGTTTAAGTGT[G/T]GAGAT TGAATCGAGTAAGGTGGGTGGGTTATTTGGGATCAT CTTCATTCTCGACAACCACAAGAGGCTCTAAGTTCTG GTCCTGCACTAGTGTGCATCAT-3'
Gm08-18	Gm08_41650869_A_C	-	SoySNP50K	41650869	42292701	-
Gm08-19	MSUSNP08-52	KASP™	SNP discovery	41696347	42338179	5'GGCATATGACTCGGTGTCGTGGGATTTTCTGTTATA TATGTTGAAGAGAACAGGCTTTAGTTCTAAATGGAT AAGGTGGATGGAAGGGTGTTTGAATTCT[G/A]TCTCA ATTTCAGTCTTGGTAAATGGCAGCCCCACAACGGAG TTCATACCTCAAAGGGGTCTTAGACAAGAGGACCCT TTAGCTCCATTTTTATTCAATGT-3'
Gm08-20	Gm08_41807810_A_G	-	SoySNP50K	41807810	42448802	-

Table 3.8a (cont'd)

Gm08-21	MSUSNP08-53	KASP™	SNP discovery	41880642	42521634	5'CATGGCAAAAATATTTTAAAGATATATAATCTTGT GTATTTTTTTCTTACAATATAAATTAATATGGATATA ATAAATTATGATTGGTCACTTAAATAGA[C/T]AGACA TAAAAAATGGATGAAATTTATTAGGTTGTTGAAACC CACTTTGATAAAATCTATGGATTGGGCTTTAAATTTT AAAATTGAATCAATGGGAGCTA-3'
Gm08-22	MSUSNP08-54	KASP™	SNP discovery	41984985	42636560	5'CAGAGGATCAATTTTTGGGTTATTTTGGGTTGTTTT ATGAAATTCAATTCCATTCTTGTGTTTTTAATCATGG ATTGATTGTGTTTGACGGATCAATTGG[C/T]GTCCCAA TGCGAAATTGTTTTGAAATTGGTATGTTTTTGTGTTA AGTATGAATCCTAGGAATTAGGTTTTTTTTTCTTCTA TTTAGTGTGAATTGTTGA-3'
Gm08-23	MSUSNP08-71	KASP™	SoySNP50K	42050788	42702363	5'GAGAGGACAGGAAAAGCATTTCCTTGGTAAGTCTA ATACAATGTTTCCATATACTTTTCAAGTCCAAGGAAC ATTAGTTAGCGCAAAAATTACTAATCTA[T/G]AGTCA CACTCTAACCACAATTTTGTCCAGCCCACTGAATGGG CATATTCAATTGCAGGATTGAATTCGGTCCTAACATT TTAGAGGCTTCAAGCATAAAA-3'
Gm08-24	MSUSNP08-69	KASP™	SoySNP50K	42563620	43390745	5'TTCATACAGCTCGTACAATTAGCAACAGTATAGCTT CATTTTTTCTTTTAAAATTTTAAGTTTAAATTTTCTGT ACGATTAGTGAGGGTAGCTGGTATTG[T/C]TCACCAT CCTGACAAACTCCTCAGCACTAACCTCACGATCATCT GCATATTTTACAGCCAAAAAACAGATATAACATCTC ACAATGATGAAAATAAATTA-3'
Gm08-25	MSUSNP08-67	KASP™	SoySNP50K	43293884	44128863	5'TACCGCGAAGGATGAGATTAATCCTTAATCGCTAC CACTATATAAAAACTCGTAAATACAACTCTCACTTTT GCAACTGTTTACAACAATAATAAGTTAA[T/C]AACTC GCGTAGCGCAAGGAACAGCATAAACGACGCTGCGCT GAACAACGACGGTGTCAGACAAACGGTGCCGGTTGG TGGGGCCGCCACGCTATTACGCC-3'
Gm08-26	MSUSNP08-65	KASP™	SoySNP50K	44373623	45398297	5'TTTCATAAGCCAGTAAGATAAATGTCTTCCTTCCAT GAATTATGTTTTCGTTGCAACCAAATGCCAAAGGTAC AAGGGAGAAGGGAAAGGGAAGGCCAGT[G/A]TAGGG AAATGAGGAGAGGTAATTGTGTCACTTTACAGTTTTT TGAAAAATCTGGGGGAGAAAGGAAAACAAGGAGAT GAGAATAGAAATTCCCTGAAAAA-3'

Table 3.8a (cont'd)

Gm08-27	MSUSNP08-64	KASP™	SoySNP50K	45060561	45913059	5'CGCGGCACCACCACCACCCATCGAACCCCTCTGCT GAACCCTCCAGAAGGACGCGCTGCGGGAGCGCGCGT GGGATTGGAATAAGGGTGAGGCTTGGAGC[C/T]TGGG GTGTGTGGAACACGCGACTGGGATTGAGGCCATTAA GAAGGGAGGGGGGGTGTCGGCGGCGATGACCGTGA AGGCGCTGCCGGTGAGGTTAGGGAG-3'
Gm08-28	MSUSNP08-4	TaqMan®	SoySNP6K	45189358	46037031	5'GCAACAAGATTAGAAGGCCTAACTCTTTAAAAACA GTCCCCAACCCCTTCGGCGGGAGGGCGACGCGGGGC TCACGAGGGCATCTTCCAAGGGAGGAAGG[T/C]GCGT AGAGTCGCCACCAACGTTTATTCGAGGAAAACGTCG GAAAAACCGGAAAGGTGTGGTCTACGGACTTTAAGC GTGAAAGGTTCGGGAGTTG-3'
Gm16-1	MSUSNP16-10	KASP™	SoySNP50K	6262227	6314120	5'CCCATGATGTCATGAGGTGTAAACTTGTTAAGACA TATCAAACTTAGGGTTTAAGTTAAC[C/T]AGATCCGA AAAAGCTGCCACTATAGTGCCTTCTCTTTGAGTATGT GGTAATTATTGATTG-3'
Gm16-2	MSUSNP16-47	KASP™	SoySNP50K	6470812	6617689	5'GATACAAAATAAAGTAAATTATGAGTACATACACA TGCTTAGATCTAAAAACAATCAATATAAAATGTCAC ATATATGAAAACATGTTTGATATTGTAAA[G/A]TTAC ATAAATCAAACTTCTAAGACTAAATTTTCAATCTACA ATCTCCCTCTTTTTGGTTTTTGAGAATGCCAAATCAA AATGATGTGTATTGATGTTTTC-3'
Gm16-3	Gm16_6474663_A_G	-	SoySNP50K	6474663	6621540	-
Gm16-4	MSUSNP16-127	KASP™	SoySNP50K	6510537	6657416	5'CTCCAAGACTAGACGAACTCTTCAAGCTTTTCTCCA ACTCCAAAACTCACTAAAAAACCTCACAAAATCAAT AACTTTTCTCTACTTGGTACTAGTAGCT[A/G]GTGTGA AATGAGCAATGGTTGAGGCTCTATTTGCAGGGGCAG ATGAAGGTCCTAGAAGGTGTTGCCTGAAGCTTGGTCT AGGGAAGATGGCAAGGATGGC-3'
Gm16-5	MSUSNP16-178	KASP™	SoySNP50K	6624879	6771675	5'GATGTATCTTGTGTGGTGGCGGTGGTGGCCCAAGG CCGCGGTGTGTCGCGTGACTGCGTGAGTCGTGTCCAC GGTGAGGAGAAGAAGATGAGAAGAAATG[C/T]TGTA AGAGGAGAAGAATAAAGCAAGGTACTAGTCCTTAAA GTGGTACTAGTCCAATGGTTCTTAAAGTGAAAAAGA AAAAATCCAACGTACTAAAACTAA-3'
Gm16-6	MSUSNP16-180	KASP™	SoySNP50K	6716691	6871009	5'GAGAATTATTACTCTGCGAGGGCCTTGAACAACTG TTGCTTTATAAAGTGCAGTAGTTCCTGGATAGACCGC CAAAACATGTTTTCCAGGAGGGAAATCA[G/A]GAGC

Table 3.8a (cont'd)

						GCTAGAAGGATCATTACTCTTGGGGAAAGGAATGAT ATTTGCCATGGGAAGCTTGTATTGTCTGTGAATGCAG CAGGAGAAACCAATGAAAATAAA-3'
Gm16-7	MSUSNP16-137	KASP™	SoySNP50K	6729421	6883739	5'ATTGATATGCTTTGTTAATTATGGTGGTACAGAAAT CTCGCACTTTGGTTGTTGTTGTTGCCTCTCCTTTT CCCATTCGTGTATGTGTTTTTTTTGG[G/A]TTCCTTAT AATTGAAGCCACGTATTAGGTTGTGTAGTACCATGTT TCATGTTTTTGTTTGTTGGTACTTGATAAAAAAAAAA AAAAAAGTGAAGAGGGAG-3'
Gm16-8	MSUSNP16-85	KASP™	SoySNP50K	7070805	7228568	5'ATGCAAGGGAAGCAGCTGCAAGAGATGCAAGGGA TGCAAAGGTGGAGGCGAGAGATGTAAAGAGAACAA CAGTGACAGCAACAACCGCAACCGCATGAAC[G/A]T GATGAGTATTAATGTGTTGTTATGAACTTATGATGTT GGTTTATGTGGGGAAATAAATGATGTATGTACCTCTT CTTGCCTATGTAGTAGGTTTGGGTG-3'
Gm16-9	MSUSNP16-15	TaqMan®	SoySNP50K	8051585	8208418	5'TCCGTTTCATGTGTTTCACAATATCCTTATACTTAG AGCTATCAAAATGGGTCAGCCCGG[T/C]CTACATGGG CTGACCCGCAACGGGTTGAGCTAAAAGTGGGCTAGT CCAGCTCGGCTCACT-3'

[a] Position is according to Glyma.Wm82.a1 (Schmutz et al. 2010)
[b] Position is according to Glyma.Wm82.a2 assembly on SoyBase (Grant et al. 2010)
[c] Flanking seuquences were mined from Glyma.Wm82.a1 (Schmutz et al. 2010)
[d] SNP was discovered by mapping the reads from the whole genome re-sequencing data of E12901 to Glyma.Wm82.a1 (Schmutz et al. 2010)

Table 3.9a SNPs and INDELs discovered from the whole genome exome-capture sequencing of E12901 in the *Rag6* and *Rag3c* fine-mapped regions

