# IDENTIFICATION AND MAPPING OF ANTHRACNOSE RESISTANCE GENES IN SORGHUM [*SORGHUM BICOLOR* (L.) MOENCH]

by

**Xiaochen Xu**

**A Thesis**

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

**Doctor of Philosophy**


Department of Agronomy

West Lafayette, Indiana

December, 2019

# THE PURDUE UNIVERSITY GRADUATE SCHOOL

# STATEMENT OF COMMITTEE APPROVAL

**Dr. Gebisa Ejeta, Chair**

Department of Agronomy

**Dr. Tesfaye Mengiste**

Department of Botany and Plant Pathology

**Dr. Jianxin Ma**

Department of Agronomy

**Dr. Rebecca Doerge**

Mellon College of Science, Carnegie Mellon University

**Approved by:**

Dr. Ronald F. Turco

For my dear wife Shuyin Liang and our lovely son Gabriel Xu

For my father Zhiqiang Xu and mother Zhangping Shen

# ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my major professor, Dr. Gebisa Ejeta, and co-advisor, Dr. Tesfaye Mengiste, for their incessant guidance, encouragement and support during all my years at Purdue University. Their rigorous academic attitude and enthusiasm for science profoundly impressed me.

I would also like to thank my advisory committee, Dr. Rebecca Doerge, Dr. Yiwei Jiang and Dr. Jianxin Ma, for their patient and valuable suggestions and inspiration.

My sincere thanks go to Dr. Patrick Rich, Dr. Adedayo Adeyanju and Dr. Fuyou Fu for teaching me experimental methods, analysis approaches and writing skills. Terry Lemming helped me enhance my plant breeding skills and advance my research population. Present and former members in Professor Ejeta's lab, Dr. Daniel Gobena, Dr. Patrick Ongom, Dr. Luis Rivera, Megan Khangura, Amanda Easterly, Stephanie Loehr and Mike White, also provided tremendous support to my research. Thank you to all. I would also like to thank the faculty, staff and graduate students from the Agronomy Department and other Purdue departments who provided insights and direction to my research and assistance during my graduate studies, including Patricia Oliver, Karen Clymer, Connie Foster, Dr. Chao-Jan Liao, Demeke Bayable, Habte Chikssa, Weiran Li, Longfei Wang, Weidong Wang and William Foster.

Last but not least, heartfelt thanks go to my family, especially my parents, my wife and our son, for their love, care and unconditional support.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

Xu, Xiaochen. Ph.D., Purdue University, December 2019. Identification and Mapping of Anthracnose Resistance Genes in Sorghum [*Sorghum bicolor* (L.) Moench]. Major Professor: Gebisa Ejeta.

*Colletotrichum sublineolum* is the causal agent of sorghum anthracnose, a very common and destructive fungal disease in warm and humid areas, especially in West and Central Africa. Use of host plant resistance is considered as the most important and effective control option for sorghum diseases. To achieve this goal, identification and mapping resistance genes is essential. In this study, we used an isolate of *C. sublineolum*, CsGL1, to screen our sorghum germplasm and identified a resistant inbred line, P9830. We developed a mapping population from a cross between P9830 and a susceptible line, TAM428, for this research. The population was advanced to the $F_6$ generation. Progenies were phenotyped at $F_2$, $F_3$ and $F_6$ generations for disease resistance against the pathogen, CsGL1. In the $F_2$ generation, 460 individuals showed resistance and 149 individuals showed susceptibility to CsGL1. This result fits the 3:1 segregation pattern expected for resistance controlled by a single gene. Bulked segregant analysis with next generation sequencing was used on selected $F_6$ recombinant inbred lines. A significant peak containing 153 SNPs was observed on the distal end of the long arm of chromosome 8. To verify resistance to CsGL1 was controlled by genes in this region, indel and SNP markers were used between 59.4Mbp and 60.6Mbp on chromosome 8 to fine map the resistance locus. One SNP marker located in the gene *Sobic.008G166400* co-segregated with resistance, and another two indel markers were discovered to be tightly linked to the resistance locus. These three PCR-based SNP markers would be useful for marker-assisted selection for improving anthracnose resistance against CsGL1. Two candidate genes, *Sobic.008G166400* and *Sobic.008G166550*, were found in the locus. Both of the genes encode LRR proteins implicated in plant disease defense response. The identity of DNA sequence between these two candidate genes is 94.1%, possibly the result of tandem duplication. Another

possible ortholog in the region is *Sobic.008G167500*. Quantitative PCR analysis showed that the expression level of *Sobic.008G166400* didn't change significantly in a resistant RIL, 17-12 but was induced in a susceptible RIL, 13-31, after CsGL1 infection. In conclusion, we mapped two candidate genes conferring resistant to CsGL1 on chromosome 8, and *Sobic.008G166400* is more likely of the two to be determined as the gene controlling resistance to CsGL1.

# CHAPTER 1. LITERATURE REVIEW

## 1.1 Introduction

This chapter provides a general background for the rational and used approaches and results obtained in this research aimed at identifying genes controlling anthracnose, a major constraint to sorghum production. This information includes literature in the concept of plant disease and plant pathology, mechanisms of plant disease resistance, anthracnose resistance in sorghum and other crops, and techniques for QTL (quantitative trait loci) mapping.

Sorghum (*Sorghum bicolor* (L.) Moench) is one of the most important crops in the grass family Poaceae. It ranks sixth in area harvested among all the crops in the world in 2017, only behind wheat (*Triticum aestivum* L.), maize (*Zea mays* L.), rice (*Oryza sativa* L.), soybean (*Glycine max* L.) and barley (*Hordeum vulgare* L.) (FAO 2017). Sorghum is well adapted to drought prone and hot environments, possibly due to its tropical and sub-tropical origin in Africa (Doggett, 1988; Smith and Frederiksen, 2000). Despite its excellent drought and heat tolerance, sorghum also has many other advantages including a relatively low input requirement, C4 photosynthesis, superior water use efficiency and comprehensive usages. The crop is widely planted in Africa, the Americas, Asia and Australia (Mall *et al.*, 2011). It is an important source of food, feed and beverage in Africa, Central America, South Asia, and China, but is primarily used for feed and cellulosic biofuel in the USA and Europe.

Plants, as sessile organisms, have to face unfavorable or stressful environments during growth. These conditions can be divided into two categories: biotic and abiotic stresses. Abiotic stress is the negative impact caused by non-living factors including drought, extreme temperatures, nutrient deficiency, salinity and toxic metals in the soil (Zhu, 2016). Biotic stress is caused by living organisms, including fungi, bacteria, viruses, nematodes, pests and weeds (Suzuki *et al.*, 2014). Diseases caused by fungi, are a major source of biotic stress in crop plants resulting in significant yield losses.

