QTL MAPPING AND MOLECULAR CHARACTERIZATION OF THE SUGARCANE APHID [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] RESISTANCE GENE IN SORGHUM [*Sorghum bicolor* (L.) Moench]

By

JIAN HUANG

Bachelor of Science in Biotechnology
Wuhan Polytechnic University
Wuhan, Hubei
2008

Master of Science in Biochemistry and Molecular Biology
Shenyang Normal University
Shenyang, Liaoning
2012

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2019

QTL MAPPING AND MOLECULAR

CHARACTERIZATION OF THE SUGARCANE APHID

[*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)]

RESISTANCE GENE IN SORGHUM [*Sorghum bicolor*

(L.) Moench]

Dissertation Approved:

Dr. Yinghua Huang

Dissertation Advisor

Dr. Michael Anderson

Committee Chair

Dr. Liuling Yan

Dr. Andrew Mort

ACKNOWLEDGEMENTS

This work would not have been possible without those people that have earned my gratitude for their contribution to my Ph.D. career in OSU, I would like to thank them for everything they have done for me, and I am truly blessed to have them all in my life.

First and foremost, I would like to express my appreciation and gratitude to my thesis advisor, Dr. Yinghua Huang. He offered me a great opportunity to join his group in summer 2013, and I know that Dr. Huang believed in me and gave me the solidest support since my first day in his lab. I am grateful to Dr. Huang for his patience, motivation, enthusiasm and knowledge, not only for his academic support, but also for giving me so many great advice and wonderful opportunities for my life, inspired me by his hardworking and passionate attitude. I could not have imaged having a better advisor and mentor for my Ph.D. career.

Similar, profound gratitude goes to my committee chair, Dr. Michael Anderson, who has been helping and encouraging me during my whole Ph.D. career, he is always there for me, I truly appreciate that and thank him from the bottom of my heart. Besides my committee chair, I am also grateful to the rest of my dissertation committee members Dr. Liuling Yan and Dr. Andrew Mort for their great support and invaluable advice.

Special thanks to Angela Phillips for technical assistance in the Lab, greenhouse and

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

field, this dissertation would not have been possible without the intelligent contribution of Angela. Special mention goes to our collaborator Dr. Zhanyuan Zhang at the University of Missouri, for his valuable support in sorghum transformation. Also, I would like to thank my lab mates: Huang Li, Hengyou Zhang, Alimamy Fornah, Suzanne Coveley, Creighton Pagett, Kaleb Blackford, and Sarah Soliman. I am also grateful to faculty, staff and fellow graduate students in Department of Plant and Soil Sciences. Last but not the least, I would like to thank my family, my relatives, and my lovely wife because they are most important people to me in the world, thank for their warm love, constant patience, and endless support.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: JIAN HUANG

Date of Degree: MAY, 2019

Title of Study: QTL MAPPING AND MOLECULAR CHARACTERIZATION OF THE SUGARCANE APHID [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] RESISTANCE GENE IN SORGHUM [*Sorghum bicolor* (L.) Moench]

Major Field: CROP SCIENCE

Abstract: Sorghum [*Sorghum bicolor* (L.) Moench] is the fifth most important cereal crop in the world. Sugarcane aphid, *Melanaphis sacchari* (Zehntner), is a new invasive pest of sorghum species in North America, which causes severe damages and significant yield loss to sorghum crop. The first objective of this research project focused on the genetic basis of the host plant resistance to sugarcane aphid in sorghum. Molecular mapping was used to analyze an $F_{2:3}$ mapping population of sorghum developed by crossing susceptible cultivar BTx623 and sugarcane aphid resistant source Tx2783, led to the identification of quantitative trait loci (QTL) linked to sugarcane aphid resistance in sorghum. One major QTL was identified on chromosome 6 of sorghum and this locus accounted for a large portion of phenotypic variation observed for plant responses to sugarcane aphid infestation. SSR markers linked to this region were developed as a tool for marker-assisted selection in future breeding programs for genetic improvement of sugarcane aphid resistance in sorghum. The second objective of this study aimed at identifying and characterizing candidate genes against this pest. For this purpose, functional genomics and DNA sequencing were used to analyze the potential candidate genes in this region and the resulted molecular data allowed the identification of the first sugarcane aphid resistance gene of sorghum. In summary, discovery of the sugarcane aphid resistance gene and development of DNA markers linked to the sugarcane aphid resistance will contribute both the resistance gene and DNA markers for further genetic studies of the resistance mechanism as well as an efficient selection tool for breeding sugarcane aphid resistance in sorghum.

TABLE OF CONTENTS

Chapter I. Introduction Page

Introduction....................1
Review of Literature....................5
Significance....................12
Objectives....................14
References....................15

Chapter II. QTL mapping of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance gene in Sorghum [*Sorghum bicolor* (L.) Moench]

Abstract....................24
Introduction....................26
Materials and Methods....................28
Results....................32
Discussion....................39
References....................42

Chapter III. Molecular characterization of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance genes in sorghum [*Sorghum bicolor* (L.) Moench]

Abstract....................47
Introduction....................49
Materials and Methods....................53
Results....................64
Discussion....................72
References....................75

APPENDICES....................82

LIST OF TABLES

Table Page

Chapter II. QTL mapping of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance gene in Sorghum [*Sorghum bicolor* (L.) Moench]

Table 2-1 List of genes and their putative function in the 500kb region between marker Xtxp06 and ESR170. Three candidate genes were underlined 38

Chapter III. Molecular characterization of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance genes in sorghum [*Sorghum bicolor* (L.) Moench]

Table 3-1 List of sequences of the gene specific primers used in real-time PCR reaction. The primers sequences were designed based on the sequence of the cloned gene (Sb1650). 56

LIST OF FIGURES

Figure Page

Chapter II. QTL mapping of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance gene in Sorghum [*Sorghum bicolor* (L.) Moench]

Fig 2-1 Agarose gel showing the partial genotypic data of SSR-Marker Xtxp204 screening with C23. The susceptible lines marked as "A" are at the upper band, whereas resistant lines "B" are at the lower band. "H" stands for heterozygous, "-" is missing data. 33

Fig 2-2 Sample of collected genotypic data of primers sb06jh7, sb06jh11, sb06jh12 for QTL mapping construction. Data format was required by software instruction. 34

Fig 2-3 Twelve linkage groups (LGs) containing 77 polymorphic markers loci were mapped by using a set of sorghum $F_{2:3}$ lines obtained from the cross BTx623 (Susceptible line) and the resistant line Tx2783 (Resistance line) 35

Fig 2-4 The major QTL identified in this study. (A) The LOD score peak profile using SIM analysis for sugarcane aphid resistance identified on chromosome 6. (B) Genomic location of QTL and linked markers on sorghum chromosome 6....... 36

Fig 2-5 Cloning and characterization of candidate gene (Map based cloning) 39

Fig 2-6 (A) QTL region on the chromosome 6. 72 genes reside in this defined QTL region. (B) Genomic location of QTL and linked molecular markers on chromosome 6; Physical location of target gene flanked by two markers with an interval. (C) Physical location of three candidate genes on chromosome 6. (D) The phenotype of BTx623 and Tx2783 with 10 dpi (day post-infestation) 41

Chapter III. Molecular characterization of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance genes in sorghum [*Sorghum bicolor* (L.) Moench]

Fig 3-1 Plant material preparation and infestation with sugarcane aphids: (A) seed germinating, (B) young seedlings grown in greenhouse, (C) Infestation with freshly

prepared aphids, and (D) co-cultivation of host plants with aphids 54

Fig 3-2 The protocol for developing transgenic sorghum plants carrying the SCA resistance gene. (A) Surface sterilization of the immature caryopsis (collected around 10-14 days after pollination). (B) Further sterilization of immature caryopsis in the conical flask with 50% bleach + few drops of tween 20. (C) Isolating immature embryos the immature caryopsis. (D) Immature embryo heat treatment and Agrobacterium infection. (E) Immature embryos on co-cultivation medium. (F) Callus induction from the initial explants. (G) Shoot regeneration. (H) Rooting, transformation of regenerated shoots to rooting medium. (I) Hardening, regenerated shoots with healthy roots are transferred to growth chamber 63

Fig 3-3 Expression of three candidate genes at 7 dpi in two parents with treatments (infestation and non-infestation). BTxI-BTxN stands for comparison of infested BTx623 and non-infested BTx623; TxI-TxN stands for comparison of infested Tx2783 and non-infested Tx2783; TxN-BTxN stands for comparison of non-infested Tx2783 and non-infested BTx623; TxI-BTxI stands for infested Tx2783 and infested BTx623 64

Fig 3-4 Interpretation of Sequencing Chromatograms of 9bp deletion within BTx623 compared to the DNA sequence in the resistant line Tx2783............................65

Fig 3-5 Sequence annotation was performed for gene molecular analysis. Sequence comparison of two parents BTx623 and Tx2783 and two progenies R193 and S194. Bold indicates the 9bp deletion region in susceptible lines, BTx623 and S194 .65

Fig 3-6 DNA amplification in SCA resistant and susceptible lines. PCR products were amplified using primers flanking the mutation region. P_1 is BTx623 SCA susceptible parent; P_2 is Tx2783 SCA resistant line; S_1, S_2, S_3 are $F_{2:3}$ susceptible progenies; R_1, R_2, R_3 are $F_{2:3}$ resistant progenies 66

Fig 3-7 Licor gel analysis of amplification products in 60 randomly selected lines at the F_6 generation, due to the deletion in DNA sequence, in susceptible lines (P_1 is susceptible parent BTx623, P_2 is resistance parent Tx2783).............................67

Fig 3-8 (A) Non-infested resistance and susceptible lines at 9dpi. (B) Non-infested resistance and susceptible lines without infestation at 12dpi. (C) Infested resistance and susceptible lines at 9dpi. (D) Infested resistance and susceptible lines at 12dpi. The treatments included three seedlings of each line (BTx623-noninfested, Tx2783-noninfested, R193-noninfested, S194-noninfested, BTx623-infested, Tx2783-infested, R193-infested, and S194-infested)...68

Fig 3-9 Seedling height increase of four different lines at 8 different time-points. The four lines were BTx623 (susceptible) and Tx2783 (resistance) and two progenies R193 (resistance) and S194 (susceptible)..69

Fig 3-10 Expression of SCA resistance gene in sorghum resistance and susceptible lines at different stages (0 dpi, 0.5dpi, 1dpi, 2dpi, 3dpi, 6dpi, and 9dpi), dpi stands for days post-infestation. R193I stands for infested resistance line 193, S194I for infested susceptible line 194, TxI for infested Tx2783, BTxI for infested BTx623. Data were the mean values and the standard deviation of more than three biological repeats...70

Fig 3-11 (A) Agarose gel showing the amplification of the candidate gene Sb1650 by PCR. (B) Agarose gel showing the products of digestion of plasmid pFGC161 and the Sb1650 (insert), and the ligation of both insert and vector. (C) Bacterial colonies carried the recombinant gene construct grown in the petri dish with LB medium+ Kanamycin. (D) PCR-confirmation of the selected colonies that carry the binary vector with the Sb1650 71

CHAPTER I

Introduction

Sorghum

Sorghum [*Sorghum bicolor* (L.) Moench], is a self-pollinated C4 species with a diploid genome (2n = 20). It is a multipurpose drought and heat tolerant C4 plant that is one of the top five cereal crops worldwide. Recently, it has been developing into a premier bioenergy crop in the U.S. because of the high sugar level and high biomass yield. Sorghum is an ancient cereal grain, which was found in Southern Egypt 8000 years ago. It was domesticated in Ethiopia and Sudan and from there moved throughout all of Africa, where it remains an important staple food. A variety of sorghum species are grown in Central and South America, Asia, and Australia, especially in dry areas as it is generally grown as a rain fed crop. Sorghum crops have become an important source of food, fiber, animal feed and fodder, and a favorable energy crop for biofuel production (Paterson et al., 2009). In the U.S., sorghum is ranked as the third most important cereal crop and claims second highest production worldwide. Likewise, the United States is the world's largest producer of grain sorghum, having planted 5.6 million acres and produced 480 million bushels in 2016 (www.usda.gov/nass 2018). Also, the United States is the number one sorghum exporter in the world market, with a share of 70% of world trade and the chief importer in 2016/2017 was China (https://grains.org/sorghum 2018).

Sugarcane aphid

The sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)], also recently known as the white sugarcane aphid, is a new invasive pest damaging sorghum species in North America. It has become one of the most important insect pests of sorghum in the southern United States and Mexico over the past years. Sugarcane aphid has been a pest of sugarcane and sorghum outside of North America, such as in Africa, Australia, and South America. Sugarcane aphid has been found on different host plants, including Johnsongrass, shattercane, sorghum-sudangrass, sudangrass, forage sorghum, and grain sorghum. It was first documented in the United States on sugarcane in Florida in 1977 and Louisiana in 1999, but did not become a pest of sorghum until 2013 (Bowling et al., 2016), after which it spread rapidly across the sorghum-producing areas, including many counties in Louisiana, Oklahoma and one county in Mississippi. By 2015, it was present in Puerto Rico and in all sorghum-producing states in the South from Texas to Florida, reaching north from Kansas to North Carolina (http://www.sorghumcheckoff.com/).

Identification and biology of sugarcane aphid

The Sugarcane aphid can either be wingless or winged, and body color ranges from gray to tan to light yellow. Wingless aphids are pale yellow to white in color with dark cornicles (tailpipes) located at their rear end (Bowling et al., 2016). Winged aphids are dark yellow in color and are generally developed by stress conditions. The sugarcane aphid is an anholocyclic, parthenogenic, viviparous species, which means that it feeds on its annual hosts (sorghum species) only in the spring and summer, and the same hosts that persist through the fall and winter months (Johnsongrass but also remnant sorghum). All aphids are female and

produce live young asexually in North America, with the exception of one report of egg production from female aphids collected from three Mexican states (Peña-Martinez et al., 2016). Sexual forms have previously been reported on sorghum in China, India, and Japan (Wang 1961; Yadava 1966; Setokuchi 1975; David and Sandhu, 1976). Based on the population density increases and food quality declines, a proportion will develop wings as adults, which enables them to fly to other nearby fields or to be carried by wind, potentially moving them long distances. Most Sugarcane aphids are female and give birth to 1 to 3 live, pregnant offspring daily. The aphid has four nymphal stadia (non-winged nymphs) (Blackman and Eastop et al., 1984). It takes about 4–12 d for development from birth to adult, depending on temperature (Chang et al., 1982). Adult longevity ranges from 10–37 d (Chang et al., 1982; Singh et al., 2004), may be apterous or alate, with a reproductive potential ranging from 34 to 96 nymphs per female depending on temperature and nutrition (Chang et al., 1982, Singh et al., 2004). Population increase is influenced by temperature and rainfall pattern (Chang et al., 1982). Sugarcane aphid population growth increases rapidly during warm, dry climatic conditions (Singh et al., 2004), which are common in much of the sorghum-producing regions of North America.

Damage to sorghum by sugarcane aphid

The damage on sorghum by Sugarcane aphids is caused in several ways, depending on the stage of sorghum plant growth when aphids infests the crop. The Sugarcane aphid causes damage by sucking plant sap from the phloem tissue of sorghum and then secretion of a large amount of honeydew over the surface of the plant. Normally, Sugarcane aphid feed mainly on the underside of leaves and the stem, then move to the upper leaves before creeping to the

grain sorghum head. Plant damage from Sugarcane aphid results from a combination of direct loss of plant nutrients and sugars during feeding, which can be aggravated by plant water stress, and reduction in photosynthetic efficiency due to sooty mold buildup from honeydew excreted by Sugarcane aphids (Singh et al., 2004). At the beginning of the infestation, the leaves turn yellow to red or brown on both sides, but infestation with heavy populations of Sugarcane aphid can kill young plants. Later infestation causes the loss of the nutrient from the plant; resulted in severe damage to the plant and significant yield loss. Symptoms of aphid damage include purpling of young plants, which can lead to stunting, chlorosis, and necrosis of maturing leaves (Singh et al., 2004). At the stage of sorghum harvest, sugarcane aphid infestation can cause harvesting issues because of honeydew on the upper leaves and with aphids in the sorghum head.

Host plant resistance to sugarcane aphid

Host plant resistance is defined as a plant's ability to resist damaging insect invasions. The effectiveness of host plant resistance can vary by location because insect pests of the same species can vary somewhat by location. There has been substantial use of sorghum hybrids with traits resistant to greenbug in North America (Michels and Burd, 2007). Research has begun to evaluate parental lines and commercially available hybrids for resistance to sugarcane aphid in North America, adding to existing international efforts (Singh et al., 2004). Recently, sorghum parental types SC110 and SC170, Tx2783, and Texas A&M sorghum lines and hybrids Tx2783, Tx3408, Tx3409, B11070, B11070, AB11055-WF1-CS1/RTx436 and AB11055-WF1-CS1/RTx437 have shown high levels of resistance to sugarcane aphid in greenhouse and field tests (Armstrong et al., 2015; Mbulwe et al., 2016).

Those characters that enable a plant to avoid, tolerate or recover from attacks of insects under conditions that would cause greater injury to other plants of the same species (Painter, 1951), we named them host plant resistance (HPR), can be an effective defense against an insect pest. HPR to insects is an effective, economical, and environment-friendly method of pest control. The most attractive feature of HPR is that farmers virtually do not need any skill in application techniques, and there is no cash investment by the resource-poor farmers.

Interaction between sugarcane aphid and sorghum

Sugarcane aphids were first discovered on sorghum in the United States in Florida as early as 1922 (Wilbrink, 1922), and later confirmed by Denmark (1988), although they also have a history of infesting sugarcane, Saccharum officinarum (L.), in Florida (Mead, 1978; Denmark, 1988) and Louisiana (Hall, 1987; White et al., 2001). In 2013, they were found feeding on sorghum for the first time in North America. Two years later, they spread rapidly across most of the sorghum-producing areas of the United States and Mexico. By 2015, they were present in Puerto Rico and in every sorghum-producing state in the South from Texas to Florida, reaching north all the way from Kansas to North Carolina (Villanueva et al., 2014). Since Sugarcane aphids have become a major issue for farmers throughout the Midwest in the past 3 years, scientists have been working on the sugarcane aphid and their occurrence on sorghum, I cited several significant kinds of literature here to introduce the situations of sugarcane aphid on sorghum, trying to identify resistant materials.

Brewer (2013) indicated that an outbreak of an invasive aphid was discovered damaging grain sorghum in Texas and neighboring states in 2013. They also found that this outbreak caused severe damage, with producers and crop consultants estimating 25%-50% yield loss

and total yield loss in some unprotected fields. In the same year, Hari C. Sharma (2013) discovered that nine genotypes (Line 61510, ICSV 12001, ICSV 12002, ICSV 12003, ICSV 12004, ICSV 12005, SLR 41, PU 10-1 and DJ 6514) revealed moderate levels of resistance to sugarcane aphid. These genotypes also showed a lower rate of aphid multiplication in clip cage experiments and leaf disc assays. The rates of aphid multiplication were lower on the genotypes IS 21807, IS 40615, IS 40616 and IS 40618 than on the susceptible control. Additionally, Hari C. Sharma (2013) identified seven lines (ICSB 215, ICSB 323, ICSB 724, ICSR 165, ICSV 12001, ICSV 12004 and IS 40615) with moderate levels of resistance to aphid damage, and although the genotypes RSV 1211, RS 29, RSV 1338, EC 8- 2, PU 10-1, IS 40617 and ICSB 695 showed a susceptible reaction to aphid damage, they suffered relatively low loss in grain yield, suggesting that these lines have tolerance to aphid damage.