Chromosome	Physical position(bp)*	Glyma.Wm82.a1	E12901
Gm08	41504561	T	C
Gm08	41504624	A	T
Gm08	41509151	ACGT	A
Gm08	41509156	GA	G
Gm08	41509158	TACCCAATA	T
Gm08	41509173	G	T
Gm08	41511432	TTGTTAA	T
Gm08	41519031	T	C
Gm08	41519484	T	C
Gm08	41519770	T	C
Gm08	41520615	T	A
Gm08	41521312	TG	GA
Gm08	41521322	C	T
Gm08	41521326	C	A
Gm08	41521441	C	T
Gm08	41521466	G	T
Gm08	41521554	A	ATCT
Gm08	41521643	G	A
Gm08	41523569	T	A
Gm08	41523617	A	C
Gm08	41523772	C	T
Gm08	41523786	T	C
Gm08	41532427	T	C
Gm08	41532432	C	A
Gm08	41532513	C	T
Gm08	41532697	C	T
Gm08	41532888	G	A
Gm08	41534025	G	A
Gm08	41537628	T	CA
Gm08	41537717	C	T
Gm08	41537978	A	G
Gm08	41538072	C	T
Gm08	41538087	C	A
Gm08	41540630	TACAA	T
Gm08	41540636	T	C
Gm08	41540664	T	C

Table 3.9a (cont'd)

Gm08	41540678	T	A
Gm08	41540683	A	G
Gm08	41541175	C	T
Gm08	41541256	T	C
Gm08	41541285	T	C
Gm08	41541311	A	T
Gm08	41541351	C	T
Gm08	41541399	G	A
Gm08	41541404	A	T
Gm08	41541407	A	G
Gm08	41541422	T	C
Gm08	41541439	C	T
Gm08	41541511	G	A
Gm08	41541532	T	C
Gm08	41541582	C	G
Gm08	41541954	T	C
Gm08	41541985	A	G
Gm08	41542026	T	G
Gm08	41542052	G	A
Gm08	41542064	GCG	AAA
Gm08	41542075	G	A
Gm08	41542109	G	A
Gm08	41542116	C	T
Gm08	41542126	G	A
Gm08	41542136	GA	AG
Gm08	41542145	G	A
Gm08	41542160	C	T
Gm08	41542210	G	A
Gm08	41542212	TG	CA
Gm08	41542215	C	T
Gm08	41542218	C	T
Gm08	41542227	G	A
Gm08	41542245	T	C
Gm16	6474663	A	G
Gm16	6474803	T	A
Gm16	6477224	T	C
Gm16	6477552	C	T
Gm16	6477668	G	A
Gm16	6477685	GG	AA

Table 3.9a (cont'd)

Gm16	6481071	C	A
Gm16	6481140	A	T
Gm16	6481246	T	A
Gm16	6481575	G	GGGA
Gm16	6481665	A	T
Gm16	6482551	A	G
Gm16	6490128	A	G
Gm16	6490135	T	G
Gm16	6490137	T	A
Gm16	6490341	G	A
Gm16	6490348	G	GAAT
Gm16	6490379	G	T
Gm16	6490388	CTA	C
Gm16	6490599	A	G
Gm16	6490605	T	C
Gm16	6490624	C	T
Gm16	6491249	C	T
Gm16	6491575	T	C
Gm16	6491630	A	C
Gm16	6491696	C	T
Gm16	6491831	G	T
Gm16	6491856	C	A
Gm16	6492005	G	A
Gm16	6492038	G	T
Gm16	6492047	T	C
Gm16	6492050	G	A
Gm16	6492339	G	A
Gm16	6492344	G	C
Gm16	6492356	C	T
Gm16	6492358	A	G
Gm16	6492371	TG	CC
Gm16	6492381	AC	GA
Gm16	6492389	G	T
Gm16	6492399	T	G
Gm16	6492448	A	C
Gm16	6492455	G	GTT
Gm16	6492460	C	T
Gm16	6492464	G	T
Gm16	6492509	CTTA	C

Table 3.9a (cont'd)

Gm16	6492539	A	T
Gm16	6496577	A	C
Gm16	6496643	C	A
Gm16	6496661	T	C
Gm16	6498489	A	C
Gm16	6498906	G	GGAA
Gm16	6507406	G	A
Gm16	6507450	C	CTT
Gm16	6507475	C	T
Gm16	6507486	C	T
Gm16	6507567	A	T
Gm16	6507754	G	A
Gm16	6511407	GA	G
Gm16	6511512	T	C
Gm16	6511797	A	T
Gm16	6511816	T	A
Gm16	6511904	TC	T
Gm16	6511906	TC	T
Gm16	6511911	C	T
Gm16	6512694	A	G
Gm16	6512944	AT	A
Gm16	6525101	GC	G
Gm16	6534261	C	T
Gm16	6534270	A	G
Gm16	6536028	G	A
Gm16	6536221	T	A
Gm16	6542809	T	C
Gm16	6542838	T	C
Gm16	6548053	C	T
Gm16	6552391	T	C
Gm16	6552513	A	AT
Gm16	6552525	G	A
Gm16	6552621	T	G
Gm16	6552765	A	T
Gm16	6555297	A	G
Gm16	6568373	T	A
Gm16	6571017	C	T
Gm16	6574259	T	A
Gm16	6574345	A	G

Table 3.9a (cont'd)

Gm16	6582889	C	T
Gm16	6582898	C	T
Gm16	6582958	T	C
Gm16	6583157	TTA	T
Gm16	6583190	A	C
Gm16	6583209	G	T
Gm16	6583975	C	G
Gm16	6583992	CTTAAA	C
Gm16	6584029	C	T
Gm16	6584194	C	A
Gm16	6584419	C	T
Gm16	6584471	G	A
Gm16	6584477	G	A
Gm16	6587151	G	C
Gm16	6587505	G	A
Gm16	6587643	T	A
Gm16	6588425	G	C
Gm16	6589456	A	T
Gm16	6589499	GT	G
Gm16	6589507	T	A
Gm16	6589532	C	G
Gm16	6606049	G	T
Gm16	6609767	G	A
Gm16	6610055	A	G
Gm16	6614967	G	T
Gm16	6615640	A	G
Gm16	6616052	T	G
Gm16	6616122	G	T
Gm16	6616627	T	C
Gm16	6617276	A	ACT
Gm16	6623320	C	T
Gm16	6627491	TA	T

* Physical position is according to Glyma.Wm82.a1 (Schmutz et al. 2010)

REFERENCES

REFERENCES

Alt J, Ryan-Mahmutagic M (2013) Soybean aphid biotype 4 identified. Crop Sci 53(4):1491-1495

Ascencio-Ibáñez JT, Sozzani R, Lee TJ, Chu TM, Wolfinger RD, Cella R, Hanley-Bowdoin L (2008) Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. Plant Physiol 148(1):436-454

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. Theor Appl Genet 126:2081-2091

Bilyeu K, Palavalli L, Sleper DA, Beuselinck P (2006) Molecular genetic resources for development of 1% linolenic acid soybeans. Crop Sci 46(5):1913-8

Beckendorf EA, Catangui MA, Riedell WE (2008) Soybean aphid feeding injury and soybean yield, yield components, and seed composition. Agron. J.100(2):237-246

Bales C (2013) Fine mapping of aphid resistance genes in soybean plant introduction (PI) 567598B. Dissertation, Michigan State University, East Lansing, MI

Clark AJ, Perry KL (2002) Transmissibility of field isolates of soybean viruses by *Aphis glycines*. Plant Dis 86(11):1219-1222

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly 6(2): 80-92

Du W (2016) Map and fine map aphid resistance genes in soybean plant introduction (PI) 567597C, 567585A and 567537. Dissertation, Michigan State University, East Lansing, MI

Faris JD, Zhang Z, Lu H, Lu S, Reddy L, Cloutier S, Fellers JP, Meinhardt SW, Rasmussen JB, Xu SS, Oliver RP (2010) A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. Proc. Natl. Acad. Sci. 107(30):13544-13549

Fehr WR, Cavinese CE (1977) Stages of soybean development. Special Report, Agriculture and Home Economics Experiment Station, Iowa State University, issue 80:11

Fernandez-Cornejo J, Nehring R, Osteen C, Wechsler S, Martin A, Vialou A (2014) Pesticide Use in U.S. Agriculture: 21 Selected Crops, 1960-2008. Washington DC: US Department of Agriculture. Economic Research Service, Economic Information Bulletin. 124

Fox CM, Kim KS, Cregan PB, Hill CB, Hartman GL, Diers BW (2014) Inheritance of soybean aphid resistance in 21 soybean plant introductions. Theor Appl Genet 127(1):43-50

Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl Acids Res 38:D843-D846. doi:10.1093/nar/gkp798

Guo X, Wang D, Gordon SG, Helliwell E, Smith T, Berry SA, St Martin SK, Dorrance AE (2008) Genetic mapping of QTLs underlying partial resistance to *Sclerotinia sclerotiorum* in soybean PI 391589A and PI 391589B. Crop Sci 48(3):1129-1139

Hartman GL, Domier LL, Wax LM, Helm CG, Onstad DW, Shaw JT, Solter LF, Voegtlin DJ, d'Arcy CJ, Gray ME, Steffey KL (2001) Occurrence and distribution of *Aphis glycines* on soybeans in Illinois in 2000 and its potential control. Plant Health Progr. doi:10.1094/PHP-2001-0205-01-HN

Hesler LS, Dashiell KE, Lundgren JG (2007) Characterization of resistance to *Aphis glycines* in soybean accessions. Euphytica 154(1-2): 91-99

Hill JH, Alleman R, Hogg DB, Grau CR (2001) First report of transmission of Soybean mosaic virus and Alfalfa mosaic virus by *Aphis glycines* in the New World. Plant Dis 85(5):561-561

Hill CB, Li Y, Hartman GL (2004a) Resistance to the soybean aphid in soybean germplasm. Crop Sci 44(1):98-106

Hill CB, Li Y, Hartman GL (2004b) Resistance of *Glycine* species and various cultivated legumes to the soybean aphid (Homoptera: Aphididae). J Econ Entomol 97:1071-1077

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(4):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(4):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 DJ, Hartman GL (2010) A new soybean aphid (Hemiptera: Aphididae) biotype identified. J Econ Entomol 103:509-515

Jakab G, Manrique A, Zimmerli L, Métraux JP, Mauch-Mani B (2003) Molecular characterization of a novel lipase-like pathogen-inducible gene family of Arabidopsis. Plant Physiol 132(4):2230-2239

Jun TH, Mian MAR, Michel AP (2012) Genetic mapping revealed two loci for soybean aphid resistance in PI 567301B. Theor Appl Genet 124:13-22

Jun TH, Mian MAR, Michel AP (2013) Genetic mapping of three quantitative trait loci for soybean aphid resistance in PI 567324. Heredity 111:16-22

Kang ST, Mian MAR, Hammond RB (2008) Soybean aphid resistance in PI 243540 is controlled by a single dominant gene. Crop Sci 48(5):1744-1748

Kim KS, Hill CB, Hartman GL, Mian MA, Diers BW (2008) Discovery of soybean aphid biotypes. Crop Sci 48(3):923-928