Sorghum anthracnose caused by *Colletotrichum sublineolum*, is prevalent in warm and humid environments in Africa, Asia, and the Americas, where sorghum is widely grown. These are major production areas accounting for over 95% of the world sorghum production in 2016 (FAO 2017). Sorghum anthracnose is prevalent and considered of primary importance in most parts of West and Central Africa (Marley *et al.*, 2005). It has been estimated that grain yield loss up to 70% may result from anthracnose infection (Singh *et al.*, 2006), though it's often difficult to precisely estimate (Ngugi *et al.*, 2000). There are several control options for sorghum anthracnose. Host plant resistance is considered as the most important and effective way to control sorghum diseases, although the resistance may be short-lived because of the highly variable pathogen forms (Rosenow and Frederiksen, 1982). Genetic control is also an environmental friendly and relatively economical method to control plant diseases. Crop-residue management is another method to control the disease because the pathogen may survive in crop debris on the soil surface between cropping seasons (Marley *et al.*, 2005). Removing crop residue can significantly reduce the incidence and severity of sorghum anthracnose. Fungicides are another control option for sorghum anthracnose. Seed dressing with the fungicide Apron Plus (metalaxyl, carboxin and furathiocarb) has been shown to effectively reduce the severity of leaf anthracnose in Nigeria (Akpa *et al.*, 1992; Gwary *et al.*, 2008). Foliar spray of benomyl is another effective fungicide to control sorghum anthracnose in Nigeria (Marley, 1996; Marley, 1997). However, controlling anthracnose by the application of fungicides is not environmental friendly and may aggravate the financial burden on African farmers.

## 1.2 The concept of plant disease and plant pathology

A useful definition of plant disease was put forth by Agrios (2005) as "a series of invisible and visible reactions and responses of plant cells and tissues to a pathogenic organism or environmental condition that lead to adverse changes in the form, function, or integrity of the plant and may cause partial impairment or death of plant parts or the entire plant." The agents causing infectious or biotic diseases in plants include pathogenic fungi, viruses, bacteria, protozoa, and nematodes. Air- or soil-borne toxic chemicals are other causal agents of noninfectious or abiotic disease. Diseases

in plants may result in both direct and indirect losses. The direct losses caused by plant disease include direct yield loss, reduction of seed or grain quality, extra cost of controlling, harvesting, grading and replanting, contamination of farming materials or tools, and soil-borne disease (Zadoks, 1967; Zadoks and Schein, 1979). Examples of indirect losses include losses to farmers, rural communities, exporters, trades, consumers, government and environment (Zadoks, 1967; Zadoks and Schein, 1979). It is believed that 20 to 40% of global agricultural productivity losses including both direct and indirect losses are caused by pathogens, insects and weeds (Teng and Gaunt, 1980; Oerke *et al*., 1994; Oerke, 2006). On average, plant disease alone is responsible for 14.1% of annual worldwide crop loss, which is valued at $220 billion (Agrios, 2005). In the United States, wheat rusts alone are estimated to result in a $5 billion yield loss annually (Savary *et al.*, 2012).

Plant pathology is defined as an integrative science that combines the knowledge of botany, biochemistry, plant physiology, genetics, molecular biology and genetic engineering, mycology, bacteriology, virology, nematology and many other branches of science applied to study the diseases and disorders of plants (Agrios, 2005). Phytopathology is a complicated discipline aimed at improving not only people's cognizance of plant diseases and their pathogenic agents, but also the methods, equipment and ability to avoid or control plant diseases. For example, phytopathologists and crop breeders spent many years identifying and introducing resistance genes into plants to reduce the use of toxic chemicals in disease management. It is still a big challenge for plant pathologists to use modern techniques and approaches to reduce the environmental cost while protecting crop yield from plant diseases and improving agricultural product quality.

## 1.3 Development of fungal disease in plants

Plants are always exposed to changing environmental conditions and innumerable microorganisms, but few of these interactions lead to diseases. The disease triangle can mainly explain this common phenomenon. The concept of the disease triangle was first introduced by George McNew in 1960 (Fig. 1.1). There are three important conditions interacting with each other to determine the potential that a disease takes place on the host: a susceptible host, a virulent

pathogen and a conducive environment for the pathogen. Other researchers (e.g. Scholthof, 2007) illustrated the triangle with six parameters: the age or growth stage of the host plant, the susceptibility of the host plant to the pathogen, virulence of the pathogen, prevalence of the pathogen, the duration of the infection period, and the severity of the environment. The disease triangle clearly demonstrates the interrelationship of the three factors in an epidemic, and it also indicates that we may limit or control a plant disease by eliminating any one of the three causal components.

Even when all the three requirements in the disease triangle are achieved, there are still a series of distinct stages that occur in succession and finally lead to the appearance, development and perpetuation of the disease and the pathogen (Agrios, 2005). These distinct stages form a closed ring called disease cycle (Fig. 1.2). The disease cycle involves changes and symptoms on the plant as well as those in the pathogen. This literature review will mainly focus on fungal diseases.

Inoculation is defined as the initial contact of a pathogen with the plant. The inoculum is the pathogen(s) that is able to land on or in contact with the plant, and it can be any part of the pathogen that initiates infection (Agrios, 2005). The inoculum could be spores, sclerotia or fragments of mycelium in fungi. Whole individuals of bacteria and viruses are also inoculums. Inoculum can be a single individual of a pathogen such as one spore, or millions of individuals of a pathogen. Inoculum can be primary or secondary. Primary inoculum survives the winter or summer and causes the original infections called primary infections in the next season. Secondary inoculum is produced from primary infections and hence causes secondary infections (Agrios, 2005).