Scott Armstrong (2015) tested 20 sorghum parental lines for phenotypic damage in the field, and 9 of them showed good resistance to the sugarcane aphid. The greenbug-resistant parental line RTx2783 that is resistant to greenbug biotypes C and E was resistant to sugarcane aphid in both greenhouse and field tests, while PI 550607 greenbug resistant to biotypes B, C, and E was highly susceptible. PI 550610 that is greenbug resistant to biotypes B, C and E maintained moderate resistance to the sugarcane aphid, while greenbug-resistant PI 264453 was highly susceptible to the sugarcane aphid. Two lines and two hybrids from the Texas A&M breeding program B11070, AB11055-WF1-CS1/RTx436 and AB11055-WF1-CS1/RTx437 were highly resistant to sugarcane aphid, as were parental types SC110, SC170, and South African lines Ent62/SADC, (Macia/TAM428)-LL9, and (SV1*Sima/IS23250)-LG15. Overall, the information that was collected from those sorghum lines can be used to

breed for sugarcane aphid resistance high grain yield potential, and also in breeding for sugarcane aphid resistance in sorghum with adaptation to the post-rainy season.

Elliott (2015) conducted research to determine if damage caused by sugarcane aphid to sorghum plants in the field of grain sorghum could be detected using multispectral remote sensing from a fixed wing aircraft. This study was conducted in commercial grain sorghum fields in the Texas Gulf Coast region in June 2014, and it discovered that the correlation between plant injury rating and plant growth stage was positive and significant indicating that plant injury from sugarcane aphid increased as plants matured.

Lloyd Mbulwe (2015) announced that Texas A&M AgriLife Research developed and released sorghum germplasm lines Tx3408 (ATx3408, Reg. No. GP-838, PI 675360; BTx3408, Reg. No. GP-839, PI 675361) and Tx3409 (ATx3409, Reg. No. GP-840, PI 675362; BTx3409, Reg. No. GP-841, PI 675363) with high levels of tolerance to the sugarcane aphid. Both of those two lines were developed from intentional crosses by using the pedigree method of plant breeding. This release provided the sorghum industry with a source of tolerance to the sugarcane aphid in an elite and diverse genetic background.

Those papers focus on the damage and losses caused by the sugarcane aphid (Melanaphis sacchari) on sorghum production fields in the United States. The economic impact of the sugarcane aphid on sorghum production were discussed and various management recommendations were presented. In order to avoid the damage caused by sugarcane aphid, pesticides have become the prime option for scientists and farmers, nevertheless, when applying pesticide, it will cause environment contaminated, especially to those non-target species, also, it requires time and effort to spray, which cost the farmers a lot of money.

Therefore, the better method is to develop the host plant resistance to sugarcane aphid, also, this is the best by understanding the genetics associated with resistance.

Development of molecular markers and QTL mapping

The sorghum genome contains ~730-megabase pairs of DNA, has a haploid chromosome number of 10 and it is a genetic model for C4 grasses due to its relatively small genome (Paterson, 2009). Reduced sequencing costs and technological advances have since enabled the sequencing and assembly of additional grass genomes, including Brachypodium distachyon (Vogel et al., 2010), corn (Schnable et al., 2009), foxtail millet (Bennetzen et al., 2012; Zhang et al., 2012), wheat (Brenchley et al., 2012), barley (Collaborative Cross Consortium, 2012b), and the desiccation tolerant Oropetium thomaeum (VanBuren et al., 2015). In addition, the genomes of 49 additional sorghum genotypes have been sequenced and assembled through alignment to the sorghum reference genome produced in 2009 (Evans et al., 2013; Mace et al., 2013; Zheng et al., 2011). Its genome size is one-third of the maize genome and 75% larger than rice (Paterson, 2009). The lower level of genome duplication and small genome size of sorghum has driven many researchers to harness genomic resources for structural and functional analyses (Paterson 2008).

Sorghum linkage maps have been developed in the last decade due to the availability of abundant polymorphic DNA markers. Linkage mapping in sorghum is relatively easy because of the diploid nature of the sorghum genome (Paterson 2008). The sequencing of whole genomes of sorghum has led to the rapid development of DNA markers followed by quantitative trait loci (QTL) mapping.

The DNA markers can be used to detect the existence of allelic variation in the genes underlying these traits. By using DNA markers to assist in plant breeding, efficiency and

precision could be greatly increased. The use of DNA markers in plant breeding is called marker-assisted selection (MAS) and is a component of the new discipline of 'molecular breeding'. DNA markers have enormous potential to improve the efficiency and precision of conventional plant breeding via marker-assisted selection (MAS), also molecular markers are neutral to the environment, cost-effective and time-saving. The large number of quantitative trait loci (QTLs) mapping studies for diverse crops species have provided an abundance of DNA marker–trait associations. The molecular marker technology was developed in the 1980s. There were different types of molecular markers that have been developed and advancement in sequencing technologies have associated crop improvement. To discover the knowledge about molecular markers, several reviews have been published in the last three decades (Stuber, 1992).

Molecular markers are neutral to the environment, cost-effective and time-saving, which have accelerated the efforts to elucidate the entire architecture of the sorghum genome. The most commonly used markers in major cereals are named simple sequence repeats (SSRs) or microsatellites (Gupta et al. 1999; Gupta and Varshney 2000) because they are the highly reliable, co-dominant in inheritance, relatively simple and cheap to use and generally highly polymorphic.

Mapping in sorghum began in 1990, and hundreds of published papers have exploited that mapping has been used in sorghum breeding for the last three decades. These published maps have included multiple markers, such as the simple sequence repeat (SSR) markers, the restriction fragment length polymorphism (RFLP)markers, the amplified fragment length polymorphism (AFLP) markers and in a few cases, the randomly amplified polymorphic DNA (RAPD) markers. There are several other works in the last ten years contributing to

sorghum maps, prominent among them is the consensus map developed by Mace (2008). Sorghum molecular marker maps, saturated with several kinds of markers, play a pivotal role in tagging loci that govern quantitative traits like drought and insect resistance. These molecular markers facilitate the dissection of the genetic mechanisms underlying insect resistance in sorghum (Yencho et al. 2000).

By using statistical methods such as single-marker analysis or interval mapping to detect associations between DNA markers and phenotypic data, genes or QTLs can be detected in relation to a linkage map (Kearsey 1998). The identification of QTLs using DNA markers was a major breakthrough in the characterization of quantitative traits (Paterson et al. 1988). In the past years, many published papers have discovered numerous DNA markers linked to genes or QTLs.

Most agricultural characters of economic interest are polygenic and quantitative in nature and controlled by many genes on the chromosome of species. The chromosomal regions that have genes for these quantitative characters are referred to as QTLs. QTL mapping is a method that is utilized to locate the genes that controls the traits of interest by using molecular markers. For a QTL experiment, molecular markers play a key role and are considered an idea tool (Angaji, 2009).

The main steps of QTL mapping include the selection of two parents that have allelic variations such as resistance and susceptible to the pest. Normally, near-isogenic lines (NILs), double haploid (DHs), backcrosses (BCs), recombinant inbred lines (RILs) and F_2 populations could be used as the mapping population (Paterson, 1996). Ideally 50-250 individuals are selected for a mapping population, but a larger size of the mapping population are needed for high-resolution and fine mapping (Mohan et al., 1997; Dhingani et al., 2015).

Different types of markers such as RFLP, AFLP, ISSR, SSR, ESTs, DArT and SNPs have been widely used for the construction of linkage map in plants (Semagn et al., 2006). The construction of a linkage map is used to discover the position and relative genetic distance between markers (Paterson et al., 1996), and QTL mapping is according to the marker segregation via chromosome recombination during meiosis, in other words, those markers which are tightly linked with each other will be transferred together more frequently during recombination as compared to those which are away from each other. The actual distance and relative order of markers are normally calculated by using odds ratio, this value is named LOD or logarithm of odds (Risch et al., 1992). In interval mapping, each locus is considered one at a time and the logarithm of the odds ratio (LOD score) is calculated for the model that the given locus is a true QTL. Based on previous research, the LOD value of >3 is considered for the construction of the linkage maps (Collard et al., 2005). There are several published papers that have reported the most important methods developed for QTL detection, such as single-marker analysis (Tanksley, 1993), simple interval analysis (Lander and Botstein, 1989), and composite interval analysis (Luciano et al., 2012). Also, several published papers have reported some important statistical programs that are widely used in QTL mapping such as in R (Broman et al., 2003), QTLNetwork (Yang et al., 2008), PLABQTL (Utz and Melchinger, 1996), QGENE (Nelson, 1997) and MapChart (Voorrips, 2002).

In the study of QTL mapping, it is essential to validate the specific QTL. For this purpose, populations are developed by crossing different parents in order to check the presence of a particular QTL in other populations with the different genetic backgrounds. NILs are commonly used for the validation of QTLs (Collard et al., 2005). NILs have been reported and used to precisely evaluate the effect of different pollen sterility loci in rice (Wu et al.,

2015), Wu used two NILs and their F_1 hybrids that are heterozygous at each of their loci to distinguish the effect of each factor of pollen sterility and gene expression profiles. The confirmation of QTLs provides the information about the marker to be used or not for MAS (Ogbonnaya et al., 2001). After QTL validation, the candidate genes can be selected and sequenced to design specific primers for PCR amplification of the candidate gene. And molecular cloning used to detect the genetic basis responsible for a mutant phenotype.

Significance

Sorghum [*Sorghum bicolor* (L.) Moench] is a multipurpose drought and heat tolerant C4 plant. Sorghum is also a subject of plant genomics research based on its importance as the fifth most important cereal crop worldwide, a biofuels crop of high and growing importance, a progenitor of one of the world's most noxious weeds, and a botanical model for many tropical kinds of grass with complex genomes. Recently, it has been developing as a premier bioenergy crop in the U.S. because of the high sugar level and high biomass yield. Sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)], recently also known as the white sugarcane aphid, is a new invasive pest damaging sorghum species. It has become one of the most important insect pests of sorghum in the southern United States and Mexico over the last five years. Sorghum injury results from a combination of loss of plant nutrients, intensifying plant water stress, induced leaf chlorosis, and reduction in photosynthetic productivity. Eventually, sorghum plants will die with a heavy infestation of sugarcane aphids (Singh et al., 2004). The loss of sorghum crop incurred due to sugarcane aphid damage ranks in the hundreds of millions of dollars to farmers. In order to protect sorghum plants and avoid the loss from the damage of sugarcane aphid, systemic pesticides are used to

control the insect. Nevertheless, when applying pesticide, it is often associated with a high cost, development of pesticide resistance and has environmentally unfriendly effects, especially apparent in a drastic reduction of beneficial insects. The undesirable effects of systemic insecticides should lead to increased interest in host plant resistance against sugarcane aphid infestation as a better alternative to chemical control measures and as an effective means of integrated pest management practices (Smith, 2004). To the best of my knowledge, no research has been reported on the breeding processes of sugarcane aphid resistant sorghum cultivars. Accordingly, there is a need to identify better resistant sources from a wide collection of germplasm, which offers a durable resistance. Better sorghum breeding strategies involving known sources of sugarcane aphid resistance would facilitate deploying of resistant cultivars into sugarcane aphid management programs. Molecular breeding takes advantage of DNA markers that are neutral to the environment and cost and time effective for development of sugarcane aphid resistant sorghum cultivars.

The DNA markers developed in this study can be used for marker-assisted selection to speed up the breeding process of sugarcane aphid resistant sorghum cultivars. Identification of sugarcane aphid resistance QTLs will also facilitate molecular cloning of the sugarcane aphid resistance genes from the sorghum plant. The newly discovered sugarcane aphid resistance genes will be a great contribution to the genetic improvement of sugarcane aphid resistance in sorghum.

QTL mapping has been detecting genes that control traits of interest in past years (Mohan et al., 1997). It’s very valuable for the genome-wide scan for QTL detection in crop plants. Pest insects are a huge problem that cause economic loss in agriculture and genes responsible for

the generation of resistance to these insects can be detected by QTL mapping (Young et al., 1992; Bonierbale et al., 1994).

This study will improve the understanding of plant genetics associated with resistance by illuminating the functions of genes expressed during the sugarcane aphid-sorghum interaction. The results from the expression studies will facilitate the identification of differentially-expressed genes between resistant and susceptible sorghum plants. This study will also identify molecular markers linked to sugarcane aphid resistance which can be further combined into the marker-assisted selection and map-based cloning of sugarcane aphid resistant genes. The results of these efforts have a larger impact on sorghum economics and the agricultural environment. Sorghum farmers will be benefited by the decreased costs incurred on pesticides, as well as from increased yield.

Objectives

Based on the above background, I performed several experiments, which are explained in detail in the following two chapters. The first project is detailed in chapter II, with a comprehensive objective of identifying quantitative trait loci (QTL) in an $F_{2:3}$ mapping population of sorghum developed by crossing susceptible cultivar BTx623 and sugarcane aphid resistant source Tx2783. We identified major loci for sugarcane aphid resistance on sorghum chromosome 6. These results were supported by previous studies which identified major loci for greenbug resistance on chromosome 9 by Wu (Wu, 2008) and Punnuri (2012). These loci accounted for a large portion of phenotypic variation observed for sugarcane aphid resistance. This project identified three candidates gene-based markers associated with insect resistance and R-gene mediated resistance. We wanted to confirm the expression of candidate genes in sorghum against sugarcane aphid. Our second project is detailed in

Chapter III. We aim to identify the first sugarcane aphid resistance gene and developed DNA markers for this resistance gene to offer both the genetic information and an efficient tool for genomics-assisted breeding of new hybrids or cultivars with genetic resistance to sugarcane aphids in sorghum.

References

Angaji SA. QTL mapping: A few key points. International Journal of Applied Research in Natural Products. 2009 Jun 15;2(2):1-3.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J. Reference genome sequence of the model plant Setaria. Nature Biotechnology. 2012 Jun;30(6):555.

Blackman RL, Eastop VF. Aphids on the world crops. John Wiley and Sons, Chichester etc. 1984:466.

Bonierbale MW, Plaisted RL, Pineda O, Tanksley SD. QTL analysis of trichome-mediated insect resistance in potato. Theoretical and applied genetics. 1994 Mar 1;87(8):973-87.

Bowling RD, Brewer MJ, Kerns DL, Gordy J, Seiter N, Elliott NE, Buntin GD, Way MO, Royer TA, Biles S, Maxson E. Sugarcane aphid (Hemiptera: Aphididae): a new pest on sorghum in North America. Journal of Integrated Pest Management. 2016 Jan 1;7(1).

Brenchley R, Spannagl M, Pfeifer M, Barker GL, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D, Kay S. Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature. 2012 Nov;491(7426):705.

Brewer M, Way M, Armstrong S, Biles S, Sekula D, Swart J, Crumley C, Knutson A, Villanueva R, Parker R, Odvody G. Outbreak of sorghum/sugarcane aphid on sorghum: first detections, distribution, and notes on management. In Proc. of 25th Texas Plant Protection Conf. Program 2013 Dec 11.

Broman KW, Wu H, Sen Ś, Churchill GA. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 2003 May 1;19(7):889-90.

Chang CP, Fang MN, Tseng HY. Studies on the life history and varietal resistance in grain sorghum aphid, Melanaphis sacchari Zehntner in central Taiwan. Chinese Journal of Entomology. 1982; 2:70-81.

Collaborative Cross Consortium. The genome architecture of the Collaborative Cross mouse genetic reference population. Genetics. 2012 Feb 1;190(2):389-401.

Collard BC, Jahufer MZ, Brouwer JB, Pang EC. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica. 2005 Jan 1;142(1-2):169-96.

David SK, Sandhu GS. New oviparous morph of Melanaphis sacchari (Zehntner) on sorghum. Entomologist's Record. 1976;88(1):28-9.

Denmark HA. Sugarcane aphids in Florida (Homoptera: Aphididae). Sugarcane aphids in Florida (Homoptera: Aphididae). 1988(302).

Dhingani RM, Umrania VV, Tomar RS, Parakhia MV, Golakiya BA. Introduction to QTL mapping in plants. Annals of Plant Sciences. 2015 Apr 30;4(04):1072-9.

Elliott NC, Backoulou GF, Brewer MJ, Giles KL. NDVI to detect sugarcane aphid injury to grain sorghum. Journal of Economic Entomology. 2015 May 5;108(3):1452-5.

Evans J, McCormick RF, Morishige D, Olson SN, Weers B, Hilley J, Klein P, Rooney W, Mullet J. Extensive variation in the density and distribution of DNA polymorphism in sorghum genomes. PLoS One. 2013 Nov 12;8(11): e79192.

Gupta PK, Varshney RK, Sharma PC, Ramesh B. Molecular markers and their applications in wheat breeding. Plant breeding. 1999 Oct;118(5):369-90.

Gupta PK, Varshney RK. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica. 2000 Jun 1;113(3):163-85.

Hall DG. The sugarcane aphid, Melanaphis sacchari (Zehntner), in Florida. Journal of the American Society of Sugarcane Technologists. 1987 Nov; 7:26-9.

Kearsey MJ, Farquhar AG. QTL analysis in plants; where are we now? Heredity. 1998 Feb;80(2):137.

Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 1989 Jan 1;121(1):185-99.

Luciano Da Costa ES, Wang S, Zeng ZB. Composite interval mapping and multiple interval mapping: procedures and guidelines for using Windows QTL Cartographer. In Quantitative Trait Loci (QTL) 2012 (pp. 75-119). Humana Press.

Mace ES, Tai S, Gilding EK, Li Y, Prentis PJ, Bian L, Campbell BC, Hu W, Innes DJ, Han X, Cruickshank A. Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. Nature Communications. 2013 Aug 27; 4:2320.

Mace ES, Xia L, Jordan DR, Halloran K, Parh DK, Huttner E, Wenzl P, Kilian A. DArT markers: diversity analyses and mapping in Sorghum bicolor. Bmc Genomics. 2008 Dec;9(1):26.

Mbulwe L, Peterson GC, Scott-Armstrong J, Rooney WL. Registration of sorghum germplasm Tx3408 and Tx3409 with tolerance to sugarcane aphid [Melanaphis sacchari (Zehntner)]. Journal of Plant Registrations. 2016 Jan 1;10(1):51-6.
Mbulwe L, Peterson GC, Scott-Armstrong J, Rooney WL. Registration of sorghum germplasm Tx3408 and Tx3409 with tolerance to sugarcane aphid [Melanaphis sacchari (Zehntner)]. Journal of Plant Registrations. 2016 Jan 1;10(1):51-6.
Mead FW. Sugarcane aphid, Melanaphis sacchari (Zehntner)—Florida—new continental United States record. Cooperative Plant Pest Report. 1978;3(34):475.

Michels GJ, Burd JD. IPM case studies: sorghum. Aphids as crop pests. (Eds HF Emden, R van Harrington) pp. 2007 Sep 1:627-37.

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T. Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding. 1997 Apr 1;3(2):87-103.

Nelson JC. QGENE: software for marker-based genomic analysis and breeding. Molecular breeding. 1997 Jun 1;3(3):239-45.

Ogbonnaya FC, Subrahmanyam NC, Moullet O, De Majnik J, Eagles HA, Brown JS, Eastwood RF, Kollmorgen J, Appels R, Lagudah ES. Diagnostic DNA markers for cereal cyst nematode resistance in bread wheat. Australian Journal of Agricultural Research. 2001;52(12):1367-74.

Painter RH. Insect resistance in crop plants. The Macmillan Company New York; 1951.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J. The Sorghum bicolor genome and the diversification of grasses. Nature. 2009 Jan;457(7229):551.

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature. 1988 Oct;335(6192):721.

Paterson AH. Genomics of sorghum. International Journal of Plant Genomics. 2008;2008.

Paterson AH. Making genetic maps. Genome mapping in plants. 1996;194.

Peña-Martínez R, Muñoz-Viveros AL, Bujanos-Muñiz R, Luévano-Borroel J, Tamayo-Mejia F, Cortez-Mondaca E. Formas sexuales del complejo pulgón amarillo del sorgo, Melanaphis sacchari/sorghi en México. Southwestern Entomologist. 2016 Mar;41(1):127-32.