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

Kim KS, Hill CB, Hartman GL, Hyten DL, Hudson ME, and Diers BW (2010b) Fine mapping of the soybean aphid-resistance gene *Rag2* in soybean PI 200538. Theor Appl Genet 121:599-610

Kim KS, Chirumamilla A, Hill CB, Hartman GL, Diers BW (2014) Identification and molecular mapping of two soybean aphid resistance genes in soybean PI 587732. Theor Appl Genet 127(5):1251-1259

Kisha TJ, Sneller CH, Diers BW (1997) Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci 37(4): 1317-1325

Kumar D, Klessig DF (2003) High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity. Proc. Natl. Acad. Sci. 100(26):16101-16106

Leister D (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet 20(3):116-122

Lemos Filho JP, Paiva ÉA (2006) The effects of sooty mold on photosynthesis and mesophyll structure of mahogany (Swietenia macrophylla King., Meliaceae). Bragantia. 65(1):11-17

Li Y, Hill CB, Hartman GL (2004) Effect of three resistant soybean genotypes on the fecundity, mortality, and maturation of soybean aphid (Homoptera: Aphididae). J. Econ. Entomol. 97(3):1106-1111

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 Breeding 19:25-34

Li Y, Zou J, Li M, Bilgin DD, Vodkin LO, Hartman GL, Clough SJ (2008) Soybean defense responses to the soybean aphid. New Phytol 179(1):185-195

Li YH, Zhou G, Ma J, Jiang W, Jin LG, Zhang Z, Guo Y, Zhang J, Sui Y, Zheng L, Zhang SS (2014) *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nature biotechnol, 32(10):1045-1052

Lorang JM, Sweat TA, Wolpert TJ (2007) Plant disease susceptibility conferred by a "resistance" gene. Proc. Natl. Acad. Sci. 104(37):14861-14866

Lundin O, Rundlöf M, Smith HG, Fries I, Bommarco R (2015) Neonicotinoid insecticides and their impacts on bees: a systematic review of research approaches and identification of knowledge gaps. PLoS One 10(8):e0136928

Malumphy CP (1997) Morphology and anatomy of honeydew eliminating organs. In: Ben-Dov and Hodgson CJ (eds) World Crop Pests. Elsevier Science B.V., Amsterdam, The Netherlands. 7:269-274

Martin GB, Bogdanove AJ, Sessa G (2003) Understanding the functions of plant disease resistance proteins. Annu Rev Plant Biol 54(1):23-61

Meng F, Han Y, Teng W, Li Y, 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 123:1459-1465

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 (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 PI243540. Theor Appl Genet 117:955-962

Mohr PG, Cahill DM (2007) Suppression by ABA of salicylic acid and lignin accumulation and the expression of multiple genes, in Arabidopsis infected with *Pseudomonas syringae* pv. *tomato*. Funct Integr Genomic 7(3):181-191

Nagy ED, Bennetzen JL (2008) Pathogen corruption and site-directed recombination at a plant disease resistance gene cluster. Genome research 18(12):1918-1923

Nickell CD, Noel GR, Cary TR, Thomas DJ, Diers BW (2001) Registration of 'Loda' soybean. Crop Sci 41(2):589

Oh IS, Park AR, Bae MS, Kwon SJ, Kim YS, Lee JE, Kang NY, Lee S, Cheong H, Park OK (2005) Secretome analysis reveals an Arabidopsis lipase involved in defense against *Alternaria brassicicola*. Plant Cell 17(10):2832-2847

Ohnesorg WJ, Johnson KD, O'Neal ME (2009) Impact of reduced risk insecticides on soybean aphid and their natural enemies. J Econ Entomol 102:1816-1826

Painter RH (1951) Insect resistance in crop plants. Macmillan, New York

Paré PW, Tumlinson JH (1999) Plant volatiles as a defense against insect herbivores. Plant Physiol 121(2):325-332

Ragsdale DW, McCornack BP, Venette RC, Potter BD, MacRae IV, Hodgson EW, O'Neal ME, Johnson KD, O'neil RJ, DiFonzo CD, Hunt TE (2007) Economic threshold for soybean aphid (Hemiptera: Aphididae). J Econ Entomol 100:1258-1267

Ragsdale DW, Landis DA, Brodeur J, Heimpel GE, Desneux N (2011) Ecology and management of the soybean aphid in North America. Annu Rev Entomol 56:375-399

Sagor GHM, Zhang S, Kojima S, Simm S, Berberich T, Kusano T (2016) Reducing cytoplasmic polyamine oxidase activity in Arabidopsis increases salt and drought tolerance by reducing reactive oxygen species production and increasing defense gene expression. Front. Plant. Sci. 7: 214

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178-183

Singh RJ (ed) (2006) Genetic resources, chromosome engineering, and crop improvement: vegetable crops (Vol. 3). CRC press, Boca Raton, FL

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. doi:10.1371/journal.pone.0054985

Studham ME, MacIntosh GC (2013) Multiple phytohormone signals control the transcriptional response to soybean aphid infestation in susceptible and resistant soybean plants. Mol Plant-Microbe Interact 26(1):116-129. doi: 10.1094/MPMI-05-12-0124-FI

Tan X, Meyers BC, Kozik A, West MA, Morgante M, St Clair DA, Bent AF, Michelmore RW (2007) Global expression analysis of nucleotide binding site-leucine rich repeat-encoding and related genes in Arabidopsis. BMC Plant Biol 7(1):56

Thompson GA, Goggin FL (2006) Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. J Exp Bot 57(4):755-766. doi: 10.1093/jxb/erj135

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Heredity 93 (1):77-78

Wang S, Basten CJ, Zeng ZB (2012) Windows QTL Cartographer 2.5. DepT of Statistics, North Carolina State University, Raleigh, NC. 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(2):209-218

Xiao L, Hu Y, Wang B, 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 126:2279-2287

Yang Z, Honda K, Wang S, Ma X (2004) Re-evaluation of *Glycine soja* germplasm from north eastern China for aphid resistance. Journal of Jilin Agricultural Sciences 29(5):3-6

Yu N, Kim M, King ZR, Harris DK, Buck JW, Li Z, Diers BW (2015) Fine mapping of the Asian soybean rust resistance gene Rpp2 from soybean PI 230970. Theor Appl Genet 128(3):387-396

Zhang G, Gu C, Wang D (2009) Molecular mapping of soybean aphid resistance genes 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

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

Zhang Z (2012) QTL mapping of Phytophthora root rot resistance and aphid resistance in soybean. Dissertation, Michigan State University, East Lansing, MI

Zhang Z, Hao J, Yuan J, Song Q, Hyten DL, Cregan PB, Zhang G, Gu C, Li M, Wang D (2014) Phytophthora root rot resistance in soybean E00003. Crop Sci 54(2):492-499

Zybailov B, Rutschow H, Friso G, Rudella A, Emanuelsson O, Sun Q, van Wijk KJ. (2008) Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. PloS One 3(4):E1994

Zhang S, Zhang Z, Bales C, Gu C, DiFonzo C, Li M, Song Q, Cregan P, Yang Z, Wang D (2017) Mapping novel aphid resistance QTL from wild soybean, *Glycine soja* 85-32. Theor Appl Genet. doi:10.1007/s00122-017-2935-z

CHAPTER 4

PYRAMIDING DIFFERENT APHID-RESISTANCE GENES IN ELITE SOYBEAN GERMPLASM TO COMBAT DYNAMIC APHID POPULATIONS

The work presented in this chapter has been submitted:

Zhang S, Wen Z, DiFonzo C, Song Q, Wang D (2017) Pyramiding different aphid-resistance genes in elite soybean germplasm to combat dynamic aphid populations. Molecular Breeding (Submitted)

Abstract

The soybean aphid, an invasive species, has posed a significant threat to soybean production in North America since 2001. Use of resistant cultivars is an effective tactic to protect soybean yield. However, the variability and dynamics of aphid populations could limit the effectiveness of host-resistance gene(s). Gene pyramiding is a promising way to sustain host-plant resistance. The objectives of this study were to determine the prevalent aphid biotypes in Michigan, and to assess the effectiveness of different combinations of aphid-resistance genes. A total of eleven soybean genotypes with known resistance gene(s) were used as indicator lines. Based on their responses, Biotype 3 was a major component of Michigan aphid populations collected during 2015 – 2016. The different performance of *Rag*-'Jackson' and *Rag1*-'Dowling' along with the break-down of resistance in plant introductions (PIs) 567301B and 567324 may be explained by the presence of Biotype 3 or an unknown virulent biotype establishing in Michigan. With the assistance of flanking markers, twelve advanced breeding lines carrying different aphid-resistance gene(s) were developed and evaluated for effectiveness in five trials across 2015 to 2017. Lines with *rag1c*, *Rag3d*, *Rag6*, *Rag3c+Rag6*, *rag1b+rag3*, *rag1c+rag4*, *rag1c+rag3+rag4*, *rag1c+Rag2+rag3+rag4* and *rag1b+rag1c+rag3+rag4* demonstrated strong and consistent resistance. Due to the variability of virulent aphid populations, different combinations of Rag genes may perform differently across geographic regions. However, advanced breeding lines pyramided with three or four Rag genes likely will provide broader and more durable resistance to diverse and dynamic aphid populations.

Introduction

Soybean [*Glycine max* (L.) Merr.] is one of the most important crops in North America because of its multiple uses as an animal feed, cooking oil, biofuel, and human protein source. In 2016, the U.S. ranked the first in world soybean production (11.73 million metric tons) with 5.52 million metric tons exported (SoyStats 2016). However, soybean production in North America has been threatened by the soybean aphid, *Aphis glycines* Matsumura (Hemiptera: Aphididae), an invasive species native to Asia (Wu et al. 2004).

Soybean aphid has aggressively dispersed to all major soybean producing areas in the U.S. and Canada (Ragsdale et al. 2011) since its discovery in southern Wisconsin in 2000 (Alleman et al. 2002). The direct aphid stylet-feeding on plant sap is the most prominent damage that can cause up to 40% soybean yield loss (Ragsdale et al. 2007). Under heavy infestations, soybean foliage can be stunted, wrinkled, distorted, and wilted; yield components, such as seed size and number, are also reduced (Wu et al. 2004). Transmissions of plant viruses by soybean aphids lead to further yield loss in soybean production (Hill et al. 2001; Clark and Perry, 2002). In addition, honeydew secreted by aphids promotes growth of sooty mold on leaves, impairing soybean photosynthesis by blocking sunlight and causing additional yield loss (Malumphy, 1997; Lemos Filho and Paiva, 2006).

Currently, insecticides are wildly used to manage soybean aphids. However, this control method increases production cost, the risk of environmental contamination and the mortality of beneficial insects (e.g., natural enemies and pollinators) (Ohnesorg et al. 2009; Lundin et al.