Unlike most of the viruses that are placed directly into plant cells by their vectors, most fungi and bacteria, usually make contact with the external surface of the plant first, and therefore need to attach to the host surface before they can penetrate and colonize the host. In fungi, adhesion of spores to the plant surface is the first step in committing a pathogen to the establishment of disease, although it is still not known how exactly fungal spores adhere to the plant surface. Currently, it is known that the fungal cell wall is not directly involved in sensing stimuli, but certain cell wall associated proteins like Msb2 (multicopy suppressor of a budding defect) are involved (Geoghegan

*et al*., 2017). Deletion of *MSB2* gene in some phytopathogenic fungi lead to defects in appressorium (a specialized cell typical of many fungal plant pathogens that is used to infect host plants) formation and development and in penetration of plant tissues (Lanver *et al.*, 2010; Liu *et al.*, 2011; Perez-Nadales and Di Pietro, 2011; Leroch *et al.*, 2015). To successfully initiate infection, appressoria must tightly adhere to the host plant surface after they are formed. Fungal spores always show organ specificity, which means normally they do not attack all parts of their host plant. Factors in determining this tissue specificity may involve surface lectins, ions or hydrophobicity. However, the exact mechanism of the interaction between spores and host surfaces is still not clear. Both signal perception of the host plant and environmental effects enable spore germination, germ tube formation and extension, and appressoria formation (Tucker and Talbot, 2001). Signaling pathways during the pre-penetration stage in different fungal pathogens might be mediated by different regulators, including cyclic adenosine monophosphate (cAMP) (Xu *et al.*, 1997; Adachi and Hamer, 1998; Ramanujam and Naqvi, 2010), motigen-activated protein kinase (MAPK) (Xu and Hamer, 1996; Kojima *et al.*, 2002) and G proteins (Takai *et al*., 2001; Li *et al.*, 2007). Hyphae are produced by appressoria, and cAMP and MAPK are required in this process.

To infect the host plants, pathogens need to penetrate plant cells. Penetration does not necessarily result in disease infection if plants are not susceptible to the pathogen. In this situation, pathogens are not able to cause disease since they cannot proceed beyond the penetration stage. Fungi can penetrate plant surfaces in several different ways such as direct penetration (Knogge, 1996), penetration through wounds (Wheeler, 1968) and penetration through natural openings (stomata, hydathodes, nectarthodes and lenticels) (Hoch *et al*., 1987; Correa Jr. and Hoch, 1995).

After penetration of plant surfaces, the following process by which pathogens initiate contact with susceptible cells or tissues of plants and absorb nutrients from them is called infection. A successful infection usually leads to the appearance of disease symptoms on the host in a few days to a few weeks after inoculation. Some viral diseases (e.g. sweet potato virus disease), however, may not show any symptoms for over one year after inoculation (Gibson *et al.*, 1998). After infection, pathogens with different lifestyles procure nutrients from hosts in different ways.

Biotrophic pathogens (e.g. *Puccinia graminis*) capture nutrients form living cells, and therefore do not kill host cells rapidly. Necrotrophic pathogens (e.g. *Botrytis cinerea*) kill host cells rapidly and obtain nutrients from dead cells. Hemibiotrophic pathogens (e.g. *Colletotrichum sublineolum*) initiate with biotrophy and then turn to necrotrophy. After infection, pathogens invade host plants through different ways and start reproducing to infect more host tissues. Most fungi produce a mycelium within the host plants they infect and produce spores on or below surfaces of the infected parts of the host, so that spores can be easily released to the environment.

Dissemination of pathogens often leads to disease aggravation and outbreak, and therefore causes significant economic loss (Anderson *et al.*, 2004). Wind is a common way of dissemination for most fungal spores. Water, insects, nematodes, pollen, seed and humans also play an important role in disseminating pathogens.

Pathogens infecting perennial plants survive in hosts during winter, summer or other stress conditions. Pathogens infecting annual plants have different mechanisms to survive under those stresses. Fungal mycelium and spores may survive on plant debris during summer and winter (Yunis and Elad, 1989; Jurick *et al.*, 2008). Some fungal pathogens like *Fusarium culmorum* are soil inhabitants with a high competitive saprophytic ability and able to survive for a long time (Sitton and Cook, 1981; Bateman and Murray, 2001). Rust fungi may survive on volunteer plants.

## 1.4 Mechanisms of resistance to pathogens

Over millions of years, plants have evolved different defense mechanisms against pathogens. Based on different classification method, we can classify those defense mechanisms into different categories.

Non-host resistance (NHR), defined as resistance shown by an entire plant species to all isolates of a specific parasite or pathogen (not pathogen-race-specific), is the most common and durable type of plant resistance to pathogens (Heath, 1985; Heath, 2000). For example, a fungal pathogen causing anthracnose in sorghum, *Colletotrichum sublineolum*, cannot infect any parts of maize or wheat at all. Similarly, *Colletotrichum graminicola*, the causal agent of anthracnose in maize, cannot lead to diseases in sorghum. Non-host resistance is attractive to researchers and

widely studied due to its long durability and broad resistance. Non-host resistance, however, is identified as a quantitative trait and extremely complex. There are multiple mechanisms explaining NHR, although many details are still not clear (Mysore and Ryu, 2004). Based on the previous research, both constitutive and induced defense, and their combination contribute to NHR (Thordal-Christensen, 2003).

Host resistance is specific to certain isolates of a pathogen and less durable, compared with NHR (Gill *et al.*, 2015). It is also less complex than NHR. Introduction of host resistance is considered as one of the most practical and important ways to control plant diseases (Rudd *et al.*, 2001). There are overlaps between host and non-host resistance.

Wax in the cuticle (cuticular) and outside the cuticle (epicuticular) plays important roles in NHR. Uppalapati *et al.* (2012) observed significant reduction of germ tubes of two rust pathogens, *Phakopsora pachyrhizi* and *Puccinia emaculata*, on the abaxial leaf surfaces of *inhibitor of rust germ tube differentation1* (*irg1*) mutants of *Medicago truncatula*. This was caused by a significant reduction of the C30 primary alcohol component of epicuticular waxes altering the surface hydrophobicity in *irg1* mutants.

The plant cell wall is a formidable physical barrier against pathogens. Additionally, the cell wall also involved in other mechanisms of plant resistance. Oxidative insolubilization of cell wall structural proteins, p33 and p100, also leads to a hypersensitive reaction (HR) and increased NHR (Brisson *et al*., 1994). Adhesion between the plant cell wall and plasma membrane contributes to defending against fungal penetration (Mellersh and Heath, 2001). The cell wall is also involved in signal transduction to induce subsequent resistance (Huckelhoven, 2007).