Punnuri S, Huang Y, Steets J, Wu Y. Developing new markers and QTL mapping for greenbug resistance in sorghum [Sorghum bicolor (L.) Moench]. Euphytica. 2013 May 1;191(2):191-203.

Risch N, Giuffra L. Model misspecification and multipoint linkage analysis. Human Heredity. 1992;42(1):77-92.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P. The B73 maize genome: complexity, diversity, and dynamics. Science. 2009 Nov 20;326(5956):1112-5.

Scott Armstrong J, Rooney WL, Peterson GC, Villenueva RT, Brewer MJ, Sekula-Ortiz D. Sugarcane aphid (Hemiptera: Aphididae): host range and sorghum resistance including cross-resistance from greenbug sources. Journal of Economic Entomology. 2015 Jan 30;108(2):576-82.

Semagn K, Bjørnstad Å, Skinnes H, Marøy AG, Tarkegne Y, William M. Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population. Genome. 2006 May 1;49(5):545-55.

Setokuchi O. Ecology of Longiunguis sacchari (Zehentner)(Aphididae) infesting sorghums. III. Occurrence in fields. InProceedings of the Association for Plant Protection of Kyushu 1975 (Vol. 21, pp. 8-10).

Sharma HC, Sharma SP, Munghate RS. Phenotyping for resistance to the sugarcane aphid Melanaphis sacchari (Hemiptera: Aphididae) in Sorghum bicolor (Poaceae). International Journal of Tropical Insect Science. 2013 Dec;33(4):227-38.

Singh BU, Padmaja PG, Seetharama N. Biology and management of the sugarcane aphid, Melanaphis sacchari (Zehntner) (Homoptera: Aphididae), in sorghum: a review. Crop Protection. 2004 Sep 1;23(9):739-55.

Singh BU, Padmaja PG, Seetharama N. Biology and management of the sugarcane aphid, Melanaphis sacchari (Zehntner)(Homoptera: Aphididae), in sorghum: a review. Crop Protection. 2004 Sep 1;23(9):739-55.

Smith MA, Wise IL, Lamb RJ. Sex ratios of Sitodiplosis mosellana (Diptera: Cecidomyiidae): implications for pest management in wheat (Poaceae). Bulletin of Entomological Research. 2004 Dec;94(6):569-75.

Stuber CW. Biochemical and molecular markers in plant breeding. Plant Breed. Rev. 1992 Apr 23; 9:37-61.

Tanksley SD. Mapping polygenes. Annual review of genetics. 1993 Dec;27(1):205-33.

Utz HF, Melchinger AE. PLABQTL: a program for composite interval mapping of QTL. J Quant Trait Loci. 1996;2(1):1-5.

VanBuren R, Bryant D, Edger PP, Tang H, Burgess D, Challabathula D, Spittle K, Hall R, Gu J, Lyons E, Freeling M. Single-molecule sequencing of the desiccation-tolerant grass Oropetium thomaeum. Nature. 2015 Nov;527(7579):508.

Villanueva RT, Brewer M, Way MO, Biles S, Sekula D, Bynum E, Swart J, Crumley C, Knutson A, Porter P, Parker R. Sugarcane aphid: a new pest of sorghum. Texas A&M AgriLife Extension publication ENTO-03. 2014.

Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon-Smith M. and Lail K. (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature. 463, 763–768.

Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. Journal of Heredity. 2002 Jan 1;93(1):77-8.

Wang YS, Pao ST, Kwei CM, Chang GS, Chu HF. Studies on the sorghum aphid, Aphis sacchari Zehntner. Acta Entomologica Sinica. 1961; 10:363-80.

White WH, Reagan TE, Hall DG. Melanaphis sacchari (Homoptera: Aphididae), a sugarcane pest new to Louisiana. Florida Entomologist. 2001;84(3):435-6.

Wilbrink G. An investigation on spread of the mosaic disease of sugarcane by aphids. Medid Procfst. Java Suikerind. 1922; 10:413-56.

Wu JW, Shahid MQ, Chen L, Chen ZX, Wang L, Liu XD, Lu YG. Polyploidy enhances F1 pollen sterility loci interactions that increase meiosis abnormalities and pollen sterility in autotetraploid rice. Plant physiology. 2015 Jan 1: pp-00791.

Wu Y, Huang Y. Molecular mapping of QTLs for resistance to the greenbug Schizaphis graminum (Rondani) in Sorghum bicolor (Moench). Theoretical and Applied Genetics. 2008 Jun 1;117(1):117.

Yadava BN, Sircar DC. Secular Landgrants of the Post-Gupta Period and Some Aspects of the Growth of the Feudal Complex in North India. Land System and Feudalism in Ancient India, University of Calcutta. 1966:72-94.

Yang J, Hu C, Hu H, Yu R, Xia Z, Ye X, Zhu J. QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. Bioinformatics. 2008 Jan 17;24(5):721-3.

Yencho GC, Cohen MB, Byrne PF. Applications of tagging and mapping insect resistance loci in plants. Annual Review of Entomology. 2000 Jan;45(1):393-422.

Young ND, Kumar L, Menancio-Hautea D, Danesh D, Talekar NS, Shanmugasundarum S, Kim DH. RFLP mapping of a major bruchid resistance gene in mungbean (Vigna radiata, L. Wilczek). Theoretical and Applied Genetics. 1992 Sep 1;84(7-8):839-44.

Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, Xie M, Zeng P, Yue Z, Wang W, Tao Y. Genome sequence of foxtail millet (Setaria italica) provides insights into grass evolution and biofuel potential. Nature Biotechnology. 2012 Jun;30(6):549.

Zheng LY, Guo XS, He B, Sun LJ, Peng Y, Dong SS, Liu TF, Jiang S, Ramachandran S, Liu CM, Jing HC. Genome-wide patterns of genetic variation in sweet and grain sorghum (Sorghum bicolor). Genome Biology. 2011 Nov;12(11): R114.

CHAPTER II

QTL mapping of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance gene in Sorghum [*Sorghum bicolor* (L.) Moench]

Abstract

The sugarcane aphid [*Melanaphis sacchari* (Zehntner)] is a new invasive pest of sorghum that can cause severe damage to the plant and yield loss. It has become one of the most important insect pests of sorghum in the southern United States and Mexico since 2013. This study was conducted to identify the genomic region contributing to resistance to the sugarcane aphid in sorghum. An $F_{2:3}$ mapping population was developed by crossing an elite line BTx623 (susceptible, seed parent) with a resistant donor Tx2783 (sugarcane aphid resistant line, pollen parent). The population was tested and scored in the greenhouse with two parallel experiments, each consisting of three replications. Significant differences in resistance were observed between the two parental lines and their $F_{2:3}$ progenies in response to sugarcane aphid feeding at 7, 10, 14 and 21 days after infestation. Among 77 polymorphic markers (50 SSRs, 21 ESRs, and 6 designed markers) were identified and used for genotyping the entire population to develop the genetic linkage map. The newly developed linkage map included 12 groups with 46 markers, which spanned a total length of 781.4 cM on 7 chromosomes. A major QTL was identified and delimited to a 500kb region on the short arm of Chromosome 6, which

associated with the plant response to sugarcane aphid feeding damage. Moreover, the markers tagged to QTL regions can be used to improve the sorghum breeding program for sugarcane aphid resistance through marker-assisted selection (MAS) and map-based cloning.

Introduction

Sorghum [*Sorghum bicolor (L.) Moench*] is a multipurpose drought and heat tolerant C4 plant, the fifth most important cereal crop worldwide. Recently it has been developing as a premier bioenergy crop in the U.S. because of the high sugar level and high biomass yield. Sorghum is an ancient cereal grain, which was found in Southern Egypt 8000 years ago. It was domesticated in Ethiopia and Sudan and from there moved throughout all of Africa, where it remains an important staple food. A variety of sorghum species are grown in Central and South America, Asia, and Australia, especially in dry areas as it is generally grown as a rain-fed crop. Sorghum crops have become an important source of food, fiber, animal feed and fodder, and favorable energy crop for biofuel production.

Plants are constantly confronted by both abiotic and biotic stresses that cause serious damage and reduce their productivity. Among all the stresses, insect pests are the main biotic stress to plants. There are more than 150 insect species reported that can infest sorghum worldwide, causing damage from the seedling to maturity stage (Huang 2013; Padmaja and Aruna. 2019). Sorghum shoot fly, stem borer, greenbug, corn planthopper, aphids, and sorghum midge are the main pest of sorghum in the field, and some pests like grain weevil, moth, and flour beetles are stored pests that can infest the stored sorghum grain. Recently, an invasive aphid pest has been attacking grain sorghum in the southern U.S. named Sugarcane aphid (*Melanaphis Sacchari*) (SCA). At present, SCA has become the most important insect pest of sorghum in the Southern U.S. and Mexico (Bowling et al., 2016).

Sugarcane aphid has been found on different host plants, including Johnsongrass, shattercane, sorghum-sudangrass, sudangrass, forage sorghum, and grain sorghum. It was first documented in the United States on sugarcane in Florida in 1977 and Louisiana in 1999 but did not become a pest of sorghum until 2013 (Bowling et al., 2016), after which it spread rapidly across the sorghum-producing areas, including many counties in Louisiana, Oklahoma, and one county in Mississippi. By 2015, it was present in Puerto Rico and in every sorghum-producing state in the South from Texas to Florida, reaching north from Kansas to North Carolina (http://www.sorghumcheckoff.com/).

The SCA causes damage by sucking plant sap from the phloem tissue of sorghum and then secreting a large amount of honeydew over the surface of the plant. Normally, SCA feed mainly on the underside of leaves and the stem, then move to the upper leaves before creeping to the grain head. At the beginning of the infestation, the leaves turn yellow to red or brown on both sides, but infestation with heavy populations of SCA can kill young plants. Later infestation causes the loss of the nutrient from the plant; resulted in severe damage to the plant and significant yield loss. At the stage of sorghum harvest, sugarcane aphid infestation can cause harvesting issues as a result of honeydew on the upper leaves and with aphids in the sorghum head.

Host plant resistance is perceived as an eco-friendly, socially acceptable and effective component of integrated pest management in deploying resistant cultivars against insect damage (Bramel-Cox et al., 1986; Andrews et al., 1993; Sharma, 1993; Sharma and Ortiz, 2002; Smith, 2004). Therefore, host plant resistance is of prime importance in the arsenal of aphid-plant interaction. Multiple mechanisms are involved in plant defense response to aphid feeding through early signaling and differential expression of gene pathways.

Materials and Methods

Development of a mapping population

A mapping population of 190 $F_{2:3}$ families derived from a cross between the susceptible cultivar BTx623 and resistant donor Tx2783. BTx623 (PI 659985 MAP) is an elite, white seeded, inbred line with the pedigree BTx3197/SC170–6-4. It was developed by Dr. Fred Miller at Texas A&M University and published by Texas A&M University in 1976 (Miller, 1977). Tx2783 is highly resistant to the sugarcane aphid (Armstrong et al., 2015) and was released in 1981. It was developed by the Texas Agricultural Experiment Station as a source of resistance to greenbug biotype C and biotype E. It is derived from Capbam which was introduced from Russia and received from DeKalb Agricultural Research Inc. in 1971 (Peterson, 1984). We produced an F_2 population of 286 individuals by selfing a F_1 plant of two parents. The F_2 and their $F_{2:3}$ families were applied in corresponding genotyping and phenotyping experiments.

F_2 seeds from the OSU Field 2014 were planted with 1 seed per hole. They were screened with sugarcane aphids, rated "S" for susceptible or "R" for resistant, and sprayed with pesticide in hopes of them surviving. If they survived, they were either transplanted to the greenhouse 2 beds or transplanted into large black pots in greenhouse 3 to get seed. The F_2 leaf tissues were collected for the DNA extraction. Nine flats (70 F_2 seeds plus 3 controls per flat) were screened over many months and 95 susceptible and 95 resistant seedlings plus the 2 parents were chosen for data analysis.

Development and use of DNA markers

In this study, all the markers applied in the QTL mapping analysis were microsatellite markers. The information of SSR primers was obtained from the published resource (Brown et al., 1996; Taramino et al., 1997; Dean et al., 1999; Bhattramakki et al., 2000; Kong et al., 2000; Cordeiro et al., 2001; Schloss et al., 2002). Additional SSR makers were developed to have a better coverage of the genome as well as to narrow-down the marker interval carrying the potential SCA resistance loci. New SSR markers were identified either *in silico* (gene-based SSR) using the sequence database from the Phytozome website (http://www.phytozome.net/ sorghum) or using collective sequence information from various EST databases (EST SSR). SSRIT, aSSR identification tool (http://www.gramene.org/db/markers/ssrtool) (Temnykh et al., 2001) was used to search the presence of microsatellites among the DNA sequences in the preliminary QTL regions. These sequences containing the SSR were then further utilized in designing primers to amplify the repeat motifs with flanking sequences. Primer 3.0 software v 0.4.0 was used with default parameters to obtain both forward and reverse flanking primer sequences (Rozen and Skaletsky, 2000). The expected PCR product size was set to 100-300 bp, 40-60% GC content with optimum of 50%, and an annealing temperature of 54◦C to 58◦C. The forward primer was extended with the M13 primer 24 sequence (5'-CACGACGTTGTAAAACGACG-3') before the 5' end of the sequence. The standard naming system for the newly developed marker was followed as proposed by De Vicente et al. (2004). Newly discovered genetic markers were designated according to the nomenclature mentioned by Punnuri et al. (2012).

All DNAs isolated from leaf samples were collected from two parental and individual $F_{2:3}$ surviveing plants for the genomic DNA isolation. Harvested tissue was placed in the oven

to dry out at 50°C for 2days. Each tissue sample was then loaded individually into a screw-cap micro-centrifuge tube with 3 chrome steel beads and ground by a BioSpec BeadBeater machine at 1200 shakes per minute for 2 min. The genomic DNA was extracted with the modified CTAB method (Murray Thompson 1980), their concentration was quantified using the Nanodrop LITE spectrophotometer (Thermo Scientific) and normalized to 10ng/ul as a working stock of PCR DNA template.

The PCR reaction volume and amplification protocol were followed as defined by Wu and Huang (2008). PCR reactions for genotyping all 190 $F_{2:3}$ individuals and two parental lines were conducted in a PTC-220 Dyad Thermal Cycler (MJ Research Inc, MA, USA) and 2720 thermal cycler (Applied Biosystems, CA, USA). The PCR reactions were performed with an initial denaturation step of 94◦C for 5 min proceeded by 13 cycles of denaturation at 94◦C for 20 sec, primer annealing at 58◦C for 1 min, primer extension at 72◦C for 30 sec, followed by 27 cycles of denaturation at 94◦C for 20 sec, annealing at 55◦C for 1 min, extension at 72◦C for 30 sec, and a final primer extension at 72◦C for 10 min. The PCR products were separated in 6% polyacrylamide gels mounted using a LICOR 4300 DNA Analyzer (LI-COR Biosciences, Lincoln, NE, USA). The electrophoretic conditions were as follows: 1500 V, 40 mA, 40 W, and 45◦C for 1.5 hr.

The two parents were added as controls either at the beginning of the 700-dye gel or at the end wells in the 800-dye gel. The DNA banding pattern in the gel was scored manually according to the lab instruction. For every single marker, we recorded whether the individual was homozygous for the BTx623 or Tx2782 allele, heterozygous (both BTx623 and Tx2783 allele present), or missing the marker amplification.

Phenotypic data collection

Phenotyping data for sugarcane aphid resistance was conducted with the $F_{2:3}$ (F_2 derived F_3 families). Within the block, each of the 190 $F_{2:3}$ was represented with two seedlings along with two parental lines. Three $F_{2:3}$ seeds from a single genetic family were planted together in a cell present in a growing tray of 12″ X 20″ X 1.75″ size filled with autoclaved Redi-Earth soil.

The sugarcane aphid colony was reared on Tx700 sorghum seedlings in isolated insect cage in the greenhouse 2 of USDA-ARS, Stillwater until they were ready for infestation. The greenhouse condition was maintained at a temperature between 23 and 28 °C with 14 h daily light and humidity 60-80% relative humidity (Wu et al. 2007; Wu and Huang 2008). For infestation, sorghum seedlings co-cultivated with sufficient sugarcane aphid were cut and placed directly between the rows of 14-day-old sorghum seedlings, correspondingly and effectively infesting all the sorghum seedlings. Foliage damage of the $F_{2:3}$ seedlings were phenotyped for sugarcane aphid feeding response at 3 dpi (day post infestation), 7 dpi, 10 dpi, and 17 dpi sugarcane aphid infestation. Evaluation of the response of sorghum seedlings to sugarcane aphid feeding was conducted using a visual estimation to indicate damage to seedling foliage with a discrete scale that 1 being alive and 5 being dead are designed for this experiment.

Linkage analysis and QTL mapping

In this study, the MAPMAKER/Exp 3.0 (Lincoln et al., 1992) was used for linkage map construction, and the logarithm of odds ratio (LOD) of 3.0 and maximum linkage threshold of 40 cM was set up for analysis (Punnuri, et al. 2012). The recombination frequency between linked loci was transformed into the genetic distance (centimorgan,

cM) using Kosambi's function (Kosambi 1944). The output files were imported into QTL Cartographer 2.5 (Wang et al. 2005) for QTL analysis. Three QTL analysis techniques, Single Marker Analysis (SMA), Simple Interval Mapping (SIM) and Composite Interval Mapping (CIM), were performed prior to Multiple Interval Mapping (MIM) analysis to have a glimpse of the significantly associated markers and the variation explained by these markers (Ravi et al., 2003; Venkateswarlu et al., 2006). We selected a new model in the MIM module and used the Regression forward selection method to search QTLs with parameter: 1.0 cM of step and 3.0 of LOD threshold. These existing QTLs were tested, and the optimum positions of significant QTLs were refined in an iterative manner. The P-value selected for declaring a significant QTL was 0.05 with LOD value obtained from 1000 permutation tests.

Candidate genes prediction

After the major QTL was localized between SSR markers, the nucleotide sequence of two markers flanking this region was used in BLAST search against Sorghum bicolor genome sequence version 3.1 (www.phytozome.net) to define the start and end position of the physical interval. All genes within this interval were subject to functional annotation and protein homology detection. The gene was a lock as a real QTL candidate only if there was experimental evidence that it is involved with sugarcane aphid resistance.

Results

Resistance source and phenotypic data

The resultant 190 lines of the C23 were available for both phenotyping and genotyping, including 95 susceptible lines and 95 resistant lines. Leaf tissues were collected from $F_{2:3}$

individuals grown in OSU Field (2014). All DNAs were extracted from survived plants using CTAB method, their concentration was quantified using the Nanodrop LITE spectrophotometer (Thermo Scientific) and normalized to 10 ng/ul.

Genotypic data was collected by scoring bands as A (parent BTx623), B (parent Tx2783), H (Heterozygous), or – (No data or unreadable) in an Excel spreadsheet. The data of SSR-Marker Xtxp204 was scored for linkage map analysis in MAPMAKER.