2015). The formation of insecticide resistance in soybean aphid populations is also an increasing concern. A more cost-effective and environmentally friendly way to managing soybean aphids is to utilize the native host-plant resistance present in soybean germplasm. Extensive screening of different soybean germplasm pools has identified ~30 plant introductions (PIs) and cultivars with antibiosis (affecting insect biology or reproduction) or antixenosis (non-preference) resistance (Hill et al. 2004; Li et al. 2004; Yang et al. 2004; Mensah et al. 2005; Hesler et al. 2007; Mian et al. 2008a; Fox et al. 2014).

Despite the high number of PIs and cultivars identified as resistant to soybean aphid, many share the same resistance genes or alleles, this due in part to the genetic bottleneck of soybean germplasm used in North America (Hyten et al. 2006). Aphid resistance QTLs identified in North America are designated as *Rag* (Resistance to *Aphis glycines*); different resistance alleles have been uncovered at six loci, *Rag1* to *Rag6*. The dominant antibiosis-resistant *Rag1/Rag* (Hill et al. 2006a, b; Li et al. 2007), the recessive antibiosis-resistant *rag1c* (Zhang et al. 2009) and *rag1b* (Bales et al. 2013) were mapped to chromosome 7 between markers Satt463 and Satt567. Additionally, *Rag1* was fine-mapped to a 115-kb interval between markers SNPKS9-3 and SNPKS5 (Kim et al. 2010a). The dominant *Rag2* (Mian et al. 2008b; Hill et al. 2009) and *Rag5* (Jun et al. 2012) were mapped to a genomic region between Satt334 and Sct_033 on chromosome 13, but they confer different resistance modality (antibiosis vs. antixenosis) (Michel et al. 2010; Jun et al. 2012). *Rag2* later was refined to a 54-kb interval between markers SNP46169.7 and SNP21A (Kim et al. 2010b). Aphid resistance in 20 PIs is associated with *Rag2*, indicating *Rag2* may be a major aphid-resistance source in the USDA soybean germplasm collection (Fox et al. 2014). The recessive antibiosis *rag4* was mapped to a different location

(between Satt649 and Satt348) on chromosome 13 (Zhang et al. 2009). Jun et al. (2013) identified two major QTLs (QTL_13_1 and QTL_13_2) near *Rag2* and *rag4*, and a minor QTL (QTL_6_1) on chromosome 6; these three QTLs suggested PI 567324 has oligogenic antixenosis resistance to soybean aphids. Five aphid-resistance QTLs/alleles were detected in a region between markers Satt285 and Satt654 on chromosome16, and designated *Rag3* (antixenosis), *Rag3b* (antibiosis), *rag3* (antibiosis), *Rag3c* (antibiosis), *Rag3d* (antibiosis) and *Rag3e* (antixenosis) (Zhang et al. 2010, 2013; Bales et al. 2013; Du 2016; Zhang et al. 2017a). Additionally, *Rag3c* was delimited to a 150-kb interval between markers Gm16_6474663_A_G and MSUSNP16-178 (Zhang et al. 2017b). The antibiosis-resistance gene *Rag6* was refined to a 49-kb interval between markers MSUSNP08-100 and MSUSNP08-101 on chromosome 8 (Zhang et al. 2017a, b).

The biggest concern of employing host-plant resistance is the breakdown of single resistance genes by virulent biotypes. To date, four different soybean aphid biotypes have been discovered in North America. Biotype 1 is avirulent to all Rag genes (Hill et al. 2004). Biotype 2 can reproduce on soybean plants with *Rag1* (Kim et al. 2008). Biotype 3 readily colonizes soybeans with *Rag2;* it also reproduces on soybeans with *Rag1* in choice tests (Hill et al. 2010). A recent multi-year study reported that the occurrence of soybean aphid biotypes was highly variable across both locations and years in the Midwestern U.S. (Cooper et al. 2015). The variability and dynamics of aphid populations could limit the durability of effectiveness of a single resistance gene. In this study, PI 567541B (a natural pyramid of *rag1c*/*rag4*) and PI 567598B (a natural pyramid of *rag1b*/*rag3*) demonstrated the widest spectrum of resistance to aphids across locations and years (Cooper et al. 2015). Similarly, other studies showed that soybean lines with

artificial pyramids of *Rag1/Rag2* had significantly lower aphid colonization than lines with the *Rag1* or *Rag2* gene alone (Wiarda et al. 2012; McCarville et al. 2014). However, Alt and Ryan-Mahmutagic (2013) reported a new soybean aphid biotype, Biotype 4, capable of colonizing PI 567541B, PI 567598B and soybean lines with the pyramid of *Rag1*/*Rag2*. There are likely more virulent biotypes not yet discovered. Therefore, integrating cultivars with multiple resistance genes, particularly with different modes of action, is important to achieve a broader and more durable resistance against different aphid populations.

The Soybean Breeding and Genetics Program at Michigan State University (MSU) has identified seven soybean accessions carrying resistant alleles at four resistance loci, including *Rag1*, *Rag3*, *Rag4*, and *Rag6* (Bales et al. 2013; Zhang et al. 2009, 2010, 2013; Zhang et al. 2017a; Du, 2016). Zhang et al. (2017b) refined *Rag6* to a 49-kb interval between markers MSU08-100 and MSUSNP08-101, and *Rag3c* to a 150-kb interval between markers Gm16_6474663_A_G and MSUSNP16-178. Fine mapping studies of five other aphid-resistance QTLs (*rag1b*, *rag1c*, *rag3*, *Rag3d* and *rag4*) refined their genomic locations and identified closely linked SNP markers (Unpublished data). With the assistance of these SNP markers, a pool of improved soybean germplasm with different combinations of aphid-resistance genes was developed. The objectives of this study were to 1) assess the introgression of aphid-resistance gene(s) using the Illumina Infinium SoySNP6K iSelect BeadChip, 2) determine the prevalence of soybean aphid biotypes in Michigan, 3) assess the effectiveness of different Rag gene combinations against Michigan aphid populations.

Materials and methods

Plant materials

A total of eleven resistant soybean genotypes, including 'Jackson', LD05-16060 (*Rag1*-'Dowling'), PI 243540, PI 567543C, PI 567585A, PI 567597C, PI 567598B, PI 567541B, PI 567301B, E08934 (derived from *G. soja* 85-32) (Zhang et al. 2017a), and PI 567324, were used as indicator lines to screen for aphid biotypes in field-cage trials during the summers of 2015 and 2016. LD05-16060 was an advanced breeding line carrying the *Rag1* gene from 'Dowling' and was developed by Dr. Brian Diers at University of Illinois Urbana-Champaign (UIUC).

In total, twelve advanced breeding lines (Table 4.1) carrying different Rag gene(s) were developed through marker-assisted selection (MAS) with markers flanking the initial-mapped or fine-mapped regions (Li et al. 2007; Hill et al. 2009; Kim et al. 2010a, b; Zhang et al. 2017a, b; Unpublished data). LD05-16657a with *Rag1* and LD08-12430a with *Rag2* were developed by Dr. Brian Diers at UIUC while 'E' lines were developed at MSU in East Lansing, Michigan with different combinations of *rag1b, rag1c*, *Rag2, Rag3c, Rag3d*, *rag3*, *rag4, Rag6* (Hill et al. 2009; Zhang et al. 2009; Bales et al. 2013; Du, 2016; Zhang et al. 2017a) (Table 1). E00003 has been consistently susceptible to Michigan aphids over the years (Zhang et al. 2017a, b), and it served as a susceptible check in this study.

Table 4.1 Pedigree information for advanced breeding lines integrated with different Rag genes

Indicator Line	Rag gene(s)	Pedigree information
E00003	*none*	C95001 (AP1995) x C94043 (PIO 9281)
LD05-16657a	*Rag1*	Dwight (3) x (Dowling x Loda)
E14922	*rag1c*	[E00003 x (SDX00R-039-42 x PI 567541B)] x E00003
LD08-12430a	*Rag2*	LD02-4485(2) x (Ina x PI 200538)
E11950	*rag3*	(Titan x PI 567598B) x LD05-16060
E12904	*Rag3d*	(Skylla x PI567585A) x Skylla
E14923	*Rag6*	(Skylla x LD01-7323) x [E00003 x (Jiyu 71 x *G.soja* 85-32)]
E14912	*rag1b, rag3*	[LD01-5907 x (Titan x PI 567598B)] x LD02-4485
E13369	*rag1c, rag4*	E07051 x {[E00003 x (SDX00R-039-42 x PI 567541B] x E00003)}
E14902	*Rag3c*, *Rag6*	(Skylla x LD01-7323) x [E00003 x (Jiyu 71 x *G.soja* 85-32)]
E13901	*rag1c, rag3*, *rag4*	{(Skylla x PI 567598B) x [Skylla x (SDX00R-039-42 x PI 567541B)]} x E07051
E13903	*rag1c*, *Rag2, rag3*, *rag4*	{[Skylla x PI 567598B] x [Skylla x (SDX00R-039-42 x PI 567541B)]} x LD08-12430a
E14919	*rag1b*, *rag1c, rag3*, *rag4*	[E00003 x (SDX00R-039-42 x PI 567541B)] x [LD01-5907 x (Titan x PI 567598B)]

* Donor of each Rag gene was indicated with an underline

DNA extraction and the Illumina Infinium SoySNP6K iSelect BeadChip genotyping analyses to assess the effectiveness of MAS

Leaf tissue was collected from a seedling of each advanced breeding line. Genomic DNA from each sample was extracted using the modified CTAB protocol described by Kisha et al. (1997), and genotyped using the Illumina Infinium SoySNP6K iSelect BeadChip (Illumina, San Diego, CA), which consists of 5,403 single nucleotide polymorphisms (SNPs) selected from the Illumina Infinium SoySNP50K iSelect BeadChip (Song et al. 2013). The genome-wide SNP distribution of the Illumina Infinium SoySNP6K iSelect BeadChip was visualized with R (R Development Core Team 2016) (Figure 4.4a). Genotypes were called using the program GenomeStudio (1.9.4 version, Illumina, San Diego, CA). Each SNP was coded based on the standard codes for nucleotides derived from the International Union of Pure and Applied Chemistry. The quality of each SNP was checked as previously reported (Yan et al. 2010). SNPs with call rate <80% across all samples were removed from the dataset. The genome-wide SNP data of each advanced breeding line was compared to that of the original aphid-resistance-gene(s) donor, mined from the public SoySNP50K iSelect BeadChip data on SoyBase (Grant et al. 2010) except for E12901. Graphic representation of genomic regions of interest from each sample were drawn with the program FlapJack (Milne et al. 2010). SNP markers that are monomorphic between the original donor line and the elite parental line were filtered. At each SNP of the advanced breeding line, the allele same as that of the original donor was assigned with the black color, and the alternative allele was assigned with the gray color.