Secondary metabolites confer resistance to pathogens as chemical barriers. Plants produce over 100,000 secondary metabolites with distinct forms (Dixon, 2001). Phytoalexins are typical secondary metabolites closely associated with plant disease resistance. Camalexin is a phytoalexin widely studied in *Arabidopsis*, and the production and accumulation of camalexin can be triggered by cytolytic toxins such as fumonisin B1 produced by *Fusarium* spp. (Stone *et al.*, 2000). Camalexin may also be induced by pathogen- or microbe-associated molecular patterns

(PAMPs/MAMPs) such as necrosis and ethylene-inducing peptide1–like proteins (Qutob *et al.*, 2006) and abiotic stresses such as abiotic elicitor α-amino butyric acid (Zhao *et al.*, 1998). Camalexins seem to have damaging effects on bacterial and fungal membranes (Rogers *et al.*, 1996; Joubert *et al.*, 2011), although the details of that injury are not fully understood. Fungal programmed cell death in *Botrytis cinerea* are induced by camalexins, produced by *Arabidopsis* (Shlezinger *et al.*, 2011). More types of phytoalexins were detected and isolated from crop plants, and most of them contribute to plant defense against pathogens (Ahuja *et al*., 2012). For example, 3-Deoxyanthocyanidins are phytoalexins which are significantly induced by *Cochliobolus heterostrophus* in sorghum seedlings, and these phytoalexins confer resistance to *Colletotrichum sublineolum*, the causal agent of sorghum anthracnose (Lo *et al*., 1999; Liu *et al.*, 2010). Other secondary metabolites (e.g. phenolic compounds) are also involved in mechanisms against plant diseases (Mazid *et al*., 2011).

Both non-host and host resistance involve in pathogen associated molecular patterns (PAMP)-triggered immunity (PTI) and effector-triggered immunity (ETI). Pathogen- or microbe-associated molecular patterns are defined as conserved molecular structures produced by microorganisms and recognized as foreign by the receptors of the innate immune system. Pathogen effectors are proteins expressed and secreted by pathogens and released into plant cells to aid infection of specific plant species. A four phased 'zigzag' model was developed by Jones and Dangl (2006) to describe the plant immune system.

Pathogen associated molecular patterns-triggered immunity was considered as the first line of plant active defense (Chisholm *et al.*, 2006). Pathogen associated molecular patterns are first recognized by transmembrane pattern recognition receptors (PRRs). Flagellin-sensing 2 (FLS2) is a widely studied PRR containing leucine-rich repeat (LRR) domain in *Arabidopsis* (Zipfel *et al.*, 2004). Flagellin-sensing 2 is able to recognize flg22, a typical PAMP containing highly conserved domain, and lead to PTI (Gomez-Gomez and Boller, 2000). Pathogen associated molecular patterns-triggered immunity can be activated in a very short time and usually does not cause HR. The very early responses of PTI (seconds to minutes) could be ion fluxes, oxidative bursts,

activation of MAPKs and changes in protein phosphorylation (Schwessinger and Zipfel, 2008). Within one hour after PAMP treatment, receptor endocytosis, induction of ethylene biosynthesis and stomatal closure were observed as parts of PTI (Ron and Avni, 2004; Melotto *et al.*, 2006). Multiple signaling pathways are involved in PTI, such as $Ca^{2+}$ bursts, reactive oxygen species (ROS) bursts, protein kinases and phytohormones (salicylic acid (SA), jasmonic acid (JA) and ethylene (ET)) (Bigeard, Colcombet and Hirt, 2015).

Effectors released by pathogens can interfere with PTI, and lead to effector-triggered susceptibility (Jones and Dangl, 2006). If effectors can be recognized by nucleotide binding and leucine-rich repeat (NB-LRR) proteins, Effector-triggered immunity will initiate to build a robust resistance against pathogens. Effector-triggered immunity often causes programmed cell death and HR in plants. Effectors can be directly and indirectly recognized by NB-LRR proteins, or recognized by NB-LRR protein pairs. It was reported that a fungal effector AvrL567 produced by *Melampsora lini* (the causal agent of flax rust disease) can be directly recognized by L5 and L6 alleles of *L* locus resistance proteins (Dodds *et al.*, 2006; Ravensdale *et al.*, 2012). Kim *et al.* (2005) found that two resistance proteins, RPS2 and RPM1, recognized type III effectors produced by *Pseudomonas syringae* and guarded RIN4 protein in *Arabidopsis*, which is an indirect recognition model. Williams *et al.* (2014) reported that two NB-LRR proteins, RPS4 and RPS1, formed a RPS1/RPS4 effector recognition complex, and each of these proteins plays a distinct role in recognition and signaling. Although ETI shares some signaling mechanisms such as ROS, MAPKs and phytohormones with PTI, there are certain differences in these signaling mechanisms between these two immunity systems (Tsuda and Katagiri, 2010). For example, it is possible that SA, JA and ET signaling can all be activated during some cases of PTI and ETI. In PTI, however, synergistic relationships among the signaling sectors were found while in ETI, compensatory relationships among the signaling sectors were identified, and these results explained why ETI is more robust than PTI (Tsuda *et al.*, 2009).

### 1.5 Anthracnose resistance in sorghum and other plants

Anthracnose disease caused by *Colletotrichum* species is a worldwide issue and affects over

42 genera of plants in the family Poaceae (Crouch and Beirn, 2009) and members of other plant families such as Fabaceaeand Solanaceae (Kelly and Vallejo, 2004; Hussain and Abid, 2011).

Sorghum anthracnose, caused by *Colletotrichum sublineolum*, is a fungal disease prevalent in warm and humid environments in Africa, Asia, and the Americas, causing significant sorghum yield losses. The pathogen responsible for sorghum anthracnose was considered to be hypervariable, which made it difficult to study. Cardwell *et al*. (1989) reported 12 isolates of *C. sublineolum* from Texas, Georgia and Puerto Rico, and these isolates showed eight pathotypes. Based on the study at six different locations in India, Thakur *et al.* (1998) classified isolates of *C. sublineolum* into six pathotypes. Moore *et al*. (2008) tested 98 isolates and identified 13 pathotypes.

Several important sorghum inbred lines showing stable resistance to anthracnose were reported, and studies on anthracnose resistance were conducted based on these lines. Mehta *et al.* (2005) reported a sorghum inbred line SC748-5 with stable resistance to *C. sublineolum* across five environments, and SC748-5 was also used by other researchers in anthracnose resistance studies. For example, Perumal *et al.* (2009) found a dominant resistance gene *Cg1* closely linked with an amplified fragment length polymorphism (AFLP ) marker *Xtxa6227* on the distal arm of LG-05 (chromosome 5) in SC748-5. Burrell *et al.* (2015) developed a recombinant inbred population by crossing SC748-5 to BTx623, and also mapped a major effect QTL with conferring anthracnose resistance on chromosome 5. Cuevas *et al.* (2018) identified 75 resistant accessions (including SC748) in a sorghum association panel containing 335 accessions. They also found three candidate genes on chromosome 5 (consistent with the previous conclusions made by Metha *et al*. 2005 and Burrell *et al*. 2015) based on the result of a genome-wide association study (GWAS), and one of the candidate genes (*Sobic.005G182400*) encodes a LRR protein and may play a role in signal transduction in disease defense response. SC112-14 (PI 533918) was reported to be consistently resistant to *C. sublineolum* isolates from Arkansas, Georgia, Texas and Puerto Rico, and several resistance loci were mapped on chromosome 5 based on the population developed by SC112-14 (Cuevas *et al*., 2014). In that study, three resistance loci were not associated with *Cg1*, although all of them are at the distal region of chromosome 5. Patil *et al.* (2017) reported a major

anthracnose-resistance QTL on chromosome 9 in SC155-14E, which was associated with strong resistance to an inoculum comprised of multiple isolates of *C. sublineolum*.