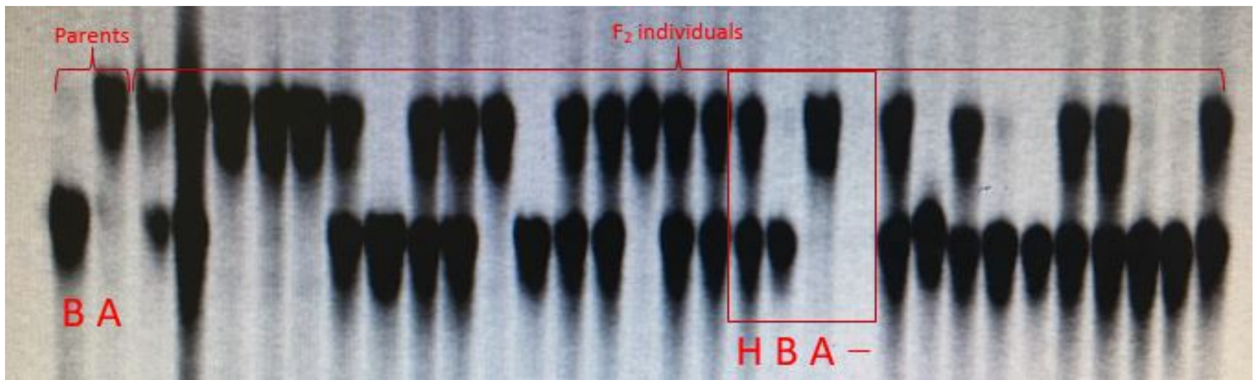


Fig 2-1. Agarose gel showing the partial genotypic data of SSR-Marker Xtxp204 screening with C23. The susceptible lines marked as “A” are at the upper band, whereas resistant lines “B” are at the lower band. “H” stands for heterozygous, “-” is missing data.

```
*sb06jh7AABAAAAAAAABBAAAAAAAAAAAAABABAHAABAAAAAAAAAAAAAABAAAAAAAAAAABAAAAAAAHAAA
AAAAAAABAAAAAAAABAAAAAAAAABAAAHAABBAAAAABBHABAAAABAABAHHBBABBAHHHAABAAAAAAA
ABAAAAAAAAAHBAAAAAAAAAAABBBAAAAAAAAAAAAAABAAAABAAHHHA

*sb06jh11BHAHBBBHHHAAHBHBHHBHBBBBABABAHBABBH-HBHBHBBHHABHHHHB-
HBHHHHBHBBHBBBHBBBHBHAHBBBBBHBBBBHBBBHHBBBH-
HAAHHHHHBHHHAHHAAAHBHHHHAABAABHHH--A-BHH-H-HA-HHHH-H-HAHHBHH-HH--AAHBHB-HH-B-
H---AHH-HAHHHHHH

*sb06jh12AHBHAAAHHHBBHAHAHHAHAAAABABABHABAAHAHAHAHAAHHBAHHAHAAHAHBHHAHAAH
AAAHAAAHAHBHAAAAAHAAAAHAAAHBAA-HBHBBHH-HHHHHHBAHHHBBHHAHHBBABBHAHH--B-
BHHBHAHH-HHHHHHA-BBHHHHAHH-HHBHAHHAHH-AAAHHHBHHA-BHHAHHH
```

Fig 2-2. Sample of collected genotypic data of primers sb06jh7, sb06jh11, sb06jh12 for QTL mapping construction. Data format was required by software instruction.

Identification of polymorphism SSR markers and construction of a genetic map

A total of 838 markers including 351 SSR markers, 302 ESR markers, 185 GSR markers, and 14 recently designed markers were surveyed for polymorphism between two parental resistant donors, Tx2783 (resistant) and BTx623 (susceptible). The percentage of polymorphism detected between the two parents was 9.2%, or 77 polymorphic markers (50 SSRs, 21 ESRs, and 6 designed markers) and were used to genotype the entire population to develop the genetic linkage map. The newly developed linkage map included 12 groups with 46 markers, which spanned a total length of 781.4 cM on 7 chromosomes. A major QTL was identified and delimited to a 500kb region on the short arm of chromosome 6. The general marker distribution was in consensus with previously published maps (Wu et al., 2006).

A linkage map was constructed by MAPMAKER/Exp 3.0 as shown below (Fig 2-3); thus, SCA resistance QTLs were identified based on the association studies of the genotyping results with the phenotypic data collected from the $F_{2:3}$ populations.

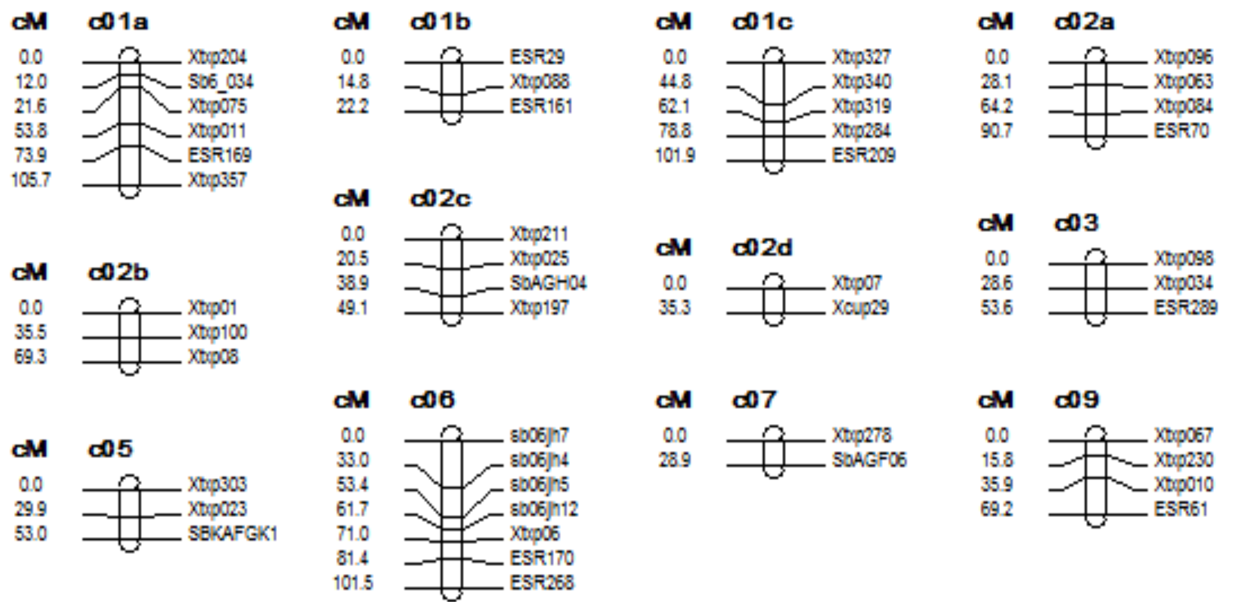
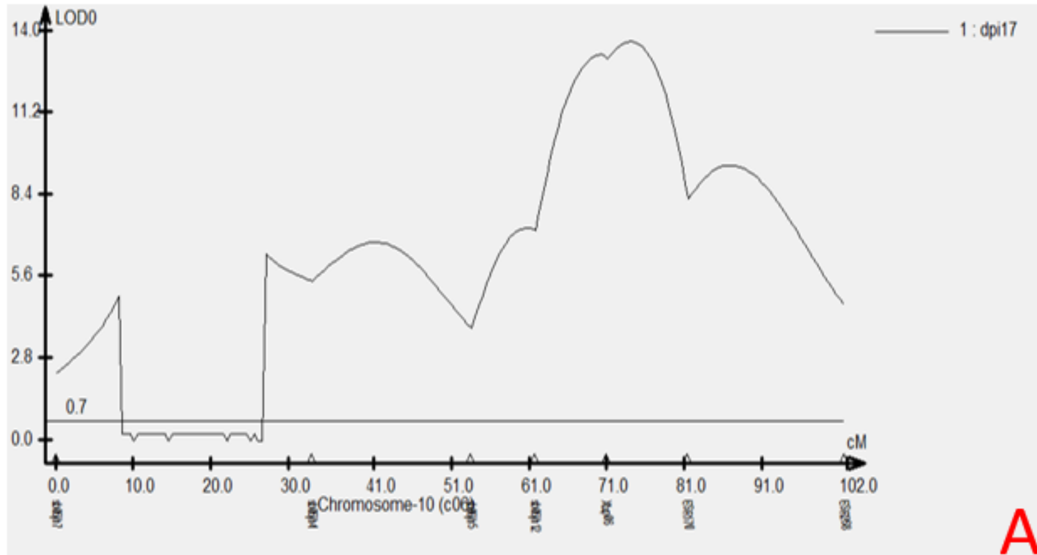


Fig 2-3. Twelve linkage groups (LGs) containing 77 polymorphic markers loci were mapped by using a set of sorghum $F_{2:3}$ lines obtained from the cross BTx623 (Susceptible line) and the resistant line Tx2783 (Resistance line).

Identification of Sugarcane aphid resistance QTL Analysis:

Based on the data information, SMA indicated that most markers significantly linked with the sugarcane aphid resistance on chromosome 6. SIM and CIM results were corroborative in the identification of the major QTL interval flanked by Xtxp06 and ESR170 on chromosome 6.



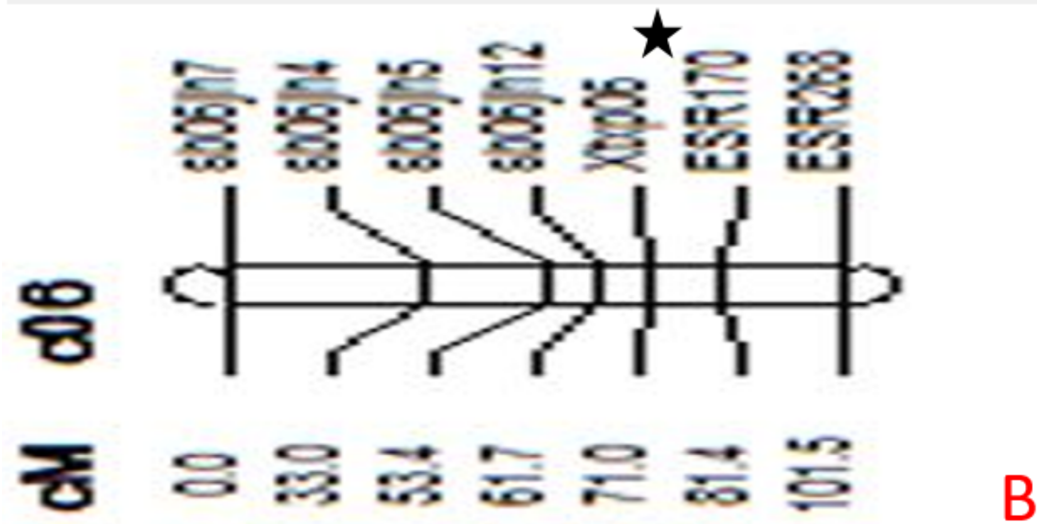


Fig 2-4. The major QTL identified in this study. (A) The LOD score peak profile using SIM analysis for sugarcane aphid resistance identified on chromosome 6. (B) Genomic location of QTL and linked markers on sorghum chromosome 6.

Candidate genes prediction

A total of 72 genes reside in this defined QTL region. Functional annotation and protein homology detection were applied to identify candidate genes. The criterion for a candidate gene was if there was experimental evidence the gene was involved in a known resistance. After all, three candidate genes were identified based on this criterion, which

may be responsible for the sugarcane aphid resistance in sorghum. A total of 7 SSR markers were mapped onto chromosome 6, the closest markers flanking the sugarcane aphid resistance locus were Xtxp06 and ESR170, which delimited a chromosome region of about 500kb containing three predicted genes, sb06g001640, sb06g001645, and sb06g001650 (Table 2-1).

Gene/marker ID	Location	Length (bp)	Gene annotation
Sb006G016350.1	2583771..2586264	2494	Unknown
Sb006G016500.1	2594601..2598442	3842	weakly similar to Arabidopsis thaliana genomic DNA, chr5, TAC clone:K9
Sb006G016500.3	2594608..2598442	3835	weakly similar to Arabidopsis thaliana genomic DNA, chr5, TAC clone:K9
Sb006G016500.5	2594612..2598428	3817	weakly similar to Arabidopsis thaliana genomic DNA, chr5, TAC clone:K9
Sb006G016500.6	2594612..2598344	3733	weakly similar to Arabidopsis thaliana genomic DNA, chr5, TAC clone:K9
Sb006G016500.4	2594612..2596454	1843	weakly similar to Arabidopsis thaliana genomic DNA, chr5, TAC clone:K9
Sb006G016600.1	2598568..2601087	2520	similar to OSJNBa0019G23.8 protein
Sb006G016700.1	2630055..2658854	28800	RNA-DEPENDENT RNA POLYMERASE 3-RELATED
Sb006G016700.2	2630536..2658873	28338	RNA-DEPENDENT RNA POLYMERASE
Sb006G016800.1	2659594..2661095	1502	PROTEASE FAMILY S26 MITOCHONDRIAL INNER MEMBRANE PROTEASE-R
Sb006G016900.1	2661605..2666186	4582	similar to Cysteine synthase
Sb006G016950.1	2663711..2664449	739	Unknown
Sb006G017000.1	2673410..2681019	7610	similar to Os03g0339100 protein
Sb006G017000.2	2673410..2681019	7610	similar to Os03g0339100 protein
Sb006G017100.1	2681149..2684265	3117	similar to Putative uncharacterized protein
Sb006G017100.2	2681149..2684265	3117	similar to Putative uncharacterized protein
<u>Sb006G017200.1</u>	2685411..2688955	3545	weakly similar to Os04g0166000 protein
Sb006G017332.1	2709784..2720442	10659	Unknown
Sb006G017266.1	2714620..2715516	897	Unknown
<u>Sb006G017400.1</u>	2724122..2727716	3595	weakly similar to Os04g0166000 protein
<u>Sb006G017500.1</u>	2748316..2751917	3602	weakly similar to Os04g0166000 protein
Sb006G017600.2	2910347..2915087	4741	DNA-DIRECTED RNA POLYMERASE
Sb006G017600.3	2910347..2915087	4741	DNA-DIRECTED RNA POLYMERASE
Sb006G017700.1	2921525..2922490	966	Unknown
Sb006G017800.1	2928365..2933751	5387	similar to OSIGBa0136O08-OSIGBa0153H12.2 protein
Sb006G017900.1	2998172..2999058	887	Unknown
Sb006G018000.1	3001970..3002551	582	Unknown
Sb006G018100.1	3006463..3007735	1273	similar to Germin-like protein 1
Sb006G018200.1	3042695..3044335	1641	similar to OSIGBa0136O08-OSIGBa0153H12.9 protein
Sb006G018200.2	3042695..3044397	1703	similar to OSIGBa0136O08-OSIGBa0153H12.9 protein
Sb006G018300.1	3049591..3051266	1676	similar to OSIGBa0136O08-OSIGBa0153H12.9 protein

Sb006G018400.1	3076736..3082312	5577	PROTEIN PHOSPHATASE 2C
Sb006G018400.2	3076063..3082312	6250	PROTEIN PHOSPHATASE 2C
Sb006G018500.1	3078591..3079356	766	Unknown
Sb006G018600.1	3107423..3109081	1659	Unknown
Sb006G018650.1	3143325..3145250	1926	ZINC FINGER CCHC DOMAIN CONTAINING PROTEIN
Sb006G018701.1	3168053..3170465	2413	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOM CONTAINING
Sb006G018751.1	3171979..3172293	315	Unknown
Sb006G018800.1	3178404..3183167	4764	similar to Ethylene receptor homologue
Sb006G018900.1	3197796..3201017	3222	similar to P0650D04.6 protein
Sb006G019000.2	3255983..3257770	1788	Putative gypsy type transposon (Transposase_28)
Sb006G019100.1	3259923..3264878	4956	Predicted protein
Sb006G019200.2	3307823..3308869	1047	Auxin responsive protein (Auxin_inducible)
Sb006G019300.1	3311087..3324205	13119	similar to Os07g0405100 protein
Sb006G019300.2	3312543..3324205	11663	similar to Os07g0405100 protein
Sb006G019400.1	3327259..3328842	1584	Unknown
Sb006G019600.1	3329226..3339781	10556	RIBOSOMAL PROTEIN L7AE FAMILY MEMBER
Sb006G019801.1	3342418..3344074	1657	Unknown
Sb006G019700.1	3340441..3341987	1547	similar to OSJNBa0039G19.9 protein
Sb006G019901.1	3346635..3347147	513	Unknown
Sb006G020000.2	3377287..3385838	8552	CLEAVAGE/POLYADENYLATION FACTOR IA SUBUNIT CLP1F
Sb006G020000.4	3376865..3385862	8998	CLEAVAGE/POLYADENYLATION FACTOR IA SUBUNIT CLP1F
Sb006G020000.3	3377286..3385862	8577	CLEAVAGE/POLYADENYLATION FACTOR IA SUBUNIT CLP1F
Sb006G020200.1	3409601..3411297	1697	weakly similar to B1168G10.9 protein
Sb006G020300.2	3443109..3447665	4557	weakly similar to Os04g0194400 protein
Sb006G020300.11	3443104..3447667	4564	weakly similar to Os04g0194400 protein
Sb006G020300.4	3443105..3447665	4561	weakly similar to Os04g0194400 protein
Sb006G020300.12	3443108..3447667	4560	weakly similar to Os04g0194400 protein
Sb006G020300.10	3443109..3447668	4560	weakly similar to Os04g0194400 protein
Sb006G020300.3	3443109..3447668	4560	weakly similar to Os04g0194400 protein
Sb006G020300.9	3443109..3447665	4557	weakly similar to Os04g0194400 protein
Sb006G020300.8	3443110..3447665	4556	weakly similar to Os04g0194400 protein
Sb006G020300.6	3443110..3447665	4556	weakly similar to Os04g0194400 protein
Sb006G020300.13	3443110..3447665	4556	weakly similar to Os04g0194400 protein
Sb006G020300.7	3443111..3447665	4555	weakly similar to Os04g0194400 protein
Sb006G020300.5	3443111..3447665	4555	weakly similar to Os04g0194400 protein
Sb006G020350.1	3462480..3464798	2319	Predicted protein
Sb006G020400.1	3495099..3497678	2580	weakly similar to Expressed protein
Sb006G020500.1	3524767..3526535	1769	similar to H0207B04.6 protein
Sb006G020600.1	3532500..3534184	1685	Very-long-chain 3-oxoacyl-CoA synthase/beta-ketoacyl-CoA synthase
Sb006G020700.1	3564784..3572233	7450	Predicted protein
Sb006G020700.2	3565064..3572233	7170	Predicted protein

Table 2-1. List of genes and their putative function in the 500kb region between marker Xtxp06 and ESR170. Three candidate genes were underlined.

According to my results, we have constructed a linkage map, the resistance QTL was mapped to a specific location of the chromosome, and we have found three candidate genes in this region which are involved in SCA resistance in sorghum. To identify the genes responsible for the target QTL, gene transformation strategies will be used in this study for mutation analysis (Fig 2-5).

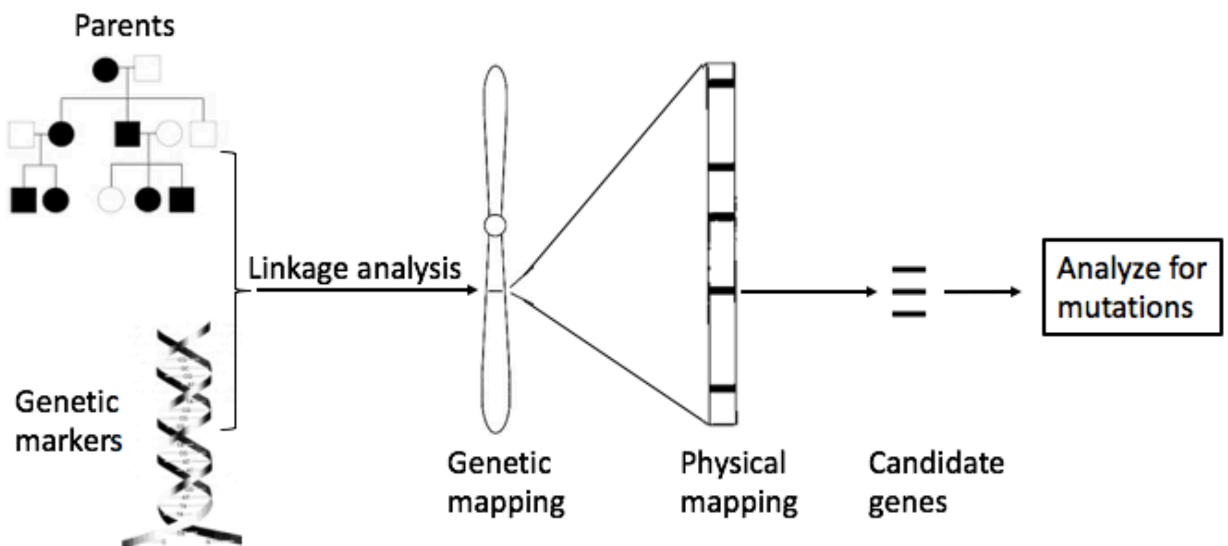


Fig 2-5. Cloning and characterization of candidate gene (Map based cloning)

Discussion

Sorghum is the fifth most important cereal crop worldwide, but its yield loss by insect pests is significant. Currently, there is a new aphid pest that has infested sorghum in the U.S named Sugarcane aphid. It was first discovered in the United States on sugarcane in Florida in 1977 and Louisiana in 1999, but did not become a pest of sorghum until 2013 (Bowling et al., 2016), after which it spread rapidly across the sorghum-producing areas.