Evaluation for soybean aphid resistance

Indicator lines and the advanced breeding lines were evaluated in choice tests in field-cage trials (Mensah et al. 2005) during the summers of 2015 and 2016. All the lines were planted in a randomized complete block design with three replications in a 12.2 x 18.3 m aphid- and predator-proof polypropylene cage (Redwood Empire Awning Co., Santa Rosa, CA) on the Agronomy Farm of MSU, East Lansing, Michigan. In each replication, fifteen seeds from each line were planted in a single 60 cm long plot with 60 cm row spacing.

The advanced breeding lines were also evaluated in the greenhouse choice-tests (Mensah et al. 2005) in the Plant Sciences greenhouse at MSU during Fall 2015, Spring 2016 and Spring 2017. Eight seeds from each line were planted in a 125-mm deep, 105-mm diameter plastic pot. All the lines were arranged in a randomized complete block design with three replications. The greenhouse was maintained at 26/15 ºC day/night with supplemental light (14 hours/day) provided by sodium vapor lights.

Soybean aphids were collected from multiple locations across Michigan in the early summer of each testing year, and maintained on susceptible soybean plants (E00003) in field-cages or the greenhouse. In each trial, each plant was artificially infested with two wingless aphids at the soybean V_2 stage (Fehr and Caviness, 1977). Each plant was visually rated for aphid resistance using a 0-4 scale (Mensah et al. 2005) when the susceptible check reached rating of 3.0 (usually three weeks after the initial infestation). Criteria of the 0-4 scale are as follows: 0 = no aphids; 0.5 = fewer than 10 aphids; 1 = 11-100 aphids; 1.5 = 101-150 aphids; 2 = 151-300 aphids; 2.5 =

301-500 aphids; 3 = 501-800 aphids, leaves and stems are covered with aphids, leaves appear slightly curly and shiny; 3.5 = more than 800 aphids, the plant appears stunted with curled yellow leaves, the plant is covered with few cast skins, no sooty mold; 4 = more than 800 aphids, the plant appears stunted with severely curled yellow leaves, the plant is covered with cast skins and sooty mold (Mensah et al. 2005). A damage index (DI) for each replication of each line was calculated as DI (%) = ∑ (rating value x no. of plants in the category) / (4 x total no. of plants) x 100 (Mensah et al. 2005). The DI ranged from 0% (no infestation) to 100% (most severe infestation). In each trial, the average DI of each line from three replications were analyzed with one-way analysis of variance (ANOVA) at a significance level of 0.05 followed by paired-wise comparisons using the PROC GLM function in SAS 9.4 (SAS Institute, Cary, NC). Lines with DI less than 37.5% were considered as aphid-resistant (Zhang et al. 2017 a, b).

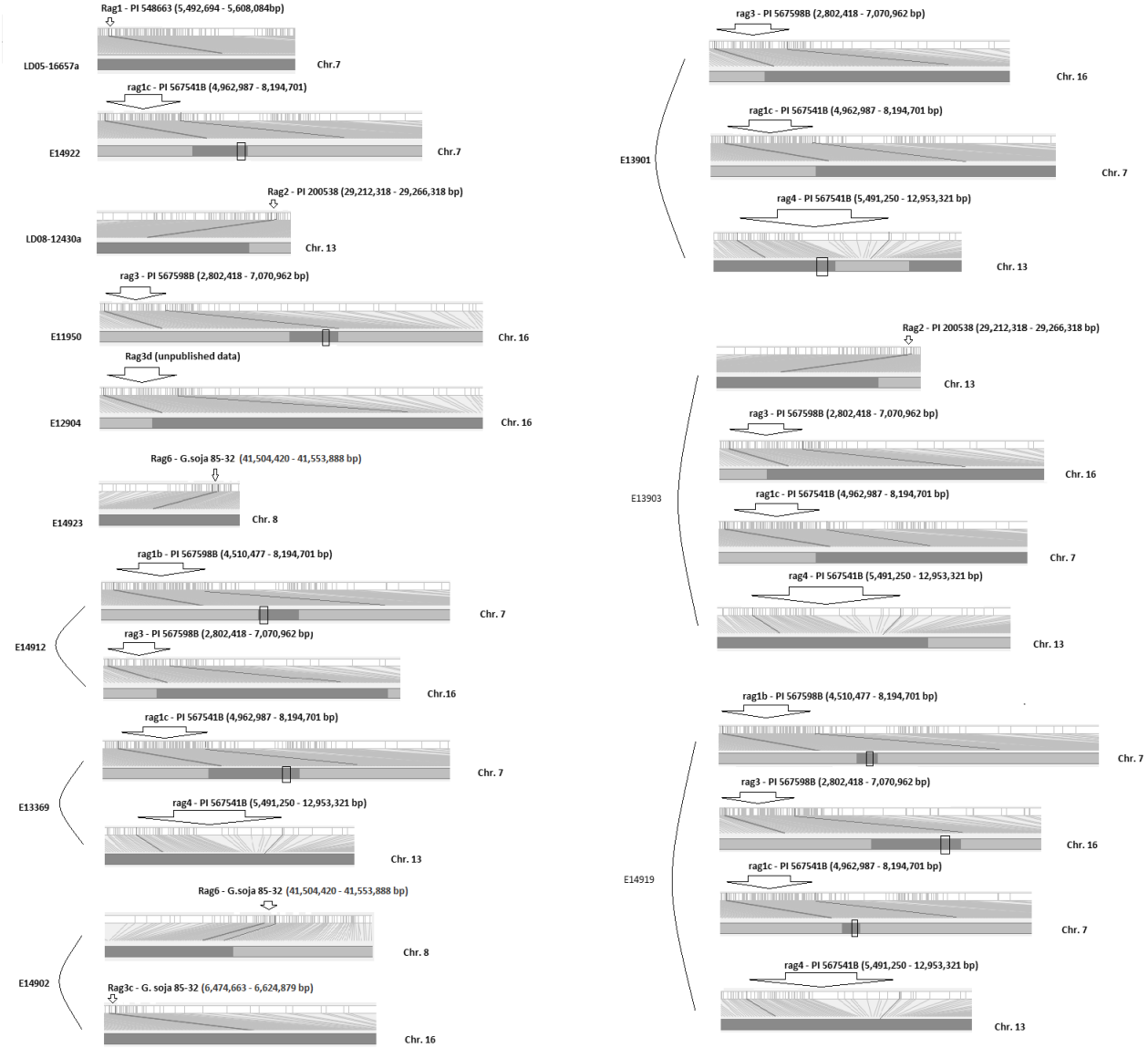
Results and discussion

Data from the Illumina Infinium SoySNP6K iSelect BeadChip verified the successful introgressions of all targeted aphid-resistance genes

The advanced breeding lines were visualized as graphical genotypes where genomic regions inherited from the original donor are indicated with the black color whereas genomic regions from the elite germplasm are presented in gray color (Figure 4.1). Targeted aphid-resistance genes with their published genomic locations (Glyma.Wm82.a1) were listed for each advanced breeding line. Unpublished fine-mapped regions of some Rag genes (including *rag1b*, *rag1c*, *rag3*, *rag4*) were indicated with rectangle boxes. When inspecting the regions of interest, all targeted aphid-resistance genes were successfully integrated into these advanced breeding lines,

which verified the different Rag gene combination in each of the advanced breeding lines. The original genome-wide SNP data of each advanced breeding line along with E12901 (the donor of *Rag6* and *Rag3c*) were presented in Table 4.4a (an electronic supplementary file).

Figure 4.1 Graphic representation of genomic region(s) of interest for each advanced breeding line. Genomic regions inherited from the original donor(s) of the aphid-resistance gene(s) are presented in black while genomic regions from the susceptible elite background are presented in gray. Targeted aphid-resistance genes with their published genomic locations are listed for each advanced breeding line. Unpublished fine-mapped regions of some Rag genes (including *rag1b*, *rag1c*, *rag3*, *rag4*) are indicated with rectangle boxes. The genomic locations are according to Glyma.Wm82.a1 on SoyBase (Grant et al. 2010)



Indicator lines suggested Biotype 3 and undescribed virulent biotype(s) prevailing in Michigan

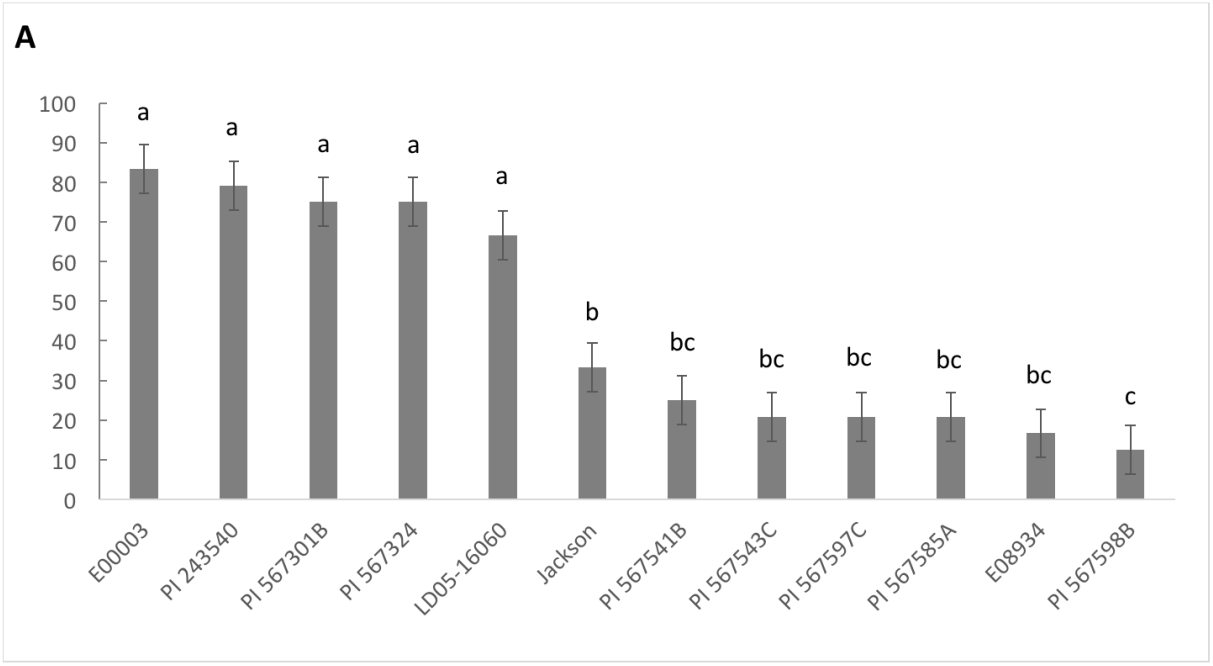
In both the 2015 and 2016 field-cage trials, LD05-16060 (*Rag1*), PI 243540, PI 567301B and PI 567324 were heavily colonized by aphids collected from Michigan fields and their DIs (ranging from 61.7 to 79.2%) were not significantly different from the susceptible check, E00003 (DIs ~ 79.2 to 83.3%) (Figure 4.2 and Table 4.2). PI 567585A was moderately resistant in 2016 (DI of 43.3%), although it performed better in 2015 (Figure 4.2 and Table 4.2). The remaining soybean genotypes, including 'Jackson' showed strong resistance (DIs ranging from 12.5 to 33%) to the same aphid populations in both field trials (Figure 4.2 and Table 4.2).