Anthracnose in common bean (*Phaseolus vulgaris*) is a very widespread and severe disease caused by *Colletotrichum lindemuthianum* (Kelly and Vallejo, 2004). Research on anthracnose resistance in common bean reported resistance genes such as *Co*-genes in the past decades. Young and Kelly (1997) developed several random amplified polymorphic DNA (RAPD) markers closely linked to anthracnose resistance genes, *Co-1*, *Co-5* and *Co-6*, for use in marker-assisted selection (MAS). Young *et al.* (1998) reported a highly anthracnose resistant and photoperiod sensitive line G2333 and another two dominant resistance genes, *Co-4*$^2$ and *Co-7*, based on their bulked segregant analysis (BSA). Mendez-Vigo *et al.* (2005) found that *Co-3* and *Co-9* are actually different alleles of the same resistance gene, rather than two different genes according to the allelism tests that they conducted.

Anthracnose fruit rot, usually caused by the fungal agent *Colletotrichum acutatum* (although *C. fragariae* and *C. gloeosporioides* have also been reported as causal agents but not as frequently as *C. acutatum*), is also a widely spread disease threatening strawberry production and quality. *Colletotrichum acutatum* mainly infects the fruit of strawberry but also observed to affect other organs such as flowers, crowns, leaves, petioles and roots (Mertely *et al*., 2009). Lerceteau-Kohler *et al*. (2005) conducted a BSA on anthracnose fruit rot, where they reported four AFLP markers tightly linked to the resistance gene *Rca2*, two of which were converted into sequence characterized amplified regions (SCAR) markers for MAS. Silva *et al.* (2015) reported that significant reduced petiole lesion symptoms were observed in transgenic strawberry which successfully expressed the *AtNPR1* gene, indicating that the *AtNPR1* gene contributes resistance to *C. acutatum.*

## 1.6 Techniques for QTL mapping

By combining advanced molecular breeding and sequencing technologies, traditional QTL mapping techniques such as BSA have become more efficient. Genomic analysis such as GWAS has emerged as another powerful tool for QTL mapping in plant breeding and trait improvement.

Bulked segregant analysis was developed by Michelmore *et al*. (1991) as a rapid and efficient tool to detect QTL or markers associated with a certain trait. To do BSA, usually a biparental segregating population needs to be developed, where individuals in the population are phenotyped for the trait of interest, and extreme phenotypes (individuals in the two tails of the phenotypic distribution) are selected and bulked into two pools (segregating bulks), respectively. The two pools will be screened for genotype frequencies. No significant differences in allele frequencies between the two bulks should be observed in genomic regions beyond loci controlling the trait. In genomic regions containing loci for the desirable trait, allele frequencies in two bulks show obvious differences. One of the advantages of BSA is that it just needs to genotype two bulks rather than genotyping every individual in the population, which can significantly decrease the cost and time. Bulked segregant analysis, however, may ignore loci with minor effects and may not be appropriate for complex quantitative traits controlled by many loci with small effects. While polymerase chain reaction (PCR) based DNA markers or microarrays were frequently used to genotype the two bulks in BSA studies in the past, genome sequencing techniques such as next-generation sequencing (NGS) are more and more commonly used in recent BSA studies (Magwene *et al*., 2011). By NGS-assisted BSA, Haase *et al.* (2015) reported 14 and 13 QTL regulating flowering time and plant height in maize, respectively. Bulked segregant RNA-seq (BSR-seq) is BSA combined with RNA-seq and also commonly used in QTL mapping. Unlike whole genome sequencing, BSR-seq can not only locate the causal gene, but it also reflects the effects of mutations on gene expression level (Liu *et al.*, 2012). However, the weakness of this technique is also very obvious. If causal mutations are located in non-expressed regions, they cannot be detected by BSR-seq analysis (Zou *et al*. 2016). Li *et al.* (2018) reported a recessive *early leaf senescence 1* (*els1*) gene controlling early leaf senescence in wheat variety M114 was mapped and several PCR-based markers were designed based on the BSR-seq study.

Genome-wide association study is a type of powerful mapping technique for QTL detection. Both of BSA and GWAS may use either F2 or RIL populations to map loci. However, unlike segregating populations in BSA, GWAS could use a population with diverse varieties of

phenotypes such as a multi-parent advanced generation inter-cross (MAGIC) population or an association panel (Huang and Han, 2014). Briefly, GWAS calculates and evaluates the association between genotypes and phenotypes of a certain desirable traits for a large number of individuals in the population (Korte and Ashley, 2013). Morris *et al.* (2013) developed a sorghum association panel with 971 worldwide accessions. They generated a genome-wide single nucleotide polymorphisms (SNPs) map with over 260,000 SNPs by genotyping-by-sequencing and located several loci and candidate genes for plant height and inflorescence architecture in sorghum by GWAS. Zhang *et al.* (2015) reported 27, 6, 18 and 27 loci for days to flowering, days to maturity, duration of flowering-to-maturity, and plant height, respectively, and several candidate genes associated with these traits in soybean based on their GWAS. In maize, Tian *et al.* (2011) identified over 200 SNPs associated with leaf angle, leaf length and width by GWAS, indicating that GWAS could be an effective method to discover the basis of key agronomic traits in plants.