The sugarcane aphid can feed on a wide range of host plants, which makes it one of the important economic pests worldwide, even though very limited information is known about its ecology and biology. Therefore, developing a sugarcane aphid resistance plant is an economically efficient strategy to avoid damage and yield losses by sugarcane aphid. In this study, we mapped and characterized a QTL affecting resistance to SCA in a sorghum population consisting of 190 $F_{2:3}$ from a cross of two parents, BTx623 (susceptible) and resistant donor Tx2783 (resistance).

In this research, we have identified the sugarcane aphid resistance QTLs in sorghum and detected a major QTL region located on chromosome 6. A linkage map spanning a total length of 781.4 cM across the genome was constructed using 77 SSR markers. A major QTL was identified and delimited to a 500kb region on the short arm of Chromosome 6. A total of 72 genes reside in this defined QTL region. Functional annotation and protein homology detection were applied to identify candidate genes. The criteria for a candidate gene was only if there was experimental evidence the gene was involved in SCA resistance. After all, three candidate genes were identified based on this criterion, which may be responsible for the sugarcane aphid resistance in sorghum. Three putative genes sb06g001640, sb06g001645 and sb06g001650 were found in this region and are potentially involved in sugarcane aphid resistance in Sorghum.

In conclusion, this study identified a major QTLs in the marker interval Xtxp06 and ESR170 on chromosome 6 for sugarcane aphid resistance in sorghum (Fig 2-6. B). A total of 72 genes reside in this defined QTL region, but after considering experimental evidence, three candidate genes were identified that are potentially involved in sugarcane aphid resistance in Sorghum (Fig 2-6. A, C). Sugarcane aphid feeding causes direct

damage (Fig 2-6. D) to sorghum plants and significant loss of yield when sorghum carries no resistance to the aphids. This project assisted in the development of molecular markers, and facilitated the identification of the location of those markers on the chromosome for future map-based cloning studies. The efforts to improve the sorghum breeding program for sugarcane aphid resistance management can be accelerated by using these labeled molecular markers. Mapping the QTLs for resistance to sugarcane aphid facilitates both identifications of sugarcane aphid resistance gene in sorghum and development of marker-assisted selections for breeding resistance new cultivars.

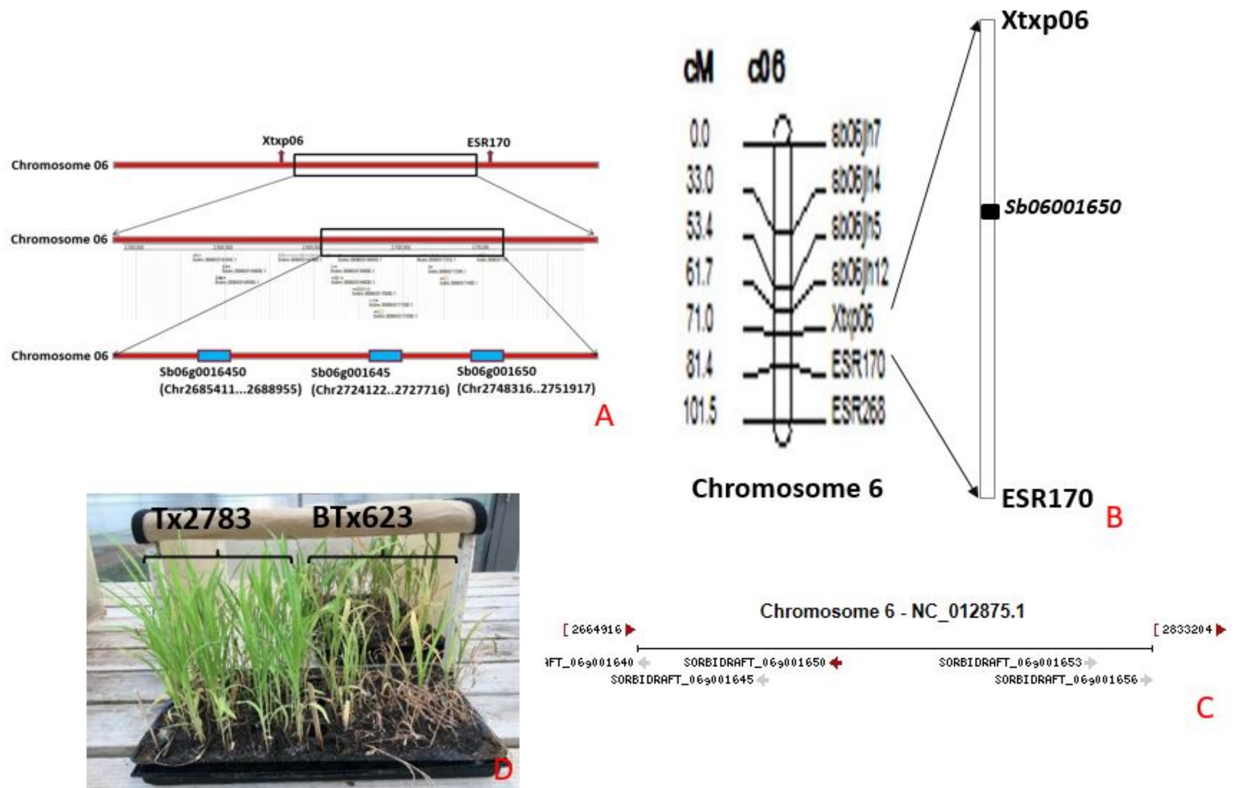


Fig 2-6. (A) QTL region on the chromosome 6. 72 genes reside in this defined QTL region. (B) Genomic location of QTL and linked molecular markers on chromosome 6; Physical location of target gene flanked by two markers with an interval. (C) Physical location of three candidate genes on chromosome 6. (D) The phenotype of BTx623 and Tx2783 with 10 dpi (day post-infestation).

References

Andrews DJ, Bramel-Cox PJ, Wilde GE. New sources of resistance to greenbug, biotype I, in sorghum. Crop Science. 1993;33(1):198-9.

Bhattramakki D, Dong J, Chhabra AK, Hart GE. An integrated SSR and RFLP linkage map of Sorghum bicolor (L.) Moench. Genome. 2000 Dec 1;43(6):988-1002.

Bowling RD, Brewer MJ, Kerns DL, Gordy J, Seiter N, Elliott NE, Buntin GD, Way MO, Royer TA, Biles S, Maxson E. Sugarcane aphid (Hemiptera: Aphididae): a new pest on sorghum in North America. Journal of Integrated Pest Management. 2016 Jan 1;7(1).

Bramel-Cox PJ, Dixon AO, Reese JC, Harvey TL. New approaches to the identification and development of sorghum germplasm resistant to the biotype-E greenbug. In Report of annual Corn and Sorghum Research Conference (USA) 1986.

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez-Candelas F, Kresovich S. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [Sorghum bicolor (L.) Moench]. Theoretical and Applied Genetics. 1996 Jul 1;93(1-2):190-8.

Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ. Microsatellite markers from sugarcane (Saccharum spp.) ESTs cross transferable to erianthus and sorghum. Plant Science. 2001 May 1;160(6):1115-23.

De Vicente C, Alercia A, Metz T. Descriptors for genetic marker technologies. IPGRI, Rome, Italy.

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE, Kresovich S. Genetic redundancy and diversity among 'Orange'accessions in the US national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Science. 1999;39(4):1215-21.

F.R. Miller. Release of A& BTx622, 623, 624. Report of Technical Committee on Seed Release and Increase. Tex. Agric. Exp. Stn., College Station, TX. Release of A& BTx622, 623, 624. Report of Technical Committee on Seed Release and Increase. Tex. Agric. Exp. Stn., College Station, TX. 1977.

Huang Y, Sharma HC, Dhillon MK. Bridging conventional and molecular genetics of sorghum insect resistance. In Genomics of the Saccharinae 2013 (pp. 367-389). Springer, New York, NY.

Kong L, Dong J, Hart GE. Characteristics, linkage-map positions, and allelic differentiation of Sorghum bicolor (L.) Moench DNA simple-sequence repeats (SSRs). Theoretical and Applied Genetics. 2000 Aug 1;101(3):438-48.

Kosambi DD. The estimation of map distance. Ann. Eugenics. 1944; 12:505-25.

Lincoln S, Daly M, Lander ES. Constructing genetic maps with MAPMAKER.

Miller FR. Release of A/BT× 623. TAES Form 97-72. Report of Technical Committee on Seed Release and Increase. Texas Agri. Expt. Stn., College Station, TX. 1976.

Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research. 1980 Oct 10;8(19):4321-6.

Padmaja PG, Aruna C. Advances in Sorghum Insect Pest Resistance. In Breeding Sorghum for Diverse End Uses 2019 Jan 1 (pp. 293-312). Woodhead Publishing.

Peterson GC, Johnson JW, Teetes GL, Rosenow DT. Registration of Tx2783 greenbug resistant sorghum germplasm line. Crop Science. 1984;24(2):390.

Punnuri S, Huang Y, Steets J, Wu Y. Developing new markers and QTL mapping for greenbug resistance in sorghum [Sorghum bicolor (L.) Moench]. Euphytica. 2013 May 1;191(2):191-203.

Ravi M, Geethanjali S, Sameeyafarheen F, Maheswaran M. Molecular marker based genetic diversity analysis in rice (Oryza sativa L.) using RAPD and SSR markers. Euphytica. 2003 Aug 1;133(2):243-52.

Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. In Bioinformatics methods and protocols 2000 (pp. 365-386). Humana Press, Totowa, NJ.

Schloss S, Mitchell S, White G, Kukatla R, Bowers J, Paterson A, Kresovich S. Characterization of RFLP probe sequences for gene discovery and SSR development in Sorghum bicolor (L.) Moench. Theoretical and Applied Genetics. 2002 Nov 1;105(6-7):912-20.

Scott Armstrong J, Rooney WL, Peterson GC, Villenueva RT, Brewer MJ, Sekula-Ortiz D. Sugarcane aphid (Hemiptera: Aphididae): host range and sorghum resistance including cross-resistance from greenbug sources. Journal of Economic Entomology. 2015 Jan 30;108(2):576-82.

Sharma HC, Ortiz R. Host plant resistance to insects: an eco-friendly approach for pest management and environment conservation. Journal of Environmental Biology. 2002 Apr 1;23(2):111-36.

Sharma HC. Host-plant resistance to insects in sorghum and its role in integrated pest management. Crop Protection. 1993 Feb 1;12(1):11-34.

Smith CM. Insect/host plant resistance in crops. In Encyclopedia of Plant and Crop Science (Print) 2019 Jul 17 (pp. 605-608). Routledge.

Taramino G, Tarchini R, Ferrario S, Lee M. Characterization and mapping of simple sequence repeats (SSRs) in Sorghum bicolor. Theoretical and Applied Genetics. 1997 Jul 1;95(1-2):66-72.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S. Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Research. 2001 Aug 1;11(8):1441-52.

Venkateswarlu M, Urs SR, Nath BS, Shashidhar HE, Maheswaran M, Veeraiah TM, Sabitha MG. A first genetic linkage map of mulberry (Morus spp.) using RAPD, ISSR, and SSR markers and pseudotestcross mapping strategy. Tree Genetics & Genomes. 2006 Dec 1;3(1):15-24.

Wang S, Basten CJ, Zeng ZB. Windows QTL Cartographer version 2.5. Statistical genetics. North Carolina State University, Raleigh. 2005.

Wu Y, Huang Y. Molecular mapping of QTLs for resistance to the greenbug Schizaphis graminum (Rondani) in Sorghum bicolor (Moench). Theoretical and Applied Genetics. 2008 Jun 1;117(1):117.

Wu YQ, Huang Y, Porter DR, Tauer CG, Hollaway L. Identification of a major quantitative trait locus conditioning resistance to greenbug biotype E in sorghum PI 550610 using simple sequence repeat markers. Journal of Economic Entomology. 2014 Sep 26;100(5):1672-8.

Wu YQ, Huang Y. An SSR genetic map of Sorghum bicolor (L.) Moench and its comparison to a published genetic map. Genome. 2006 Dec;50(1):84-9.

CHAPTER III

Molecular characterization of the sugarcane aphid [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] resistance genes in sorghum [*Sorghum bicolor* (L.) Moench]

Abstract

Sugarcane aphid is a new invasive pest of sorghum that can cause severe damage to the crop, and the economic impact of sugarcane aphids can be quite serious. This project was conducted to identify potential candidate genes within the QTL region responsible for sugarcane aphid resistance in an intercross population. In the early stage of this study, we mapped the QTLs linked to sugarcane aphid resistance in sorghum using the mapping population derived from two parents BTx623 (Sugarcane aphid susceptible line) and Tx2783 (Sugarcane aphid resistance line). Totally, 68 genes reside in this region of chromosome 6, and three genes, Sb1640, Sb1645, and Sb1650 were identified as the candidate genes for sugarcane aphid resistance. Based on analysis of DNA sequences, a 9bp insertion was discovered in the Sb1650 of the resistant line when compared with the susceptible line. Gene-specific primers were designed based on the DNA difference between BTx623 and Tx2783, which can differentiate sorghum lines with or without SCA resistance. Furthermore, gene expression assays were performed to examine Sb1650 expression patterns of the sugarcane aphid resistance gene between resistant and

susceptible lines using the real-time PCR method, which showed that expression of Sb1650 was significantly higher in resistant plants and its expression was highly induced by sugarcane aphid. This evidence suggests that the Sb1650 is responsible for sugarcane aphid resistance in sorghum. Together these molecular results from DNA analysis and gene expression patterns proved the genetic contribution of sugarcane aphid resistance in sorghum. Finally, identification of the first sugarcane resistance gene and developed DNA markers for this resistance gene offer both the genetic information and an efficient tool for genomics-assisted breeding of new hybrids or cultivars with genetic resistance to sugarcane aphids in sorghum.

Introduction

Sorghum [Sorghum bicolor (L.) Moench], is the fifth most important grain crop grown based on tonnage, after maize, wheat, rice, and barley (www.fao.org). Generally, it produces around 70 million tons of grain from about 50 million ha of land, and it is the dietary staple food for more than 750 million people in more than 30 countries. Also, it's a well-known drought and an economically important C4 grass grown for food, feed, sugar, and lignocellulosic biomass production for biofuel (Huang 2012). Sorghum has a wide range of usages, which include human food, animal feed, and the production of alcoholic beverages, and biofuels. Nutritionists also label sorghum as very healthy as it is rich with essential nutrients that are important in the body.

Sorghum is cultivated for grain and is a major food crop in much of South Asia, Africa, and Central America. The United States is the world's leading producer of sorghum with a production total of 11.5 million metric tons (Food and Agriculture Organization of the United Nations (FAOSTAT)). As a cereal, sorghum is the 3rd most produced grain in the country. The country has several regions that yield large sorghum harvests, including Colorado, South Dakota, Oklahoma, Texas and Kansas. Farmers mainly use the crop as livestock food and in the production of ethanol. India followed in ranking with an average of 7.5 million metric tons of sorghum; Nigeria ranked as the third largest sorghum producer in the world, with FAOSTAT data indicating that it produces an average 7.4 million metric tons. Forth to tenth sorghum producing countries are Mexico (6.1 million metric tons), Sudan (4.4 million metric tons), Sudan (former) (3.7 million metric tons), China (3.3 million metric tons), Argentine (2.9 million metric tons),

Ethiopia (2.2 million metric tons) and Australia (1.9 million metric tons) (http://www.fao.org/home/en).

As a food and feed crop, sorghum is an important "guaranteed" in the global agroecosystem. Worldwide, sorghum is remarkably tolerant to low input levels, an essential trait for areas such as Northeast Africa and the US Southern Plains that receive too little rainwater for most other grains. Increased demand for limited fresh water supplies, increasing use of marginal farmland, and global climatic trends, all suggest that dryland crops such as sorghum will be of growing importance to feed the world's expanding populations. Recently it has been developing as a premier bioenergy crop in the U.S. because of the high sugar level, high biomass yield, lower water demands, and market price. Sorghum plays an important role in the development and evolution of dedicated energy crops.

Meeting the food and fuel production challenges of the coming century will require production gains from traditional crop breeding, genomic selection, genome editing, and biotechnology approaches that develop plants with increased productivity and traits such as drought, pest and disease resistance, and canopies that have high photosynthetic efficiencies (Technow et al., 2015; Kromdijk et al., 2016; Mondal et al., 2016). Voytas (2013) pointed out that a new genome engineering technique can alter DNA in living plant cells which is introducing specific nucleotide substitutions in a gene that change a protein's amino acid sequence, deleting genes or chromosomal segments, and inserting foreign DNA at precise genomic locations. Mullet et al. (2014) predicted that sorghum is emerging as an excellent genetic model for the design of C4 grass bioenergy crops, and with the current trajectory of energy, sorghum genetic improvement indicates that it will

be possible to sustainably produce biofuels from C4 grass bioenergy crops that are cost competitive with petroleum-based transportation fuels. Mickelbart et al. (2015) thought crop yield reduction caused by severe climatic change could be a problem for global food security. An effective new paradigm is the targeted identification of specific genetic determinants of stress adaptation that have evolved in nature and their precise introgression into elite varieties. Ort et al. (2015) proposed an array of redesigns to improve photosynthetic efficiency, which can aid the productivity of food and bioenergy in plants.

Sugarcane aphid (*Melanaphis Sacchari*) (SCA) is an invasive aphid pest, and since 2013, it has become the most important insect pest of sorghum in the Southern U.S. and Mexico (Bowling, 2016). The sugarcane aphid was first discovered in Texas in 2013, and has rapidly moved north and west to Oklahoma, Kansas, and California. In 2013, there were only 38 counties across 4 states that had been infested with SCA, but by 2015, populations migrated to cover 417 counties across 17 states. Sugarcane aphid (SCA) is a new pest for sorghum growers, but its presence has made an impact. Aphid infestations can cause up to 100 percent crop loss if left unprotect according to a Mississippi State University study (Bowling et al., 2016).

The sugarcane aphid feeds on the sorghum only in the spring and summer, therefore, its termed as an anholocyclic, parthenogenic, viviparous species (Bowling et al., 2016). Most SCA are female, although the sexual forms have been reported separately in China (Wang, 1961); India (David and Sandhu, 1976) and Japan (Setokuchi, 1975). SCA can reproduce without mating, giving birth daily to 1-3 pregnant offspring. SCA nymphs take about 4-12 days to reach reproductive adult stage (Chang et al., 1982), with their life

length is around 28 days (Singh et al., 2004). There are two types of SCA including apterous and alate, with winged adults developing under stress conditions and environmental impacts such as high population density and declined food quality. This enables them to cross long distances (Chang et al., 1982; Singh et al., 2004).