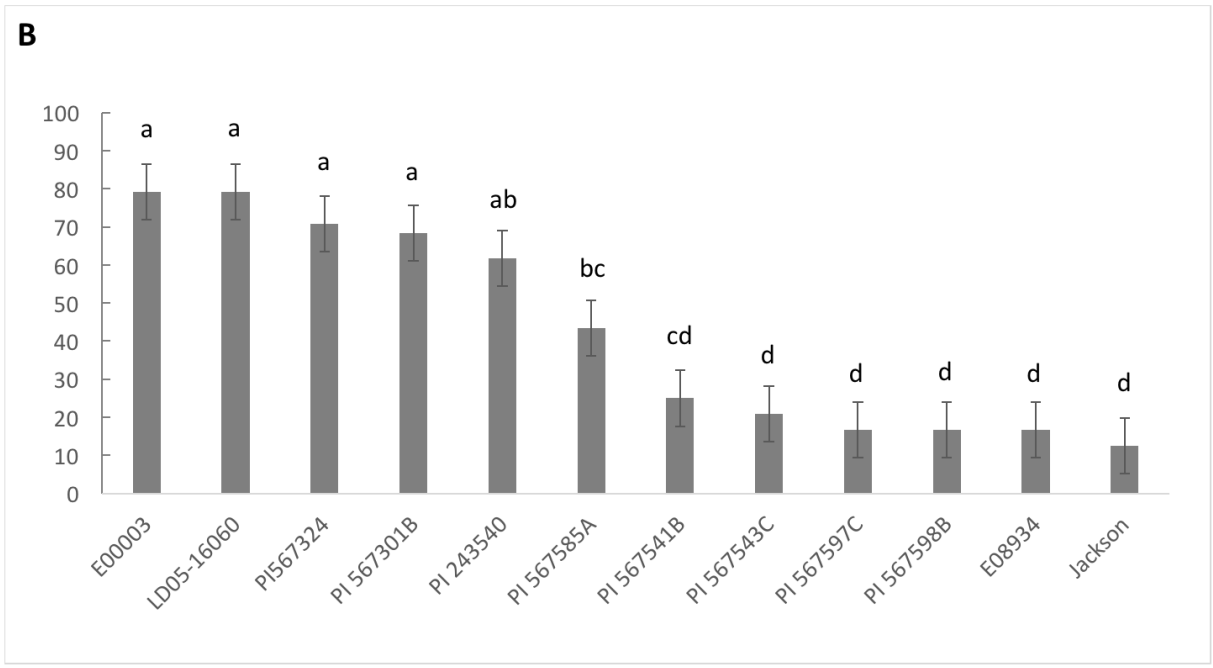
Table 4.2 Aphid damage indices (%) for indicator lines in field trials in Michigan, 2015-2016

Line	Rag genes	Mean soybean aphid damage index (%)*	
		Field 2015	Field 2016
E00003	*None*	83.3a	79.2a
LD05-16060	*Rag1*	66.7a	79.2a
PI 243540	*Rag2*	79.2a	61.7ab
PI 567301B	*Rag5 + QTL_8*	75a	68.3a
PI 567324	*Rag2′ + rag4′ + QTL_6_1*	75a	70.8a
Jackson	*Rag*	33.3b	12.5d
PI 567541B	*rag1c +rag4*	25bc	25cd
PI 567543C	*Rag3*	20.8bc	20.8d
PI 567597C	*Rag3e*	20.8bc	16.7d
PI 567585A	*Rag3d*	20.8bc	43.3bc
E08934	*Rag6+Rag3c*	16.7bc	16.7d
PI 567598B	*rag1b+rag3*	12.5c	16.7d

* DI (%) followed by same letter(s) are not significantly different at $P < 0.05$ in each trial

Figure 4.2 Aphid damage indices (%) of a susceptible check (E00003) and indicator lines used to screen for soybean aphid biotypes in (A) 2015 and (B) 2016 field-cage trials. Bars with same letter(s) are not significantly different at $P < 0.05$ in each trial





'Dowling'(*Rag1*) and 'Jackson'(*Rag*) were reported as overcome by Biotype 2 in both choice and no-choice tests (Kim et al. 2008). Biotype 3 aphids readily colonized *Rag2* soybeans in choice and no-choice tests as well as *Rag1* soybeans in choice tests (Hill et al. 2010). Alt and Ryan-Mahmutagic (2013) discovered a new biotype, Biotype 4, capable of colonizing PI 567541B and PI 567598B. In our study, the *Rag1* (LD05-16060) and *Rag2* (PI 243540) lines were readily colonized by aphids; in contrast, the *Rag* line ('Jackson'), PI 567541B and PI 567598B maintained strong resistance. This suggests that Biotype 3 aphids were a major component of the collected aphid populations in Michigan during 2015 and 2016.

The response of 'Jackson' to Biotypes 3 or 4 is unknown as it was not included in the previous aphid biotype studies by Hill et al. (2010) and Alt and Ryan-Mahmutagic (2013). In our study, 'Jackson' performed differently than LD05-16060 (carrying *Rag1*-'Dowling') in both years; it showed a strong resistance in 2015 and a very strong resistance in 2016 whereas LD05-16060 was consistently as susceptible as E00003. In a regional investigation conducted by Cooper et al. (2015), 'Jackson' was characterized as resistant in multiple states (SD, IA, MI and OH) whereas 'Dowling' was susceptible in all ten participating states in the year of 2010. Zhang et al. (2017a) also observed that 'Jackson' was resistant whereas 'Dowling' was susceptible in Michigan during 2010. Combining the evidences from Cooper et al. (2015) and Zhang et al. (2017a), the different reactions of these two varieties to aphid populations in some years (2010, 2015, and 2016) suggested that *Rag* and *Rag1* themselves are different, despite being mapped to a similar genomic region (Li et al. 2007). They could be allelic at a same locus or different QTLs located closely. 'Jackson' showed strong resistance to aphid populations that were primarily Biotype 3 in our field trials during 2015 and 2016, which suggests Biotype 3 is likely not able to overcome

the resistance in ‘Jackson’. Further study on the response of ‘Jackson’ to Biotype 3 is needed to exam this hypothesis. It is also possible that the different performance of *Rag1* and *Rag* in the present study was due to an undescribed aphid biotype capable of colonizing *Rag1* but not *Rag* soybeans. Single clones of Michigan aphids will be tested on ‘Dowling’ and ‘Jackson’ to explore this possibility.

Mian et al. (2008a) reported that PI 567301B had strong antixenosis resistance to Biotypes 1 and 2, controlled by a major QTL (*Rag5*) and a minor QTL on chromosome 8 (Jun et al. 2012). Similarly, Mian et al. (2008a) reported that PI 567324 showed moderate antixenosis resistance to Biotype 1 and strong resistance to Biotype 2, contributed by QTL13_1 mapped closely to *Rag2*, QTL13_2 mapped closely to *rag4* and a minor QTL_6_1 on chromosome 6 (Jun et al. 2013). Jun et al. (2013) suggested that the oligogenic resistance in PI 563724 would provide broader and more durable aphid resistance compared to lines with a single aphid resistance gene. However, in our field trials during 2015 and 2016, both PI 567301B and PI 563724 were heavily colonized by aphids. Although the reaction of these PIs to other biotypes has not been tested, their high damage indices (ranging from 68.3 to 75%) in our study could be explained by their susceptibility to Biotype 3 aphids which appeared to predominate the aphid population in 2015 and 2016; it also could be due to an undescribed virulent biotype in Michigan. PI 567301B and PI 563724 will be tested with Biotype 3 and/or single clones isolated from Michigan aphid populations to further investigate the hypotheses.

Table 4.3 Aphid damage indices (%) for advanced breeding lines in field and greenhouse trials in Michigan, 2015-2017

Line	Rag genes	Mean soybean aphid damage index (%)*				
		Field 2015	Greenhouse 2015	Field 2016	Greenhouse 2016	Greenhouse 2017
E00003	*None*	83.3a	68.5A	79.2a	75A	70.8a
LD05-16657a	*Rag1*	66.8b	70A	70.5ab	72.8A	68.3a
LD08-12430a	*Rag2*	83.3a	73.5A	76.7a	75A	67.5a
E11950	*rag3*	42.2c	12.5B	60.0b	25B	13.5c
E14923	*Rag6*	22.2de	19.5B	23.9c	23.6BC	36.0b
E12904	*Rag3d*	27.4d	12.5B	12.5c	12.5C	12.5c
E14922	*rag1c*	12.5e	12.5B	12.5c	12.5C	21.7c
E14902	*Rag3c + Rag6*	12.5e	12.5B	12.5c	12.5C	13.3c
E14912	*rag1b+rag3*	20.8de	12.5B	12.5c	12.5C	16.7c
E13369	*rag1c+rag4*	12.5e	14.1B	12.5c	12.5C	15.8c
E13901	*rag1c+rag3+rag4*	14.1e	19.8B	12.5c	12.5C	12.5c
E13903	*rag1c+Rag2+rag3+rag4*	14.2e	15.6B	12.5c	13.3C	16.7c
E14919	*rag1b+rag1c+rag3+rag4*	12.5e	13.0B	13.3c	13.3C	18.2c

* DI (%) followed by same letter(s) are not significantly different at $P < 0.05$ in each trial

Lines with rag1c or Rag3d or Rag6 or pyramided Rag genes showed strong and broad resistance

Several soybean lines with a single aphid-resistance gene were readily colonized by aphids in our study (Figure 4.3). LD05-16657a with *Rag1* and LD08-12430a with *Rag2* had severe aphid damages (DI ~ 66.8 to 88.3%) in all trials across 2015 – 2017 (Table 4.3), which was consistent with the performance of indicator lines, LD05-16060 (*Rag1*) and PI 243540 (*Rag2*). E11950 with *rag3* showed strong resistance in all the greenhouse trials but had moderate aphid damages (DI ~ 42.2 to 60%) in the field trials (Table 4.3), whereas the original donor, PI 567598B, had very strong resistance in the field trials (DI ~ 12.5 to 16.7%) (Table 4.2). PI 567598B also had the lowest frequency (18%) of aphid colonization across eleven locations during 2008-2010 (Cooper et al. 2015). Combining the results from Cooper et al. (2015) and the present study, the pyramid of *rag1b*/*rag3* is critical to provide soybean with broad and durable resistance.