## 1.7 References

Adachi K and Hamer JE (1998) Divergent cAMP signaling pathways regulate growth and pathogenesis in the rice blast fungus *Magnaporthe grisea*. *The Plant Cell* 10: 1361–1373

Agrios G (2005) Plant Pathology. 5th edn. Elsevier, Inc., New York

Ahuja I, Kissen R and Bones AM (2012) Phytoalexins in defense against pathogens. *Trends in Plant Science* 17: 73–90

Akpa AD, Gana SD and Alabi O (1992) Foliar fungicidal spray for the control of some common foliar diseases of sorghum in Nigeria. *Nigerian Journal of Plant Protection* 14: 92-96

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR and Daszak P (2004) Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology and Evolution* 19: 535–544

Bateman GL and Murray G (2001) Seasonal variations in populations of *Fusarium* species in wheat-field soil. *Applied Soil Ecology* 18: 117–128

Bigeard J, Colcombet J and Hirt H (2015) Signaling mechanisms in pattern-triggered immunity (PTI). *Molecular Plant* 8: 521–539

Brisson LF, Tenhaken R and Lamb C (1994) Function of oxidative cross-linking of cell wall structural proteins in plant disease resistance. *The Plant Cell* 6: 1703–1712

Burrell AM, Sharma A, Patil NY, Collins SD, Anderson WF, Rooney WL and Klein PE (2015) Sequencing of an anthracnose-resistant sorghum genotype and mapping of a major QTL reveal strong candidate genes for anthracnose resistance. *Crop Science* 55: 790–799

Cardwell KF, Hepperly PR and Frederiksen RA (1989) Pathotypes of *Colletotrichum graminicola* and seed transmission of sorghum anthracnose. *Plant Disease* 73: 255–257

Chisholm ST, Coaker G, Day B and Staskawicz BJ (2006) Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell* 124: 803–814

Correa Jr. A and Hoch HC (1995) Identification of thigmoresponsive loci for cell differentiation in *Uromyces* germlings. *Protoplasma* 186: 34–40

Crouch JA and Beirn LA (2009) Anthracnose of cereals and grasses. *Fungal Diversity* 39: 19–44

Cuevas HE, Prom LK, Cooper EA, Knoll JE and Ni X (2018) Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in the U.S. sorghum association panel. *The Plant Genome* 11: 170099

Cuevas HE, Prom LK and Erpelding JE (2014) Inheritance and molecular mapping of anthracnose resistance genes present in sorghum line SC112-14. *Molecular Breeding* 34: 1943–1953

Dixon RA (2001) Natural products and plant disease resistance. *Nature* 411: 843–847

Dodds PN, Lawrence GJ, Catanzariti A, The T, Wang CA, Ayliffe MA, Kobe B and Ellis JG (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences* 103: 8888–8893

Doggett H (1988) Sorghum. 2nd edn. John Wiley & Sons, Inc., New York

FAO (2017) Food and Agriculture Organization of the United Nations, Statistics Division. http://www.fao.org/faostat/en/#data/QC (Cited 20 February 2019)

Geoghegan I, Steinberg G and Gurr S (2017) The role of the fungal cell wall in the infection of plants. *Trends in Microbiology* 25: 957–967

Gibson RW, Mpembe I, Alicai T, Carey EE, Mwanga ROM, Seal SE and Vetten HJ (1998) Symptoms, aetiology and serological analysis of sweet potato virus disease in Uganda. *Plant Pathology* 47: 95–102

Gill US, Lee S and Mysore KS (2015) Host versus nonhost resistance: Distinct wars with similar arsenals. *Phytopathology* 105: 580–587

Gomez-Gomez L and Boller T (2000) FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis. Molecular Cell* 5: 1003–1011

Gwary DM, Bwatanglang N and Bdliya BS (2008) Integrated management of sorghum anthracnose through the use of fungicides, crop varieties and manipulation of sowing dates in Sudan Savanna of Nigeria. *International Journal of Agriculture & Biology* 10: 661–664

Haase NJ, Beissinger T, Hirsch CN, Vaillancourt B, Deshpande S, Barry K, Buell CR, Kaeppler SM and de Leon N (2015) Shared genomic regions between derivatives of a large segregating population of maize identified using bulked segregant analysis sequencing and traditional linkage analysis. *G3 Genes|Genomes|Genetics* 5: 1593–1602

Heath MC (1985) Implications of nonhost resistance for understanding host-parasite interactions. In: Groth JV and Bushnell WR (eds) *Genetic basis of biochemical mechanisms of plant disease*. APS Press, St. Paul, 25-42

Heath MC (2000) Nonhost resistance and nonspecific plant defenses. *Current Opinion in Plant Biology* 3: 315–319

Hoch HC, Staples RC, Whitehead B, Comeau J and Wolf ED (1987) Signaling for growth orientation and cell differentiation by surface topography in *Uromyces*. Science 235: 1659-1662

Huang X and Han B (2014) Natural variations and genome-wide association studies in crop plants. *Annual Review of Plant Biology* 65: 531–551

Huckelhoven R (2007) Cell wall–associated mechanisms of disease resistance and susceptibility. *Annual Review of Phytopathology* 45: 101–127

Hussain F and Abid M (2011) Pests and diseases of chili crop in Pakistan: A review. *International Journal of Biology and Biotechnology* 8: 325–332

Jones JDG and Dangl JL (2006) The plant immune system. *Nature* 444: 323–329

Joubert A, Bataille-Simoneau N, Campion C, Guillemette T, Hudhomme P, Iacomi-Vasilescu B, Leroy T, Pochon S, Poupard P and Simoneau P (2011) Cell wall integrity and high osmolarity glycerol pathways are required for adaptation of *Alternaria brassicicola* to cell wall stress caused by brassicaceous indolic phytoalexins. *Cellular Microbiology* 13: 62–80

Jurick II WM, Narvaez DF, Brennan MM, Harmon CL, Marois JJ, Wright DL and Harmon PF (2008) Winter survival of the soybean rust pathogen, *Phakopsora pachyrhizi*, in Florida. *Plant Disease* 92: 1551–1558

Kelly JD and Vallejo VA (2004) A comprehensive review of the major genes conditioning resistance to anthracnose in common bean. *HortScience* 39: 1196–1207

Kim MG, da Cunha L, McFall AJ, Belkhadir Y, DebRoy S, Dangl JL and Mackey D (2005) Two *Pseudomonas syringae* type III effectors inhibit RIN4-regulated basal defense in *Arabidopsis. Cell* 121: 749–759

Knogge W (1996) Fungal infection of plants. *The Plant Cell* 8: 1711-1722

Kojima K, Kikuchi T, Takano Y, Oshiro E and Okuno T (2002) The mitogen-activated protein kinase gene *MAF1* is essential for the early differentiation phase of appressorium formation in *Colletotrichum lagenarium*. *Molecular Plant-Microbe Interactions* 15: 1268–1276

Korte A and Farlow A (2013) The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* 9: 29

Lanver D, Mendoza-Mendoza D, Brachmann A and Kahmann R (2010) Sho1 and Msb2-related proteins regulate appressorium development in the smut fungus *Ustilago maydis*. *The Plant Cell* 22: 2085–2101