Damage caused by SCA on sorghum produced for grain and forage happens in several ways. Generally, SCA feed on the stem and underside of leaves causing the plant to change color from green to yellow or red, to brown as it dies with an increased SCA population. SCA causes damage to sorghum by sucking the sap from plant and releasing a sticky liquid excrement called honeydew. This honeydew serves as a food source for the fungus sooty mold, which turn the plant leaves black in color, thus reducing the plant's photo synthetic capability, ultimately impacting plant growth, and causing yield and economic loss (Evers, 2018; Armstrong at el., 2015; Bowling et al., 2016).

Plant breeding has played a key role in the coordinated effort for increased crop production with molecular markers improving the efficiency and precision of conventional plant breeding via marker-assisted selection (MAS) (Collard and Mackill, 2008). Normally, the process of developing new crop varieties can take almost 25 years. Now, biotechnology has significantly shortened the period to 10 years for new crop varieties to be released to the market. One of the tools which can make it easier and faster for scientists to select plant traits is marker-assisted selection (MAS). Marker assisted selection (MAS) refers to the use of DNA markers that are tightly-linked to target loci as a substitute for or to assist phenotypic screening (Ben-Ari and Lavi, 2012).

MAS was discovered 32 years ago by C. Smith and P. Simpson (1986) and by Soller and Beckmann (1983), however, Neimann-Sorensen and Robertson (1961), Soller (1978), and Stuber et al. (1982) suggested the concept of using DNA markers for the purpose of selection in breeding programs. Moreover, Tanksley et al. (1981) published the first MAS experimental study.

The advantages of MAS include three main reason. First of all, it is a simpler method compared to phenotypic screening, especially for traits with laborious screening, which also saves time and reduces cost. MAS increases reliability with no environmental effects, it can discriminate between homozygotes and heterozygotes, and is more accurate and efficient with selection of specific genotypes. Thus, leading to an accelerated variety development (Collard and Mackill, 2007; Collard et al., 2005; Hash et al., 2003; Hasan et al., 2015; Kelly and Miklas, 1999; Ribaut and Hoisington, 1998; Ejeta and Knoll, 2007; Knoll and Ejeta, 2008).

Materials and Methods

Plant material and growth condition

Two parents BTx623 (susceptible to sugarcane aphid) and Tx2783 (resistant to sugarcane aphid) and two progenies number 193 (Resistance progeny) and 194 (Susceptible progeny) were grown in pots in the greenhouse with a separated growing tray (12″ X 20″ X 1.75″) filled with autoclaved Redi-Earth soil. One tray was for sugarcane infestation, and another one served as control with non-infestation. Fifty fungicide-treated (VPG Hi-Yield, Captan 50W Fungicide) seeds were planted in 10 holes for each line. The sorghum seedlings were infested with sugarcane aphids amplified on susceptible sorghum

seedlings when they were three weeks old. The greenhouse was maintained at a temperature between 23 and 28 °C with 14 h light and 60-80% relative humidity.

The plant samples from the whole plant including leaf tissue and stem were collected at 0.5 days, 1 day, 2 days, 3 days, 6 days, 9 days and 12 days after infestation. Samples were also collected at 0 day before infesting sugarcane aphid, which served as controls (calibrator sample) for the real-time PCR experiments. More than three seedlings for every line from each replication were combined and weighed to make more than 2.0 g of tissue for subsequent RNA extractions. Eight samples per replication were collected, frozen in liquid nitrogen, and stored at -80◦C for RNA extraction. All samples were collected twice from different pots, which represented two technical replications.

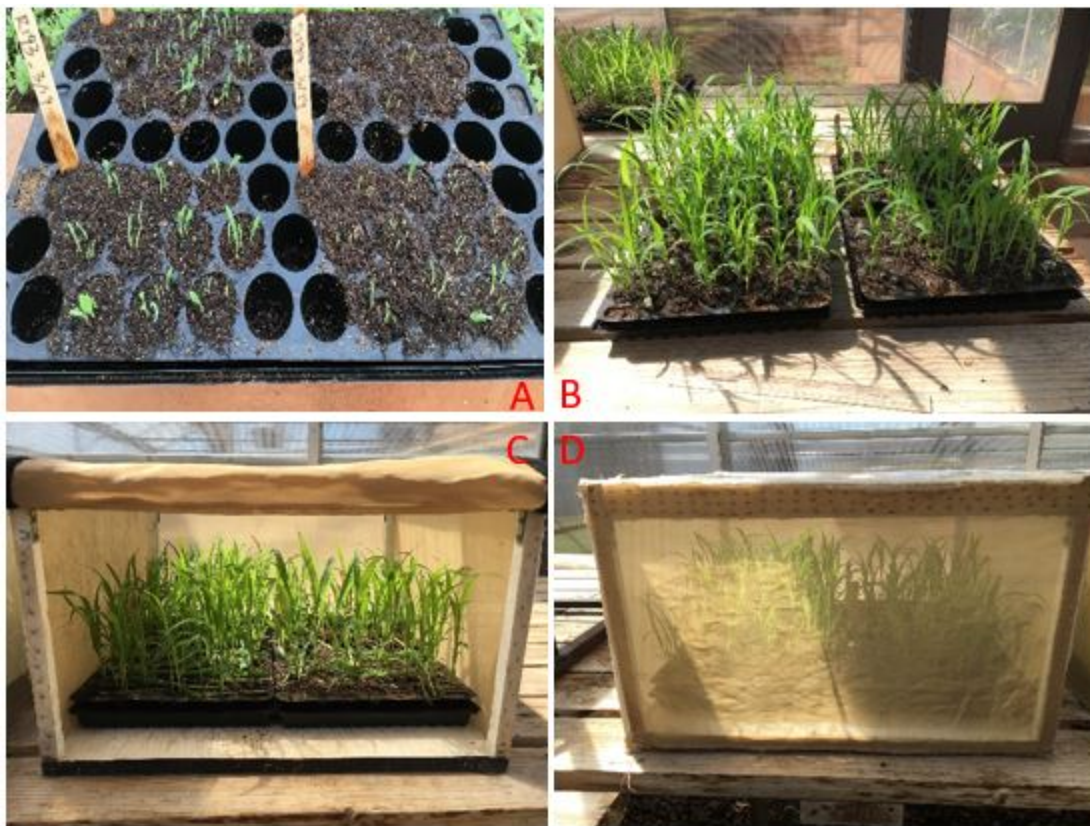


Fig 3-1. Plant material preparation and infestation with sugarcane aphids: (A) seed germinating, (B) young seedlings grown in greenhouse, (C) Infestation with freshly prepared aphids, and (D) co-cultivation of host plants with aphids.

RNA extraction and cDNA synthesis

Total RNA was extracted from the eight different samples using TRIzol reagent (Invitrogen, CA). RNA samples were quantified using a NanoDrop spectrophotometer and diluted to an equal concentration of 1 μg/ul of RNA for all eight samples. The cDNAs were synthesized from mRNAs with reverse transcriptase using the QuantiTect Reverse Transcription Qiagen Kit according to the manufacturer's instruction (Qiagen, CA). The kit also contains a prior step of removing genomic DNA contamination by treating with DNAase, therefore the cDNA synthesized was free of genomic DNA. All cDNAs and primers were checked using regular PCR.

Primer design as gene-specific markers

Identification of the candidate SCA resistance gene in Tx2783

Based on the mapping result, there are three hypothetical proteins in the SCA resistant QTL region on chromosome 6, including Sb06g001640, Sb06g001645, and Sb06g001650. Thus, using real-time PCR method, expression of the three potential candidate genes was analyzed in the four samples including two parents and two progenies (R193, and S194) in order to evaluate whose expression was related to host plant resistance to infestation from the sugarcane aphids. Primers specific for three candidate genes in the QTL region were designed using the software Primer 3.0 (Rozen and Skaletsky 2000). The cDNA sequences for the three genes were retrieved from phytozome (a database for sorghum (http://www.phytozome.net/sorghum)). The forward and reverse sequences of these genes were selected using Primer 3.0 (http://www-genome.wi.mit.edu/genome_software/other/primer3.html) with the following criteria: length of the primer is 18-23 nt, optimum = 20 nt; expected product size = 150-300 bp;

annealing temp = 58-60◦C and GC = 45-60 % with an optimum of 50 %. These primers were checked at the sequence level to avoid primer dimers.

After RT-PCR, the candidate gene was determined to be Sb06g001650. A 3.1 kb DNA fragment (Sb1650) was cloned from Tx2783 and sequenced using the service provided by the core facility of Oklahoma State University. The corresponding DNA sequences from BTx623 were retrieved from the sorghum genome database (www.phytozome.net/sorghum) as the reference sequence for sequence comparison. Also, the candidate genes in the two progenies R193 (resistant progeny) and S194 (susceptible progeny) were sequenced for this sequence analysis.

Primer name	Sequence	GC content	Tm value
Sb1640qPCR_F	GGCAGAACACTCACAACTTTA	43%	62
Sb1640qPCR_R	CCCAATCCGTCTGGTTCATATCC	52%	67
Sb1645qPCR_F	CCTATGGGTGCTAGATATGAGCC	52%	67
Sb1645qPCR_R	TATTACCAGATAGATCGAGGAGC	43%	62
Sb1650qPCR_F	GTGTACTCAAGTTATGCGGTTGC	48%	66
Sb1650qPCR_R	CCTTCCCTTCCTTACATGAACC	50%	65

Table 3-1. List of sequences of the gene specific primers used in real-time PCR reaction. The primers sequences were designed based on the sequence of the cloned gene (Sb1650).

Expression of the candidate gene

Real time PCR

Total RNA was extracted from each collected tissue of each sample using TRIzol Reagent (Invirogen) following the manufacturer's instructions, 1 ug aliquots were treated with DNase I (Promega) and then reversely transcribed into cDNA using the Reverse Transcription Kit (Promega). The transcriptional profile was analyzed by RT-PCR using the SYBR Green PCR Master Mix (Applied Biosystems) and the Biorad Real-Time PCR System. Primer specificity was verified by cloning and sequencing PCR products obtained from BTx623 cDNA. The melting curves analyses and the evaluation of primer

amplification efficiency were also performed before running the cDNA samples. Relative expression levels were calculated from the cycle threshold using 2-ΔΔCt method as described by Livak and Schmittgen (2001). A sorghum housekeeping gene TUBLIN was used as an internal control. The data were an average of the three biologic replicates.

The expression level of the candidate gene was analyzed by using real-time PCR experiments. The housekeeping gene (tubulin) was co-amplified along with the candidate gene for all samples, which was further used in normalization during gene expression analysis. Firstly, a standard curve was made for every primer along with the tubulin primer, for eight data points using 10-fold cDNA dilutions. The 15 µl final volume of PCR mix was made using 3µl of cDNA template dilution obtained from the standard curve, 12 µl mix of the primer of 1µM and SYBR *Premix Ex Taq TM*green I dye fromTaKaRa mix (TAKARA BIO INC, CA). The amplification protocol was composed of initial denaturation at 94◦C for 5 min, followed by (94◦C for 10 sec, 58◦C for 30 sec, 72◦C for 30 sec) for 40 cycles, 94◦C for 1 min, 55◦C for 1 min, (55◦C for 10 sec) for 81 cycles and finally a hold at 4◦C. The melt curve was started at 55◦C with increase in 0.5◦C until 94◦C. The *C*T values were obtained from the MyiQ software. The level of expression of tubulin was constant in two parent lines and two progenies at four different time points. The formula for 2-ΔΔCT method (Livak and Schmittgen 2001) for relative gene quantification was used as specified in Bio-Rad Real-Time PCR applications guide using delta delta *C*T values (Bio-RAD Laboratories, Inc. Hercules, CA). For each gene, a two-way fixed factor statistical model was used to assess the differences in expression using delta *C*T values between families (two parents), time points, and their interaction. The statistical analyses were calculated using the GLIMMIX procedure in SAS version

9.2. Least-squares means (or means) and differences of the means were calculated using the LSMEANS command in SAS and the type I error rate was controlled by Tukey's adjustment. We had two biological replications, each consisting of eight samples including controls. PCR experiments were repeated twice (two technical replications) for each biological replication. In the technical replications, each sample was assayed in triplicate and the values were averaged as one data point.

Molecular cloning and transformation

Amplification of the candidate gene by PCR

The gene Sb1650 was amplified from the cDNA of the resistant line Tx 2783 using the gene specific primers fused with restriction site tails (Forward: 5'GCGTGGCGCGCCATGCAGGAAATTTTTGTG3' Reverse: 5'GCGTCCCGGGTCATCGAAGCACTGCACC3'). After PCR was completed, 5 ul of product was run on 0.8% agarose gel to ensure that the correct size band was dominant and at least present in the sample. PCR/gel DNA cleanup kit (Qiagen) was used to purify the DNA obtained from the PCR mixture, then quantified the concentration of the resultant cDNA clone.

Amplicon digestion by a restriction enzyme

For insert preparation, 1 ug of the purified cDNA from above was digested with AscI and XmaI for 1 hour, and stopped by deactivation at 80°C for 20 minutes. After digestion was completed, the digestion product was run on a 0.8% agarose gel and checked under low-frequency UV light. Then the target DNA band on the gel (at about 3k in size) was cut out and put into a pre-weighed 1.5 ml microcentrifuge tube. The final product of the

insert was purified using the Cleanup kit and quantified after cleanup. The resultant insert [cDNA of (Sb1650)] was stored at -20°C until needed for ligation.

Plasmid digestion by a restriction enzyme

Similarly, the binary vector pFGC161 DNA was digested with AscI and XmaI. After digestion was complete, the digested plasmid was checked on a 0.8% agarose gel. The linearized plasmid was cut out from the gel (should be at 12,746 bp in size) and purified using the PCR/gel DNA cleanup kit. The pre-cut vector was stored in -20°C until needed for ligation.

Development of the gene construct

For ligation, the precut vector and insert DNA were added to the reaction buffer at a mole ratio of insert to vector 3:1. This was calculated based on the length of the plasmid DNA, which should be 12,746 bp and the length of insert DNA should be about 3,081 bp. Finally, the enzyme T4 ligase was added to the reaction tube. The tube was mixed gently by flicking and then briefly spun down. The reaction mix was incubated at room temperature for 2 hours and then deactivated at 25°C for 2 hours followed by a ramp up to 70°C for 10 minutes, then back down to 4°C.

Transformation of the gene construct into *E. coli* cell

For *E. coli* transformation, the E. coli competent cells (DH5α) were thawed on ice. Then 5 ul of the mixture from the ligation reaction (also chilled on ice) was added and mixed thoroughly by flicking. It was kept on ice for 1 hour, with occasionally mixing every 20 minutes or so. After 1 hour, the tube was transferred onto a heat block pre-set at 42°C for

30 seconds and then immediately transferred back to ice. After 5 minutes back on ice, 900 ul of room temperature liquid SOC media was added to the tube and mixed by inverting, then transferred to a shaker pre-set to 200 rpm at 37°C for 1 hour. Finally, ~100 ul was plated onto a plate containing LB media + agar that was supplemented with 50 ug/ml of kanamycin, then incubated overnight in an oven at 37°C.

Confirmation of successful sequence insertion

The next morning Colony PCR was conducted on the bacteria with the internal primers. Briefly, each colony was picked from the plate and added to a tube for PCR amplification. After the PCR amplification was completed, the amplification products were checked using gel electrophoresis in order to select bacterial colonies that contained the correct size for SCA resistance.

The selected colonies were cultured in 10 mL of liquid LB media in a culture tube or a 50mL Falcon tube containing 50ug/mL kanamycin and cultured overnight. This large volume of bacterial culture was used to satisfy the extraction of a large amount of the plasmid DNA.

Diagnostic restriction digestion

A miniprep kit was used to extract the plasmid DNA from each sample. The plasmid DNA concentration of each sample was quantified after extraction and noted. Restriction digests with AscI and XmaI were performed in a 25 uL total volume reaction at 37 ^{o}C for 1 hour followed by 80 ^{o}C for 20 minutes, DNA of the original pFGC161 plasmid (un-digested) was used as a negative control and digested it with AscI and XmaI. Each sample, including the control, was run on a 0.8% agarose gel. All samples, including the

control, should have a large band at 12,746 bp size, however, samples that were successfully inserted with the gene of interest should also have a smaller band at 3,081 bp size, while the pFGC161 negative control should have a small band at 1,184 bp in size. DNA stocks of any samples that did not show a positive pattern for the gene of interest in the plasmid backbone were discarded.

Sanger sequencing.

DNA samples with a positive restriction digest result were sent to the HBRC Sequencing Core Facility for confirmation.

Specific primers were designed that matched to the sequence of the pFGC161 backbone just upstream of where and downstream the gene of interest was inserted. These primers confirmed the correct sequence near the junctions.

Since the sequence of interest was 3 kb in length, primers were designed along the predicted sequence every 600 to 800 bp or so to get the full sequence. However, since Q5 polymerase very rarely makes copying error, we were able to sequence only the junction sites.

After sequencing confirmation, the DNA was ready to be mobilized to *Agrobacterium* for transformation.

Transformation of sorghum plant

The transformation work was done in collaboration with the Plant Biotechnology Innovation Laboratory of University of Missouri. The process started with collection of immature embryos (1.5 mm in length) from sorghum genotype P898012 harvested 10 to

14 days after pollination. *Agrobacterium tumefaciens* strain AGL1 carrying our constructed plasmid was grown out onto YEP media plates supplemented with antibiotics rifampicin and kanamycin and incubated at 28 degrees for 3 days before the transformation. On the day of transformation, bacteria were collected from the plate and re-suspended in inoculation medium at OD600= ~0.4 and then shaken at low speed for 3 to 4 hours. Immature embryo explants were collected from inoculation medium and then after washing with more inoculation medium to remove starchy endosperm, embryos were inoculated with the *Agrobacterium* suspension for 10 minutes at room temperature. Embryos were then transferred onto co-cultivation medium plates and spaced out on the surface of the medium scutellum side facing upward. These were then incubated at 25 degrees in the dark for three days until explants were transferred to resting medium plates and incubated at 28 degrees in the dark for 7 to 10 more days. After this time, explants were transferred to selection medium containing the herbicide glufosinate, which could select for tissues that were successfully transformed with the target construct. After 2 weeks on this medium, visible somatic embryos formed on the callus surface. They were transferred to shoot induction medium and incubated at room temperature on a 16/8 photoperiod. They were kept on shoot induction medium for 4 weeks (with a subculture at 2 weeks). Subsequent, healthy shoots were transferred to rooting medium and subcultured every 2 weeks until they were sufficiently robust to be transferred to small soil pots. After a hardening period, plants were matured in small pots until they were large enough to be transferred to large pots in the greenhouse. At this time, transgenic plants were evaluated by leaf painting with the Bialophos herbicide. Finally, at least 5

independent transformation events would be ready for characterization of the transgenic plants and validation of the SCA resistance gene.

Fig 3-2. The protocol for developing transgenic sorghum plants carrying the SCA resistance gene. (A) Surface sterilization of the immature caryopsis (collected around 10-14 days after pollination). (B) Further sterilization of immature caryopsis in the conical flask with 50% bleach + few drops of tween 20. (C) Isolating immature embryos the immature caryopsis. (D) Immature embryo heat treatment and Agrobacterium infection. (E) Immature embryos on co-cultivation medium. (F) Callus induction from the initial explants. (G) Shoot regeneration. (H) Rooting, transformation of regenerated shoots to rooting medium. (I) Hardening, regenerated shoots with healthy roots are transferred to growth chamber.