Figure 4.3 Aphid damage indices (%) of a susceptible check (E00003) and the advanced breeding lines with different combinations of aphid-resistance gene(s) in (A) field-cage and greenhouse trials in 2015, (B) field-cage and greenhouse trials in 2016, and (C) a greenhouse trial in 2017. Damage indices from the field-cage trial were presented with gray bars followed by lower-case letters in (A) and (B). Damage indices from the greenhouse trial were presented with black bars followed by upper-case letters in (A) and (B). Within each trial, bars with same letter(s) are not significantly different at $P < 0.05$

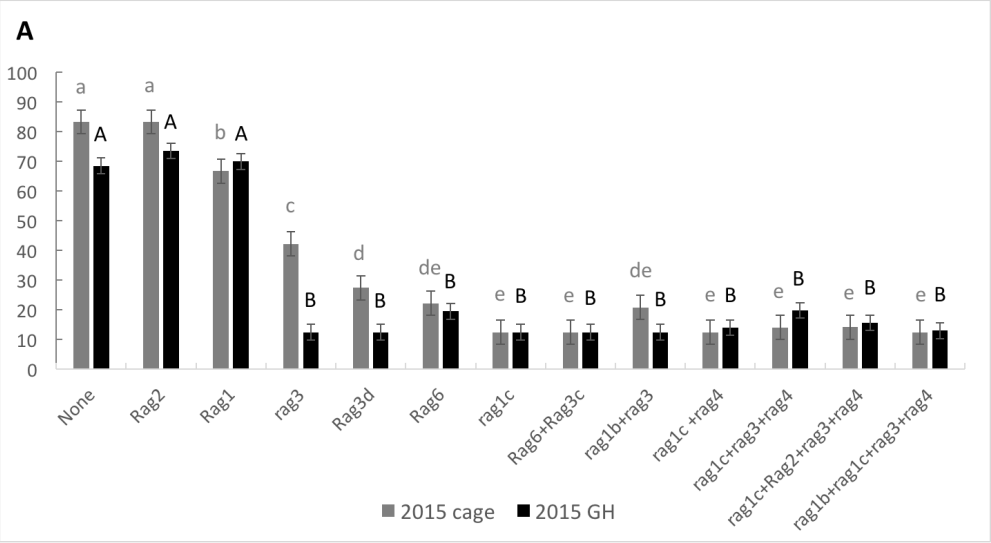
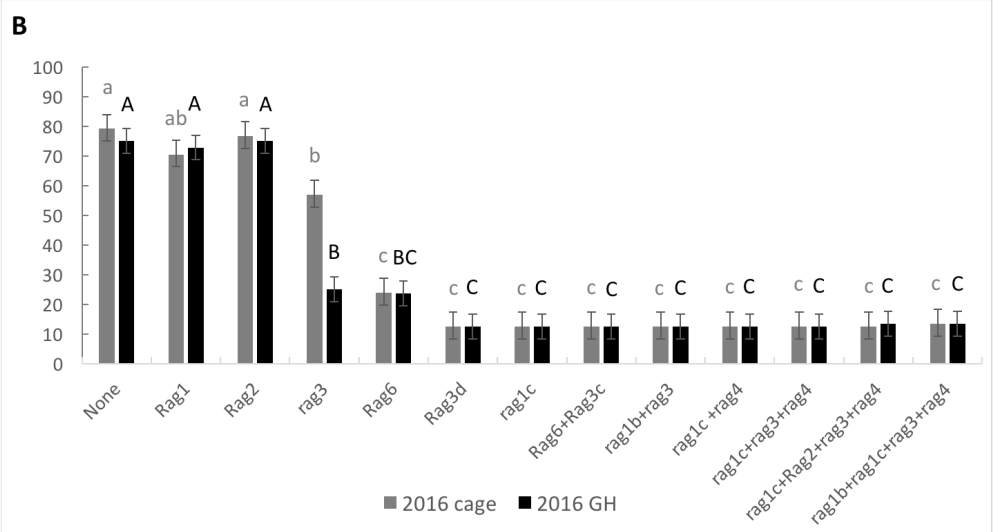
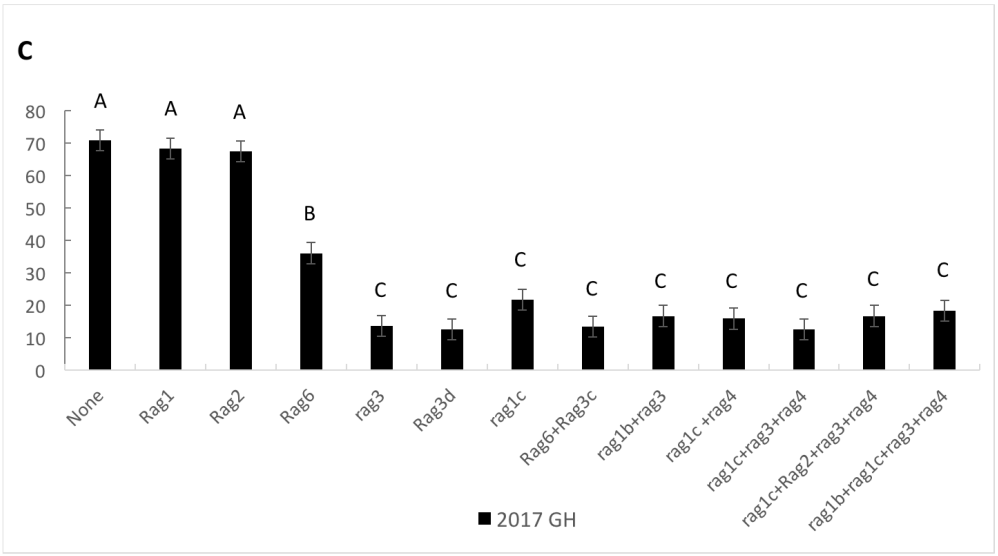


Figure 4.3 (cont'd)





E14923 with *Rag6* alone was highly resistant (DI ~ 19.5 to 23.9%) to aphids across all trials during 2015 - 2016. However, its damage index (36.0%) in 2017 greenhouse trial was slightly below the resistance threshold (DI ~ 37.5%), and it was statistically greater than those of the remaining resistant lines (Figure 4.3C and Table 4.3). The original donor, E08934 (*Rag6* + *Rag3c*), and the advanced breeding line, E14902 (*Rag6* + *Rag3c*), exhibited very strong and

consistent resistance (DI ~ 12.5 to 16.7%) across all trials (Tables 4.2 and 4.3). Collectively, *Rag6* alone offers a strong resistance, however, the pyramid of *Rag6*/*Rag3c* provides a stronger and more durable resistance.

E12904 with *Rag3d* appears to have a more consistent strong resistance compared to its original donor, PI 567585A. It displayed a strong resistance (DI ~ 12.5 to 27.4 %) across all five trials during 2015 - 2017 (Figure 4.3 and Table 4.3). However, PI 567585A had moderate aphid damage (DI ~ 43.3%) in 2016 field trial even though it had a lower damage index (20.8%) in 2015 field trial (Figure 4.2 and Table 4.2). The consistent strong resistance effect of *Rag3d* in E12904 may be attributed to the elite genetic background; some background gene(s) may upregulate the expression of *Rag3d.*

Across all five trials during 2015-2017, E14922 with *rag1c* showed a consistent strong resistance (DI ~ 12.5% to 21.7%) whereas LD05-16657a with *Rag1* was consistently susceptible (DI ~ 66.8 to 72.8%) (Figure 4.3 and Table 4.3). The strong resistance provided by *rag1c* alone suggested that *rag1c* is a different gene or allele from *Rag1* even though they were mapped in close proximity (Li et al. 2007; Zhang et al. 2009; Kim et al. 2010a). This conclusion is consistent with the genotypic evidence collected by Zhang et al. (2009); the band patterns of SSR markers flanking *rag1c* were distinctive between PI 567541B and 'Dowling'.

Among the resistant soybean genotypes tested by Cooper et al. (2015), PI 567541B and PI 567598B demonstrated the widest spectrum of resistance to aphid populations across North America during 2008-2010; the broad resistance was deduced contributed by the natural

pyramids of two resistance genes in these two PIs. However, PI 567541B and PI 567598B were later found fully colonized by Biotype 4 (Alt and Ryan-Mahmutagic, 2013). In our study, E14912 (*rag1b*+*rag3* from PI 567598B) and E13369 (*rag1c*+*rag4* from PI 567541B) showed very strong resistance across 2015 to 2017 (Figure 4.3 and Table 4.3), however, their resistance might be limited in geographic regions that have a higher pressure of Biotype 4 or other undescribed virulent biotypes.

rag1c and *rag3* are the two major genes controlling aphid resistance in PI 567541B and PI 567598B, respectively (Zhang et al. 2009; Bales et al. 2013). Additionally, Chandrasena et al. (2015) detected a significant additive x additive interaction between *rag1c* and *rag3*, contributing up to 24% of the phenotypic variation in aphid resistance. To achieve a broader and more durable resistance, additional aphid-resistance gene(s) were pyramided with *rag1c*+*rag3*. Advanced breeding line E13901 was pyramided with three aphid-resistance genes, including *rag1c*, *rag3* and *rag4*. Compared to E13901, E13903 has one more aphid-resistance gene, *Rag2*, to provide additional resistance. E14919 has all four genes from PI 567541B and PI 567598. All these advanced breeding lines (E13901, E13903 and E14919) pyramided with multiple aphid-resistance genes had very strong and consistent resistance to aphid populations in Michigan across 2015-2017 (Figure 4.3 and Table 4.3), and they are expected to be strong and durable when combating diverse and dynamic aphid populations across geographic regions.

Conclusion

The utilization of host-plant resistance is an effective way to control soybean aphids. However, the aphid resistance provided by *Rag1* soybeans, PI 243540 (*Rag2*), PI 567301B (*Rag5*) and PI 567324 (*Rag2'+rag4'+QTL_6_1*) was overcome by aphids in our field trials during 2015 and 2016. The high damage indices of PI 567301B and PI 567324 could be explained by their susceptibility to Biotype 3 aphids which appeared to be prevalent in our field trials. In contrast to the susceptibility of *Rag1* soybeans, 'Jackson' maintained strong resistance in the field trials during 2015 and 2016. Coupled with the similar evidences from Cooper et al. (2015) and Zhang et al. (2017a), *Rag1* and *Rag* are likely different loci or alleles, which may be distinguished by Biotype 3. In addition, it is possible that an undescribed virulent biotype prevalent in our field trials caused the susceptibility of PI 567301B and PI 567324 and the different responses from *Rag1* soybeans and 'Jackson'. Biotype 3 and single isolates of Michigan aphids will be tested on these soybean genotypes to further exam the hypotheses.