Lerceteau-Kohler E, Guérin G and Denoyes-Rothan B (2005) Identification of SCAR markers linked to *Rca2* anthracnose resistance gene and their assessment in strawberry germplasm. *Theoretical and Applied Genetics* 111: 862–870

Leroch M, Mueller N, Hinsenkamp I and Hahn M (2015) The signalling mucin Msb2 regulates surface sensing and host penetration via BMP1 MAP kinase signalling in *Botrytis cinerea.* *Molecular Plant Pathology* 16: 787–798

Li L, Wright SJ, Krystofova S, Park G and Borkovich KA (2007) Heterotrimeric G protein signaling in filamentous fungi. *Annual Review of Microbiology* 61: 423–452

Li M, Li B, Guo G, Chen Y, Xie J, Lu P, Wu Q, Zhang D, Zhang H, Yang J, Zhang P, Zhang Y and Liu Z (2018) Mapping a leaf senescence gene *els1* by BSR-Seq in common wheat. *The Crop Journal* 6: 236–243

Liu H, Du Y, Chu H, Shih CH, Wong YW, Wang M, Chu IK, Tao Y and Lo C (2010) Molecular dissection of the pathogen-inducible 3-deoxyanthocyanidin biosynthesis pathway in sorghum. *Plant and Cell Physiology* 51: 1173–1185

Liu S, Yeh C, Tang HM, Nettleton D and Schnable PS (2012) Gene mapping via bulked segregant RNA-seq (BSR-Seq). *PLoS ONE* 7: e36406

Liu W, Zhou X, Li G, Li L, Kong L, Wang C, Zhang H and Xu J (2011) Multiple plant surface signals are sensed by different mechanisms in the rice blast fungus for appressorium formation. *PLoS Pathogen* 7: e1001261

Lo SC, de Verdier K and Nicholson RL (1999) Accumulation of 3-deoxyanthocyanidin phytoalexins and resistance to *Colletotrichum sublineolum* in sorghum. *Physiological and Molecular Plant Pathology* 55: 263–273

Magwene PM, Willis JH and Kelly JK (2011) The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology* 7: e1002255

Mall TK, Dweikat I, Sato SJ, Neresian N, Xu K, Ge Z, Wang D, Elthon T and Clemente T (2011) Expression of the rice CDPK-7 in sorghum: Molecular and phenotypic analyses. *Plant Molecular Biology* 75: 467–479

Marley PS (1996) Sorghum Pathology. *Cereals Research Program Report to the 1996 Annual Cropping Scheme Meeting*. Institute for Agricultural Research, Ahmadu Bello University, Samaru, Zaria, Nigeria

Marley PS (1997) Sorghum Pathology. *Cereals Research Program Report to the 1997 Annual Cropping Scheme Meeting*. Institute for Agricultural Research, Ahmadu Bello University, Samaru, Zaria, Nigeria

Marley PS, Diourte M, Neya A and Rattunde FW (2005) Sorghum anthracnose and sustainable management strategies in West and Central Africa. *Journal of Sustainable Agriculture* 25: 43–56

Mazid M, Khan TA and Mohammad F (2011) Role of secondary metabolites in chemical defence mechanisms in plants. *Biology and Medicine* 3: 232–249

McNew GL (1960) The nature, origin, and evolution of parasitism. In: Horsfall JG and Dimond AE (eds) *Plant pathology: An Advanced Treatise*. Academic Press, New York, 19-69

Mehta PJ, Wiltse CC, Rooney WL, Collins SD, Frederiksen RA, Hess DE, Chisi M and TeBeest DO (2005) Classification and inheritance of genetic resistance to anthracnose in sorghum. *Field Crops Research* 93: 1–9

Mellersh DG and Heath MC (2001) Plasma membrane-cell wall adhesion is required for expression of plant defense responses during fungal penetration. *The Plant cell* 13: 413–424

Melotto M, Underwood W, Koczan J, Nomura K and He SY (2006) Plant stomata function in innate immunity against bacterial invasion. *Cell* 126: 969–980

Mendez-Vigo B, Rodriguez-Suarez C, Paneda A, Ferreira JJ and Giraldez R (2005) Molecular markers and allelic relationships of anthracnose resistance gene cluster B4 in common bean. *Euphytica* 141: 237–245

Mertely JC, Forcelini BB and Peres NA (2009) Anthracnose fruit rot of strawberry. Plant Pathology Department, UF/IFAS Extension: PP-207. University of Florida, IFAS, EDIS, Gainesville

Michelmore RW, Paran I and Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences* 88: 9828–9832

Moore J, Ditmore M and TeBeest D (2008) Pathotypes of *Colletotrichum sublineolum* in Arkansas. *Plant Disease* 92: 1415–1420

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES and Kresovich S (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences* 110: 453–458

Mysore KS and Ryu C (2004) Nonhost resistance: How much do we know? *Trends in Plant Science* 9: 97–104

Ngugi HK, Julian AM, King SB and Peacocke BJ (2000) Epidemiology of sorghum anthracnose (*Colletotrichum sublineolum*) and leaf blight (*Exserohilum turcicum*) in Kenya. *Plant Pathology* 49: 129–140

Oerke E (2006) Crop losses to pests. *Journal of Agricultural Science* 144: 31–43

Oerke EC, Dehne HW, Schonbeck F and Weber A (1994) Crop production and crop protection. Estimated losses in major food and cash crops. 1st edn. Elsevier, Inc, New York

Patil NY, Klein RR, Williams CL, Collins SD, Knoll JE, Burrell AM, Anderson WF, Rooney WL and Klein PE (2017) Quantitative trait loci associated with anthracnose resistance in sorghum. *Crop Science* 57: 1–14

Perez-Nadales E and Di Pietro A (2011) The membrane mucin Msb2 regulates invasive growth and plant infection in *Fusarium oxysporum. The Plant Cell* 23: 1171–1185

Perumal R, Menz MA, Mehta PJ, Katile S, Gutierrez-Rojas LA, Klein RR, Klein PE, Prom LK, Schlueter JA, Rooney WL and Magill CW (2009) Molecular mapping of *Cg1*, a gene for resistance to anthracnose (*Colletotrichum sublineolum*) in sorghum. *Euphytica* 165: 597–606

Qutob D, Kemmerling B, Brunner F, Kufner I, Engelhardt S, Gust AA, Luberacki B, Seitz HU, Stahl D, Rauhut T, Glawischnig E, Schween G, Lacombe B, Watanabe N, Lam E, Schlichting R, Scheel D, Nau K, Dodt G, Hubert D, Gijzen M and Nurnberger T (2006) Phytotoxicity and innate immune responses induced by Nep1-like proteins. *The Plant Cell* 18: 3721–3744