Results

Determination of the Sb1650 as a candidate gene for resistance to sugarcane aphid

First, quantitative RT-PCR was performed to evaluate which gene was responsible for the sorghum plant resistance to sugarcane aphid. According to the RT-PCR results (Fig. 3-3), expression of Sb1650 was highly correlated with host-plant resistance among the three potential candidate genes, Sb1640, Sb1645, and Sb1650 located in the QTL region. Expression activities of the three potential candidate genes were measured, in which the Sb1650 showed its strong reaction to aphid infestation. The Sb1650 showed significant increases in response to aphid attack and continued high expression levels during the co-cultivation. Thus, Sb1650 is involved in the sugarcane aphid resistance in sorghum.

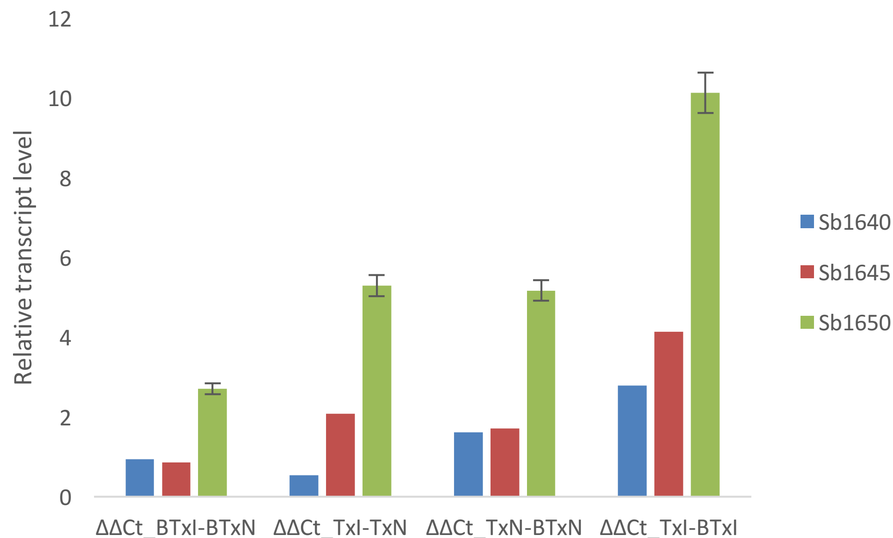


Fig 3-3. Expression of three candidate genes at 7 dpi in two parents with treatments (infestation and non-infestation). BTxI-BTxN stands for comparison of infested BTx623 and non-infested BTx623; TxI-TxN stands for comparison of infested Tx2783 and non-infested Tx2783; TxN-BTxN stands for comparison of non-infested Tx2783 and non-infested BTx623; TxI-BTxI stands for infested Tx2783 and infested BTx623.

DNA sequence analysis

The 3.1 kb DNA fragment of the Sb1650 was cloned from both the resistance donor Tx2783 and the susceptible elite line BTx623 using PCR method and sequenced, for the full open-reading frame. Sequence comparison was performed for the full open-reading frame. Interestingly, as shown in Fig.3.5, there was a nine nucleotides deletion (GTGGTTTCC) from the same gene in BTx623 in comparison to the Sb1650 in Tx2783. This deletion was also found in the susceptible progeny S194. Thus, the sequence analysis data suggests that the 9bp deletion was responsible for the sugarcane aphid resistance in sorghum.

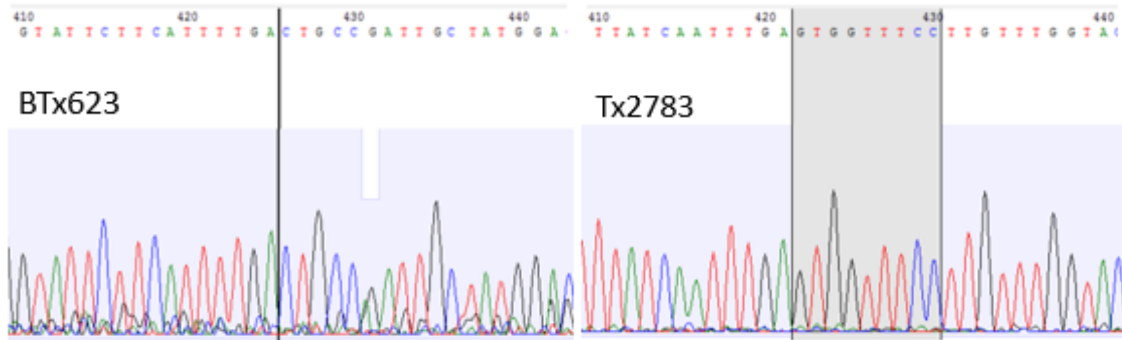


Fig 3-4. Interpretation of Sequencing Chromatograms of 9bp deletion within BTx623 compared to the DNA sequence in the resistant line Tx2783.

```
S194       ...AGATCTGGGTATTCTTCATTTTGA---------CTGCCGATTGCTATGGACCTTTAGAGG...
R193       ...AGATTTCGGCATTTATCAATTTGAGTGGTTTCCTTGTTTGGTACTATGGACCTTTAGAGG...
Tx2783     ...AGATTTCGGCATTTATCAATTTGAGTGGTTTCCTTGTTTGGTACTATGGACCTTTAGAGG...
BTx623     ...AGATCTGGGTATTCTTCATTTTGA---------CTGCCGATTGCTATGGACCTTTAGAGG...
              **** * ** ***  *** *****          **      * *****************
```

Fig 3-5. Sequence annotation was performed for gene molecular analysis. Sequence comparison of two parents BTx623 and Tx2783 and two progenies R193 and S194. Bold indicates the 9bp deletion region in susceptible lines, BTx623 and S194.

Candidate gene confirmation using specific genic marker

Based on the deletion difference of DNA sequence of the candidate gene between Tx2783 and BTx623, the gene specific primers for Sb1650 were developed, including

Forward: CGGCATTTATCAATTTGAGTGGTTTCC, and

Reverse: TCTAACGACATCAGGGTCAGG. As shown in Fig. 3-6, this pair of primers produced expected results, positive DNA amplication in resistant lines but negative in susceptible lines that can be used for plant breeding. Shown here is the agarose gel picture, two parents and 6 $F_{2:3}$ progenies including 3 susceptible lines and 3 resistant lines that were selected for marker verification. All SCA resistant lines show positive results, which indicates the marker is reasonable for marker-assisted selection.

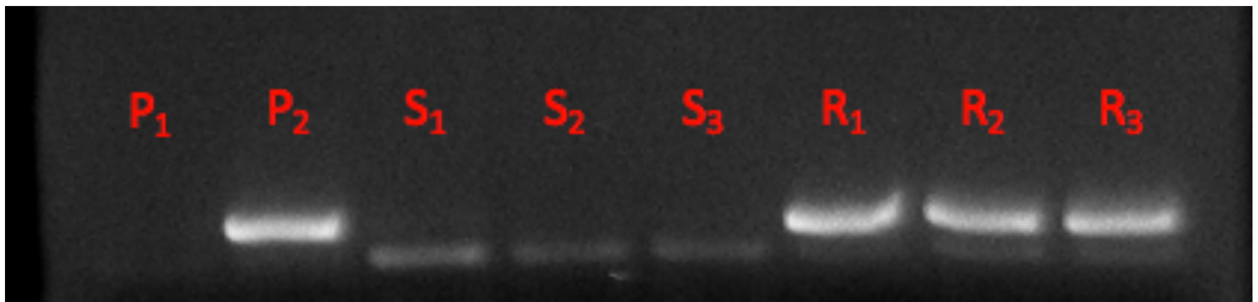


Fig 3-6. DNA amplification in SCA resistant and susceptible lines. PCR products were amplified using primers flanking the mutation region. P_1 is BTx623 SCA susceptible parent; P_2 is Tx2783 SCA resistant line; S_1, S_2, S_3 are $F_{2:3}$ susceptible progenies; R_1, R_2, R_3 are $F_{2:3}$ resistant progenies.

For further confirmation, this pair of primers was tested among the randomly selected 60 lines of this family at the F_6 generation (Table S3-1). The same type of DNA amplification was produced among these lines, with positive DNA amplification in all resistant lines tested, but no amplification in all susceptible lines (Fig. 3-7, Table S3-1). These results from the DNA amplification with the gene specific primers further confirmed that the Sb1650 is indeed the sugarcane resistance gene in sorghum. Moreover, the primers (markers) proved to be a useful tool for not only reliable identification of genetic sources with sugarcane aphid resistance, but also for marker-assisted breeding to

facilitate developing new hybrid and cultivars with resistance to sugarcane aphids.

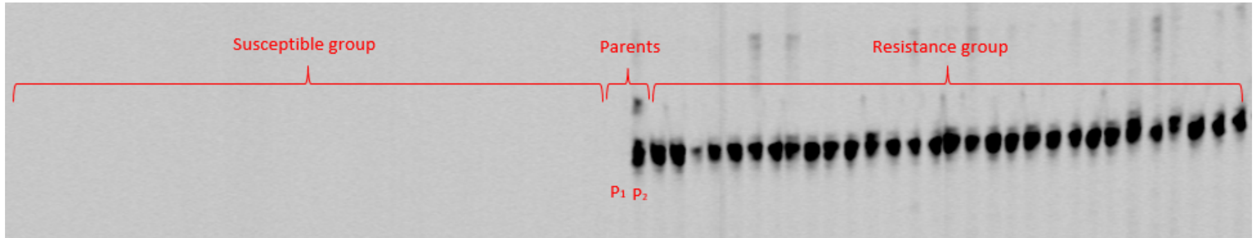


Fig 3-7. Licor gel analysis of amplification products in 60 randomly selected lines at the F_6 generation, due to the deletion in DNA sequence, in susceptible lines (P_1 is susceptible parent BTx623, P_2 is resistance parent Tx2783).

Interactions between host plant and sugarcane aphid

Phenotypic data. Sugarcane aphid causes direct damage to sorghum plants and significant loss of the yield when sorghum carries no resistance to the aphids. In the experimental trials, BTx623 and Tx2783 were growing side-by-side, then they were infested by sugarcane aphid through artificial infestation (Fig 3-8); simultaneously another group of seedlings from the two lines were growing in the same way, but without aphid infestation served as controls. Two parents, BTx623 (susceptible) and Tx2783 (resistance), and two progenies, R193 (resistance) and S194 (susceptible), were planted for interaction analysis. Three seedlings of each line were selected to measure the increase in height of the seedlings. All individual plants were evaluated at 0 dpi, 0.5 dpi, 1 dpi, 2 dpi, 3 dpi, 6 dpi, 9 dpi, and 12 dpi. Also, three representative seedlings at each stage were collected for RNA isolation.

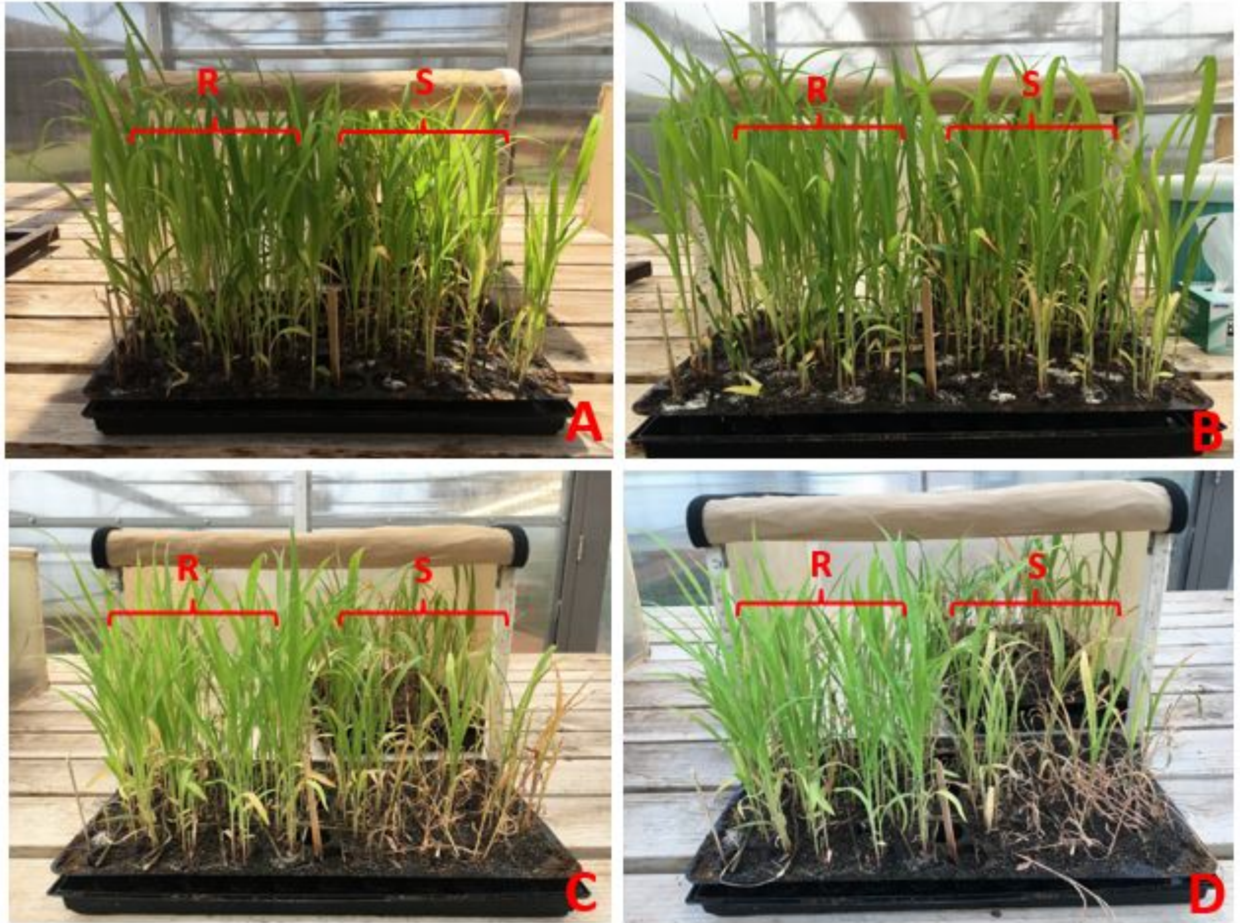


Fig 3-8. (A) Non-infested resistance and susceptible lines at 9dpi. (B) Non-infested resistance and susceptible lines without infestation at 12dpi. (C) Infested resistance and susceptible lines at 9dpi. (D) Infested resistance and susceptible lines at 12dpi. The treatments included three seedlings of each line (BTx623-noninfested, Tx2783-noninfested, R193-noninfested, S194-noninfested, BTx623-infested, Tx2783-infested, R193-infested, and S194-infested).

Following infestation, the SCA started feeding on the underside of seedling leaves and stem. Though initial colonies consisted of just a few aphids, the rapid regenerating aphids covered much of the lower parts of the seedlings on the susceptible lines, but showed small increase in aphid number on the resistant lines. Large infestations stunted the plant growth of susceptible lines. As shown in Fig 2-8, the SCA resulted in slow growth of

seedling height. Aphid feeding caused leaves/seedlings to turn yellow, and then brown as the leaf tissue died.

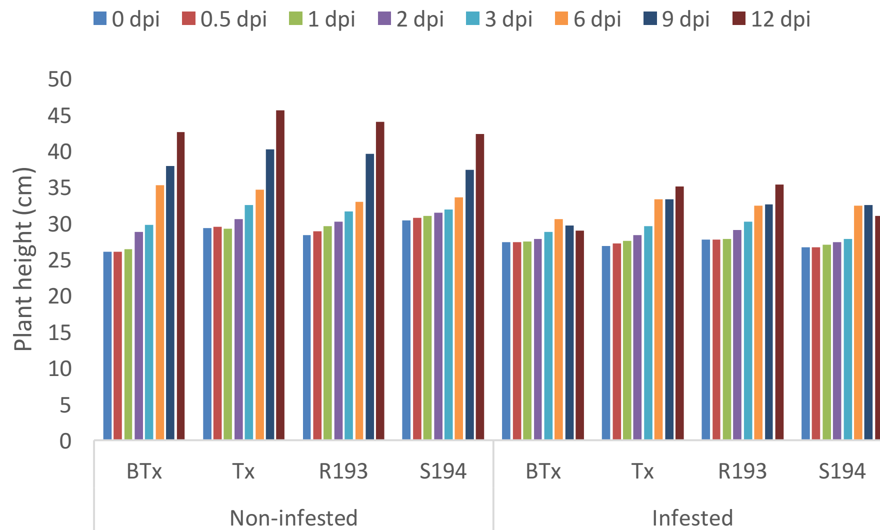


Fig 3-9. Seedling height increase of four different lines at 8 different time-points. The four lines were BTx623 (susceptible) and Tx2783 (resistance) and two progenies R193 (resistance) and S194 (susceptible).

Genotypic data. Two parents (BTx623 and Tx2783) and two progenies (R193 and S194) were arranged to check the activity of the Sb1650 in responding to attack by sugarcane aphid. In this way, the induction of Sb1650 gene by the SCA was further confirmed using quantitative real-time PCR (RT-PCR). The results of RT-PCR analyses at multiple time points indicated that the transcripts of the Sb1650 gene were SCA-specific and rapidly increased over the time course (Fig. 3-10). The level of Sb1650 expression was induced following attack by aphids and rapidly increased to a significant high level by six days post-infestation. Much stronger expression was induced by the SCA in the resistant genotype than those in susceptible plants.

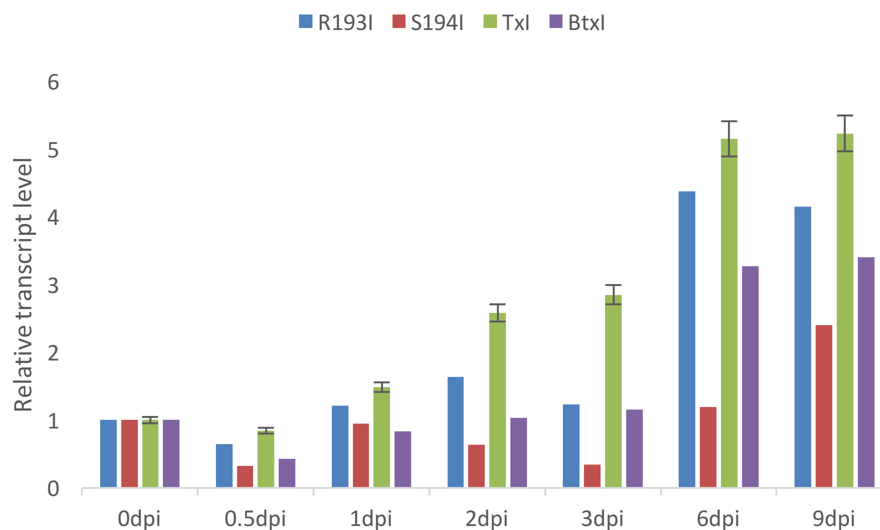


Fig 3-10. Expression of SCA resistance gene in sorghum resistance and susceptible lines at different stages (0 dpi, 0.5dpi, 1dpi, 2dpi, 3dpi, 6dpi, and 9dpi), dpi stands for days post-infestation. R193I stands for infested resistance line 193, S194I for infested susceptible line 194, TxI for infested Tx2783, BTxI for infested BTx623. Data were the mean values and the standard deviation of more than three biological repeats.

Gene function validation using transformation

Construction of a binary vector to express the SCA resistance.

The full length candidate gene has been cloned and fused to a binary vector as shown in Fig. 3-11.