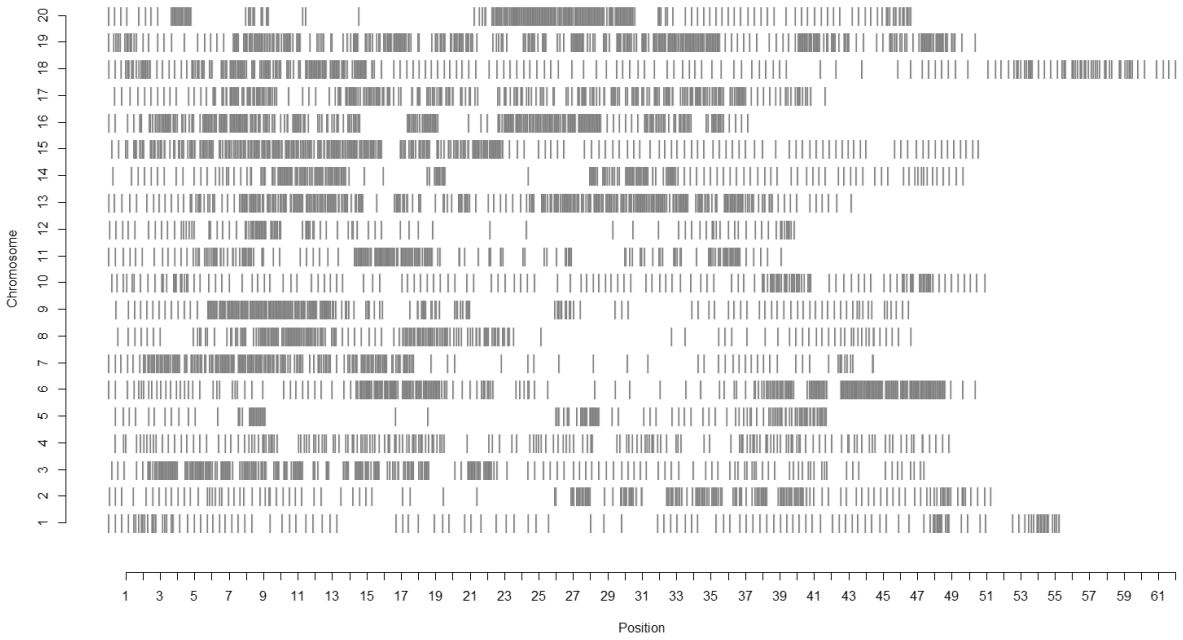
Advanced breeding lines with single aphid-resistance genes, such as *rag1c*, *Rag3d* and *Rag6* showed very strong resistance to Biotype 3 across trials during 2015 - 2017. The strong resistance provided by *rag1c* suggested that it is a different locus or allele from *Rag1* even though they were mapped closely. According to a regional study by Cooper et al. (2015), soybean aphids have a high degree of virulence diversity in North America, which means the effectiveness of a single aphid resistance gene is likely limited by soybean aphid virulence variability.

Advanced breeding lines pyramided with two aphid-resistance genes, such as *rag1b+rag3*, *rag1c+rag4*, and *Rag3c+Rag6* demonstrated strong resistance in Michigan. Although Biotype 3 dominated in our trials, there is variability in soybean aphid populations from year-to-year across the Midwest, and undescribed biotypes are likely yet to be identified. Lines with multiple Rag genes, such as *rag1c+rag3+rag4*, *rag1c+Rag2+rag3+rag4* and *rag1b+rag1c+rag3+rag4,* likely will provide broader and more durable resistance to diverse and dynamic aphid populations. The advanced breeding lines with different combinations of Rag genes developed in this study are significant resources for breeders to develop varieties to combat different aphid populations across many geographies.

APPENDIX

APPENDIX

Figure 4.4a The genome-wide SNP distribution of the Illumina Infinium SoySNP6K iSelect BeadChip visualized with R



REFERENCES

REFERENCES

Alleman RJ, Grau CR, Hogg DB (2002) Soybean aphid host range and virus transmission efficiency. In: Proceedings of Wisconsin Fertilizer, Aglime, and Pest Management Conference. Madison, WI

Alt J, Ryan-Mahmutagic M (2013) Soybean aphid biotype 4 identified. Crop Sci 53(4):1491-1495

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. Theor Appl Genet 126:2081-2091

Chandrasena D, Wang Y, Bales C, Yuan J, Gu C, Wang D (2015) Pyramiding 3, 1b, 4, and 1c aphid-resistant genes in soybean germplasm. Crop Sci. 55(5):2108-2115

Clark AJ, Perry KL (2002) Transmissibility of field isolates of soybean viruses by *Aphis glycines*. Plant Dis 86(11):1219-1222

Cooper SG, Concibido V, Estes R, Hunt D, Jiang GL, Krupke C, McCornack B, Mian R, O'Neal M, Poysa V, Prischmann-Voldseth D (2015) Geographic distribution of soybean aphid biotypes in the United States and Canada during 2008–2010. Crop Sci. 55 (6):2598-2608

Du W (2016) Map and fine map aphid resistance genes in soybean plant introduction (PI) 567597C, 567585A and 567537. Dissertation, Michigan State University, East Lansing, MI

Fehr WR, Cavinese CE (1977) Stages of soybean development. Special Report, Agriculture and Home Economics Experiment Station, Iowa State University, issue 80:11

Fox CM, Kim KS, Cregan PB, Hill CB, Hartman GL, Diers BW (2014) Inheritance of soybean aphid resistance in 21 soybean plant introductions. Theor Appl Genet 127(1):43-50

Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl Acids Res 38:D843-D846. doi:10.1093/nar/gkp798

Hesler LS, Dashiell KE, Lundgren JG (2007) Characterization of resistance to *Aphis glycines* in soybean accessions. Euphytica 154(1-2): 91-99

Hill JH, Alleman R, Hogg DB, Grau CR (2001) First report of transmission of Soybean mosaic virus and Alfalfa mosaic virus by *Aphis glycines* in the New World. Plant Dis 85(5):561-561

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

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, Li Y, Hartman GL (2004) Resistance to the soybean aphid in soybean germplasm. Crop Sci 44(1):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(4):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(4):1606-1608

Hyten DL, Song Q, Zhu Y, Choi IY, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB (2006) Impacts of genetic bottlenecks on soybean genome diversity. Proc Natl Acad Sci 103(45): 16666-16671

Jun TH, Mian MAR, Michel AP (2012) Genetic mapping revealed two loci for soybean aphid resistance in PI 567301B. Theor Appl Genet 124:13-22

Jun TH, Mian MAR, Michel AP (2013) Genetic mapping of three quantitative trait loci for soybean aphid resistance in PI 567324. Heredity 111:16-22

Kim KS, Hill CB, Hartman GL, Mian MA, Diers BW (2008) Discovery of soybean aphid biotypes. Crop Sci 48(3):923-928

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

Kim KS, Hill CB, Hartman GL, Hyten DL, Hudson ME, and Diers BW (2010b) Fine mapping of the soybean aphid-resistance gene *Rag2* in soybean PI 200538. Theor Appl Genet 121:599-610

Kisha TJ, Sneller CH, Diers BW (1997) Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci 37(4): 1317-1325

Lemos Filho JP, Paiva ÉA (2006) The effects of sooty mold on photosynthesis and mesophyll structure of mahogany (Swietenia macrophylla King., Meliaceae). Bragantia 65(1):11-17

Li Y, Hill CB, Hartman GL (2004) Effect of three resistant soybean genotypes on the fecundity, mortality, and maturation of soybean aphid (Homoptera: Aphididae). J Econ Entomol 97(3):1106-1111

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 Breeding 19:25-34

Lundin O, Rundlöf M, Smith HG, Fries I, Bommarco R (2015) Neonicotinoid insecticides and their impacts on bees: a systematic review of research approaches and identification of knowledge gaps. PLoS One 10(8):e0136928

Malumphy CP (1997) Morphology and anatomy of honeydew eliminating organs. In: Ben-Dov and Hodgson CJ (eds) World Crop Pests. Elsevier Science B.V., Amsterdam, The Netherlands. 7:269-274

McCarville MT, O'Neal ME, Potter BD, Tilmon KJ, Cullen EM, McCornack BP, Tooker JF, Prischmann-Voldseth DA (2014) One gene versus two: a regional study on the efficacy of single gene versus pyramided resistance for soybean aphid management. J Econ Entomol 107(4): 1680-1687

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

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 PI243540. Theor Appl Genet 117:955-962

Michel AP, Mian MR, Davila-Olivas NH, Cañas LA (2010) Detached leaf and whole plant assays for soybean aphid resistance: differential responses among resistance sources and biotypes. J Econ Entomol 103(3):949-957

Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WT, Flavell AJ, Marshall D (2010) Flapjack--graphical genotype visualization. Bioinformatics 26(24): 3133-3134

Ohnesorg WJ, Johnson KD, O'Neal ME (2009) Impact of reduced risk insecticides on soybean aphid and their natural enemies. J Econ Entomol 102:1816-1826

Ragsdale DW, Landis DA, Brodeur J, Heimpel GE, Desneux N (2011) Ecology and management of the soybean aphid in North America. Annu Rev Entomol 56:375-399

Ragsdale DW, McCornack BP, Venette RC, Potter BD, MacRae IV, Hodgson EW, O'Neal ME, Johnson KD, O'Neil RJ, DiFonzo CD, Hunt TE (2007) Economic threshold for soybean aphid (Hemiptera: Aphididae). J Econ Entomol 100:1258-1267

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. doi:10.1371/journal.pone.0054985

SoyStats. http://soystats.com/

Wiarda SL, Fehr WR, O'Neal ME (2012) Soybean aphid (Hemiptera: Aphididae) development on soybean with *Rag1* alone, *Rag2* alone, and both genes combined. J Econ Entomol 105(1):252-258

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(2):209-218

Yan J, Yang X, Shah T, Sánchez-Villeda H, Li J, Warburton M, Zhou Y, Crouch JH, Xu Y (2010) High-throughput SNP genotyping with the GoldenGate assay in maize. Mol Breed 25(3):441-451

Yang Z, Honda K, Wang S, Ma X (2004) Re-evaluation of *Glycine soja* germplasm from north eastern China for aphid resistance. Journal of Jilin Agricultural Sciences 29(5):3-6

Zhang G, Gu C, Wang D (2009) Molecular mapping of soybean aphid resistance genes 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

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

Zhang S, Zhang Z, Bales C, Gu C, DiFonzo C, Li M, Song Q, Cregan P, Yang Z, Wang D (2017a) Mapping novel aphid resistance QTL from wild soybean, *Glycine soja* 85-32. Theor Appl Genet. doi:10.1007/s00122-017-2935-z

Zhang S, Zhang Z, Wen Z, Gu C, An Y, Bales C, DiFonzo C, Song Q, Wang D (2017b). Fine mapping of the aphid resistance genes *Rag6* and *Rag3c* from *Glycine soja* 85-32. Theor Appl Genet (Under Review)