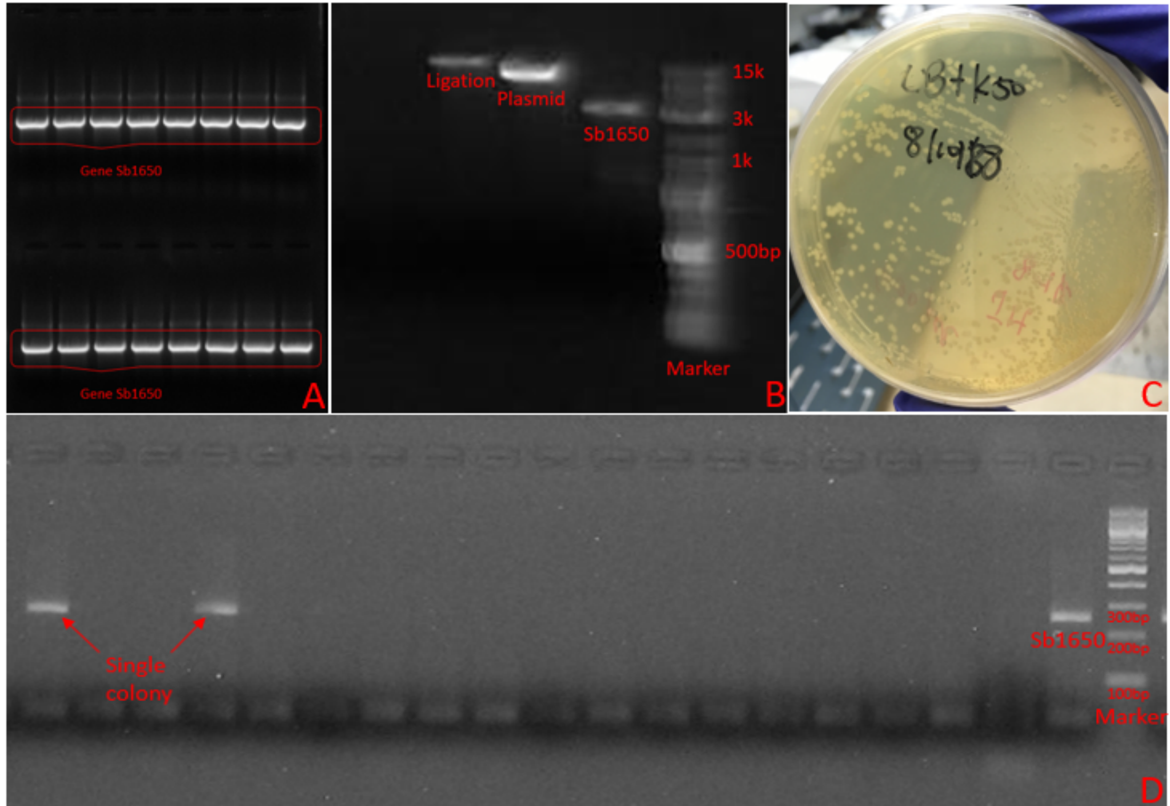


Fig 3-11. (A) Agarose gel showing the amplification of the candidate gene Sb1650 by PCR. (B) Agarose gel showing the products of digestion of plasmid pFGC161 and the Sb1650 (insert), and the ligation of both insert and vector. (C) Bacterial colonies carried the recombinant gene construct grown in the petri dish with LB medium+ Kanamycin. (D) PCR-confirmation of the selected colonies that carry the binary vector with the Sb1650.

Once the expression vector was constructed in the E. coli cells, the plasmid was isolated from E. coli, purified, and then transformed into *Agrobacterium tumefaciens*. Thus, it is ready for the next step, *Agrobacterium*-mediated transformation in order to transfer the SCA resistance gene into sorghum lines lacking such a resistance gene.

Sorghum transformation work is currently underway in cooperation with Professor Zhanyuan Zhang at the Plant Biotechnology Innovation laboratory, University of

Missouri. It is expected that transgenic sorghum plants carrying the SCA resistance gene will be regenerated and produce T_1 seeds within four to six months. With that in hand, experiments for analysis of the gene expression in transgenic plants and validation of gene function will be conducted.

Discussion:

Sorghum is affected by various biotic factors, leading to severe reduction in productivity and production in different production systems. Development of host plant resistance is the cheapest and most sustainable method for managing the pests and diseases (Huang 2011). Cereal aphids such as greenbug and sugarcane aphid are major pests of sorghum. Thus, great research efforts during the past years have led to the identification of genetic sources with natural resistance to those aphids, which promise to help manage the aphid problems (Dogimont et al. 2010). Now, sorghum growers are asking for genetically improved hybrids and cultivars that can defend against those aphid pests, which means sorghum breeders are in need of genetic information of host-plant resistance.

This dissertation research project focused on such an important topic in genetics of sorghum resistance to the new invasive sugarcane aphid. The first part of this project aimed to identify the genomic region responsible for resistance to the SCA. For this purpose, a mapping population was developed from a cross between BTx623 (susceptible to SCA) and Tx2783 (resistance to SCA). BTx623 is an elite line of grain sorghum, whose genome was sequenced for the first time in 2009 (Paterson et al. 2009); thus, it has been used as a model system to characterize and determine the genetic basis for many important traits. DNA markers (i.e. SSR) developed for sorghum in Dr. Yinghua Huang's laboratory were used for QTL mapping in the current study. Subsequently, a major QTL

was found closely linked to the genetic resistance to SCA, flanked by Xtxp06 and ESR170, on chromosome 6. Interestingly, according to the previous reports from Dr. Huang's laboratory, the major QTLs associated with greenbug resistance resided on chromosome 9 of sorghum (Punnuri et al. 2013, Wu and Huang 2008). These results indicate that the genetic basis for resistance to sugarcane aphids is different from that of greenbug resistance in sorghum. However, resistance to SCA and resistance to greenbug share the same feature as both are dominant resistance genes. Given consideration of the fact that the two genes locate on two different chromosomes, it may be easier to pyramid both genes in a hybrid through the traditional breeding; thus, the resultant new hybrid will possess the resistance to sugarcane aphid as well as greenbug.

Based on the result of the QTL mapping, there are at least three expression units (i.e. three genes) associated with the SCA resistance QTL, sb06g001640, sb06g001645, and sb06g001650. Although it is possible that all three genes are involved in the host defense against sugarcane aphids, the experimental evidence showed that Sb1650 (i.e. sb06g001650) is the major genetic component responsible for the SCA resistance because the expression activity of the Sb1650 was closely correlated to the event of SCA attack according to the RT-PCR analysis (Fig. 3-3). The determination of Sb1650 as the candidate resistance is supported by the DNA sequence data (Fig. 3-5). There was a nine nucleotides deletion (GTGGTTTCC) from the SB1650 gene in BTx623 compared to the Sb1650 in Tx2783. Thus, it is believed that this deletion links the loss of function for the SCA resistance in cultivar BTx623 and other susceptible lines. More convincing further evidence is that DNA makers designed based on the deletion sequence can reliably detect

the progenies that are segregating the SCA resistance. Together, all experimental results and molecular data provided evidence for SB1650 as the sugarcane resistance gene.

A cDNA clone of the Sb1650 gene was sequenced, which is 3,069 base pairs in length. This gene encodes a hypothetical protein, meaning that it has never been characterized at the molecular level. Based on sequence annotation from genomic databases, the Sb1650 gene shares a weakly similar (50.0%) to a rice disease resistance protein, which belongs to the CC-NBS-LRR family. The CC-NBS-LRR gene family is known to mediate host resistance through direct or indirect recognition of pathogen-associated molecular patterns or pathogen effectors (Takken and Tameling 2009, Jones and Dangl 2006). Given the importance of this novel gene controlling sugarcane aphids, there is a need to focus on a further characterization in details at the molecular level in order to gain a better understanding of the mode of action for the gene and regulatory factors associated with its expression.

According to our observation, seedlings of sorghum showed a delay response to attack by sugarcane aphids when compared to their response to greenbugs. Phenotypically, seedlings of sorghum did not exhibit any symptoms during the first 7-8 days after SCA infestation, while sorghum seedlings start to show symptoms as soon as 3-4 days following greenbug infestation (Huang 2006). Similarly, expression of aphid-induced genes was also much quicker during host plant-greenbug interaction (Park et al. 2006) compared to activity of the SCA resistance gene as shown in Fig. 3-10. It is not known why sorghum seedlings react to the SCA slower than to greenbugs. It implies that some internal factors may play important role(s) in regulation of the host-plant defense against sugarcane aphids. Thus, in future studies, identification of regulatory factors that regulate

expression of the SCA resistance gene or modulate the defense pathway will contribute to a better understanding of the molecular mechanisms of SCA resistance.

This study has led to the identification of the first sorghum resistance gene to devastating sugarcane aphids, which is a valuable genetic resource for study of host-plant defense and the genetic mechanisms of SCA resistance. The genomic data and molecular information resulted from this study will benefit sorghum researchers and breeders worldwide who are interested in sorghum improvement for aphid resistance. DNA markers developed in this study offer a useful genomic tool to facilitate breeding sorghum with resistance to sugarcane aphids using marker-assisted selection as well as to be used for evaluation of new resistance sources.

References

Beckmann JS, Soller M. Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. Theoretical and Applied Genetics. 1983 Nov 1; 67(1):35-43.

Ben-Ari G, Lavi U. Marker-assisted selection in plant breeding. In Plant Biotechnology and Agriculture. 2012 (pp. 163-184).

Bowling RD, Brewer MJ, Kerns DL, Gordy J, Seiter N, Elliott NE, Buntin GD, Way MO, Royer TA, Biles S, Maxson E. Sugarcane aphid (Hemiptera: Aphididae): a new pest on sorghum in North America. Journal of Integrated Pest Management. 2016 Jan 1; 7(1).

Chang CP, Fang MN, Tseng HY. Studies on the life history and varietal resistance in grain sorghum aphid, Melanaphis sacchari Zehntner in central Taiwan. Chinese Journal of Entomology. 1982; 2:70-81.

Collard BC, Das A, Virk PS, Mackill DJ. Evaluation of 'quick and dirty' DNA extraction methods for marker- assisted selection in rice (Oryza sativa L.). Plant Breeding. 2007 Feb; 126(1):47-50.

Collard BC, Jahufer MZ, Brouwer JB, Pang EC. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica. 2005 Jan 1; 142(1-2):169-96.

Collard BC, Mackill DJ. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 2008 Feb 12; 363(1491):557-72.

David SK, Sandhu GS. New oviparous morph of Melanaphis sacchari (Zehntner) on sorghum. Entomologist's Record. 1976; 88(1):28-9.

Dogimont C, Bendahmane A, Chovelon V, Boissot N. Host plant resistance to aphids in cultivated crops: genetic and molecular bases, and interactions with aphid populations. Comptes Rendus Biologies. 2010 Jun 1; 333(6-7):566-73.

Ejeta G, Knoll JE. Marker-assisted selection in sorghum. In Genomics-assisted crop improvement 2007 (pp. 187-205). Springer, Dordrecht.

Evers L. Sugarcane Aphid in Sorghum. Natural Sciences Education. 2018 Jan 1; 47(1).

Hasan MM, Rafii MY, Ismail MR, Mahmood M, Rahim HA, Alam MA, Ashkani S, Malek MA, Latif MA. Marker-assisted backcrossing: a useful method for rice improvement. Biotechnology & Biotechnological Equipment. 2015 Mar 4; 29(2):237-54.

Hash CT, Raj AB, Lindup S, Sharma A, Beniwal CR, Folkertsma RT, Mahalakshmi V, Zerbini E, Blümmel M. Opportunities for marker-assisted selection (MAS) to improve the feed quality of crop residues in pearl millet and sorghum. Field Crops Research. 2003 Oct 1; 84(1-2):79-88.

Huang Y. Improvement of crop protection against greenbug using the worldwide sorghum germplasm collection and genomics-based approaches. Plant Genetic Resources. 2011 Jul; 9(2):317-20.

Huang Y. Phloem feeding regulates the plant defense pathways responding to both aphid infestation and pathogen infection. InBiotechnology and Sustainable Agriculture 2006 and Beyond 2007 (pp. 215-219). Springer, Dordrecht.

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

Kelly JD, Miklas PN. Marker-assisted selection. In Common bean improvement in the twenty-first century 1999 (pp. 93-123). Springer, Dordrecht.

Knoll J, Ejeta G. Marker-assisted selection for early-season cold tolerance in sorghum: QTL validation across populations and environments. Theoretical and Applied Genetics. 2008 Feb 1; 116(4):541-53.

Kromdijk J, Głowacka K, Leonelli L, Gabilly ST, Iwai M, Niyogi KK, Long SP. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. Science. 2016 Nov 18; 354(6314):857-61.

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001 Dec 1; 25(4):402-8.

Mickelbart MV, Hasegawa PM, Bailey-Serres J. Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. Nature Reviews Genetics. 2015 Apr; 16(4):237.

Mondal S, Rutkoski JE, Velu G, Singh PK, Crespo-Herrera LA, Guzman C, Bhavani S, Lan C, He X, Singh RP. Harnessing diversity in wheat to enhance grain yield, climate resilience, disease and insect pest resistance and nutrition through conventional and modern breeding approaches. Frontiers in Plant Science. 2016 Jul 6; 7:991.

Mullet J, Morishige D, McCormick R, Truong S, Hilley J, McKinley B, Anderson R, Olson SN, Rooney W. Energy Sorghum—a genetic model for the design of C4 grass bioenergy crops. Journal of Experimental Botany. 2014 Jun 22; 65(13):3479-89.

Neimann-Sorensen A, Robertson A. The association between blood groups and several production characteristics in three Danish cattle breeds. Acta Agriculturae Scandinavica. 1961 Jan 1; 11(2):163-96.

Ort DR, Merchant SS, Alric J, Barkan A, Blankenship RE, Bock R, Croce R, Hanson MR, Hibberd JM, Long SP, Moore TA. Redesigning photosynthesis to sustainably meet global

food and bioenergy demand. Proceedings of the National Academy of Sciences. 2015 Jul 14; 112(28):8529-36.

Park SJ, Huang Y, Ayoubi P. Identification of expression profiles of sorghum genes in response to greenbug phloem-feeding using cDNA subtraction and microarray analysis. Planta. 2006 Apr 1; 223(5):932-47.

Punnuri S, Huang Y, Steets J, Wu Y. Developing new markers and QTL mapping for greenbug resistance in sorghum [Sorghum bicolor (L.) Moench]. Euphytica. 2013 May 1; 191(2):191-203.

Ribaut JM, Hoisington D. Marker-assisted selection: new tools and strategies. Trends in Plant Science. 1998 Jun 1; 3(6):236-9.

Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. In Bioinformatics Methods and Protocols 2000 (pp. 365-386). Humana Press, Totowa, NJ.

Scott Armstrong J, Rooney WL, Peterson GC, Villenueva RT, Brewer MJ, Sekula-Ortiz D. Sugarcane aphid (Hemiptera: Aphididae): host range and sorghum resistance including cross-resistance from greenbug sources. Journal of Economic Entomology. 2015 Jan 30; 108(2):576-82.

Setokuchi O. Ecology of Longiunguis sacchari (Zehentner) (Aphididae) infesting sorghums. III. Occurrence in fields. In Proceedings of the Association for Plant Protection of Kyushu 1975 (Vol. 21, pp. 8-10).

Simpson PJ, Smith CB, Rosenthal G, Lucchesi BR. Reduction in the incidence of thrombosis by the thromboxane synthetase inhibitor CGS 13080 in a canine model of coronary artery injury. Journal of Pharmacology and Experimental Therapeutics. 1986 Aug 1; 238(2):497-501.

Singh BU, Padmaja PG, Seetharama N. Biology and management of the sugarcane aphid, Melanaphis sacchari (Zehntner) (Homoptera: Aphididae), in sorghum: a review. Crop Protection. 2004 Sep 1; 23(9):739-55.

Soller M, Beckmann JS. Genetic polymorphism in varietal identification and genetic improvement. Theoretical and Applied Genetics. 1983 Nov 1; 67(1):25-33.

Soller M. The use of loci associated with quantitative effects in dairy cattle improvement. Animal Science. 1978 Oct; 27(2):133-9.

Stuber CW, Goodman MM, Moll RH. Improvement of Yield and Ear Number Resulting from Selection at Allozyme Loci in a Maize Population 1. Crop Science. 1982; 22(4):737-40.

Takken FL, Tameling WI. To nibble at plant resistance proteins. Science. 2009 May 8; 324(5928):744-6.

Tanksley D, Medina-Filho H, Rick CM. The effect of isozyme selection on metric characters in an interspecific backcross of tomato—basis of an early screening procedure. Theoretical and Applied Genetics. 1981 Sep 1; 60(5):291-6.

Technow F, Messina CD, Totir LR, Cooper M. Integrating crop growth models with whole genome prediction through approximate Bayesian computation. PloS One. 2015 Jun 29; 10(6):e0130855.

Voytas DF. Plant genome engineering with sequence-specific nucleases. Annual Review of Plant Biology. 2013 May 2; 64.

Wang YS, Pao ST, Kwei CM, Chang GS, Chu HF. Studies on the sorghum aphid, Aphis sacchari Zehntner. Acta Entomologica Sinica. 1961; 10:363-80.

Wu Y, Huang Y. Molecular mapping of QTLs for resistance to the greenbug Schizaphis graminum (Rondani) in Sorghum bicolor (Moench). Theoretical and Applied Genetics. 2008 Jun 1; 117(1):117.

APPENDICES

Table S3-1. List of F_6 progenies (30S30R) were selected for marker verification (Forward: CGGCATTTATCAATTTGAGTGGTTTCC, Reverse: TCTAACGACATCAGGGTCAGG).

Phenotype	OSU Field 2018 Stake #	Seed source	Line	Notes	Genotype
R	4	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	26	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	36	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	40	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	46	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	47	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	50	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	63	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	74	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	103	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	108	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	170	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	201	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	203	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	220	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	221	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	245	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	260	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	279	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	282	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	291	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	1196	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	32	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	33	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	71	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	174	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	200	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	235	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	272	PR 2015w F5 seed	C23	BTx623/Tx2783	+
R	1198	PR 2015w F5 seed	C23	BTx623/Tx2783	+
S	13	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	44	PR 2015w F5 seed	C23	BTx623/Tx2783	-

S	70	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	115	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	117	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	121	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	125	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	127	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	134	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	142	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	151	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	163	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	165	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	172	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	188	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	199	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	202	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	209	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	212	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	215	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	223	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	238	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	247	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	252	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	262	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	263	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	285	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	286	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	287	PR 2015w F5 seed	C23	BTx623/Tx2783	-
S	290	PR 2015w F5 seed	C23	BTx623/Tx2783	-

*PR stands for Puerto Rico.

VITA

Jian Huang

Candidate for the Degree of

Doctor of Philosophy

Thesis: QTL MAPPING AND MOLECULAR CHARACTERIZATION OF THE SUGARCANE APHID [*Melanaphis Sacchari* (Zehntner) (Hemiptera: Aphididae)] RESISTANCE GENE IN SORGHUM [*Sorghum bicolor* (L.) Moench]

Major Field: Crop Science

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Crop Science at Oklahoma State University, Stillwater, Oklahoma in May, 2019.

Completed the requirements for the Master of Science in Biochemistry and Molecular Biology at Shenyang Normal University, Shenyang, Liaoning, China in 2012.

Completed the requirements for the Bachelor of Science in Biotechnology at Wuhan Polytechnic University, Wuhan, Hubei, China in 2008.

Experience:

Research experience: Research Assistant at Oklahoma State University, Stillwater, OK, 2013-2019. Research Engineer at U-Me Biotech (Takara), Wuhan, China, 2012-2013. Visiting Research Staff at U.S. Department of Agriculture-ARS, Stillwater, OK, 2011-2012. Research Assistant at Shenyang Normal University, Shenyang, China. 2009-2011.

Teaching experience: Lab Instructor at U.S. Department of Agriculture, Stillwater, OK, 2013-2019. Teaching Assistant for Graduate Seminar in Plant & Soil Sciences, 2017. Teaching Assistant for Plant Physiology, 2009-2011.

Professional Memberships:

Member of American Association for the Advancement of Science; Member American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.