Ramanujam R and Naqvi NI (2010) PdeH, a high-affinity camp phosphodiesterase, is a key regulator of asexual and pathogenic differentiation in *Magnaporthe oryzae*. *PLoS Pathogens* 6: e1000897

Ravensdale M, Bernoux M, Ve T, Kobe B, Thrall PH, Ellis JG and Dodds PN (2012) Intramolecular interaction influences binding of the flax L5 and L6 resistance proteins to their AvrL567 ligands. *PLoS Pathogens* 8: e1003004

Rogers EE, Glazebrook J and Ausubel F (1996) Mode of action of the *Arabidopsis thaliana* phytoalexin camalexin and its role in *Arabidopsis*-pathogen interactions. *Molecular Plant-Microbe Interactions* 9: 748–757

Ron M and Avni A (2004) The receptor for the fungal elicitor ethylene-inducing xylanase is a member of a resistance-like gene family in tomato. *The Plant Cell* 16: 1604–1615

Rosenow DT and Frederiksen RA (1982) Breeding for disease resistance in sorghum. In: House LR, Mughogho LK and Peacock JM (eds) *Sorghum in the Eighties*. Proc of the International Symposium on Sorghum, Pantancheru, AP, India. 2–7 Nov 1981. ICRISAT, Patancheru, AP, India, 447–455

Rudd JC, Horsley RD, McKendry AL and Elias EM (2001) Host plant resistance genes for Fusarium head blight: Sources, mechanisms, and utility in conventional breeding systems. *Crop Science* 41: 620–627

Savary S, Ficke A, Aubertot J and Hollier C (2012) Crop losses due to diseases and their implications for global food production losses and food security. *Food Security* 4: 519–537

Scholthof KG (2007) The disease triangle: Pathogens, the environment and society. *Nature Reviews Microbiology* 5: 152–156

Schwessinger B and Zipfel C (2008) News from the frontline: Recent insights into PAMP-triggered immunity in plants. *Current Opinion in Plant Biology* 11: 389–395

Shlezinger N, Minz A, Gur Y, Hatam I, Dagdas YF, Talbot NJ and Sharon A (2011) Anti-apoptotic machinery protects the necrotrophic fungus *Botrytis cinerea* from host-induced apoptotic-like cell death during plant infection. *PLoS Pathogens* 7: e1002185

Silva KJP, Brunngs A, Peres NA, Mou Z and Folta KM (2015) The *Arabidopsis NPR1* gene confers broad-spectrum disease resistance in strawberry. *Transgenic Research* 24: 693–704

Singh M, Chaudhary K, Singal HR, Magill CW and Boora KS (2006) Identification and characterization of RAPD and SCAR markers linked to anthracnose resistance gene in sorghum [*Sorghum bicolor* (L.) Moench]. *Euphytica* 149: 179–187

Sitton JW and Cook RJ (1981) Comparative morphology and survival of chlamydospores of *Fusarium roseum* 'Culmorum' and 'Graminearum'. *Phytopathology* 71: 85–90

Smith CW and Frederiksen RA (2000) Sorghum: Origin, History, Technology, and Production. John Wiley & Sons, Inc., New York

Stone JM, Heard JE, Asai T and Ausubel FM (2000) Simulation of fungal-mediated cell death by fumonisin B1 and selection of fumonisin B1-Resistant (*fbr*) *Arabidopsis* mutants. *The Plant Cell* 12: 1811–1822

Suzuki N, Rivero RM, Shulaev V, Blumwald E and Mittler R (2014) Abiotic and biotic stress combinations. *New Phytologist* 203: 32–43

Takai Y, Sasaki T and Matozaki T (2001) Small GTP-binding proteins. *Physiological Review* 81: 153-208

Teng P and Gaunt R (1980) Modelling systems of disease and yield loss in cereals. *Agricultural Systems* 6: 131–154

Thakur RP, Mathur K, Rao VP, Chandra S, Shivramkrishnan S, Kannan S, Hiremath RV, Tailor HC, Kushwaha US, Dwivedi RR and Indira S (1998) Pathogenic and genetic characterization of six Indian populations of *Colletotrichum sublineolum*, the causal agent of sorghum anthracnose. *India Phytopathology Journal* 51: 338–348

Thordal-Christensen H (2003) Fresh insights into processes of nonhost resistance. *Current Opinion in Plant Biology* 6: 351–357

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB and Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genetics* 43: 159–162

Tsuda K, Sato M, Stoddard T, Glazebrook J and Katagiri F (2009) Network properties of robust immunity in plants. *PLoS Genetics* 5: e1000772

Tsuda K and Katagiri F (2010) Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Current Opinion in Plant Biology* 13: 459–465

Tucker SL and Talbot NJ (2001) Surface attachment and pre-penetration stage development by plant pathogenic fungi. *Annual Review of Phytopathology* 39: 385–417

Uppalapati SR, Ishiga Y, Doraiswamy V, Bedair M, Mittal S, Chen J, Nakashima J, Tang Y, Tadege M, Ratet P, Chen R, Schultheiss H and Mysore KS (2012) Loss of abaxial leaf epicuticular wax in *Medicago truncatula irg1*/*palm1* mutants results in reduced spore differentiation of anthracnose and nonhost rust pathogens. *The Plant Cell* 24: 353–370

Wheeler BEJ (1968) Fungal parasites of plants. In: Ainsworth GC and Sussman AS (eds) The fungi: An advanced treatise. Academic Press, New York, chap. 7

Williams SJ, Sohn KH, Wan L, Bernoux M, Sarris PF, Segonzac C, Ve T, Ma Y, Saucet SB, Ericsson DJ, Casey LW, Lonhienne T, Winzor DJ, Zhang X, Coerdt A, Parker JE, Dodds PN, Kobe B and Jones JDG (2014) Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* 344 299–303

Xu J, Urban M, Sweigard JA and Hamer JE (1997) The *CPKA* gene of *Magnaporthe grisea* is essential for appressorial penetration. *Molecular Plant-Microbe Interactions* 10: 187–194

Xu J and Hamer JE (1996) MAP kinase and cAMP signaling regulate infection structure formation and pathogenic growth in the rice blast fungus *Magnaporthe grisea*. *Genes and Development* 10: 2696–2706

Young RA and Kelly JD (1997) RAPD markers linked to three major anthracnose resistance genes in common bean. *Crop Science* 37: 940–946

Young RA, Melotto M, Nodari RO and Kelly JD (1998) Marker-assisted dissection of the oligogenic anthracnose resistance in the common bean cultivar, 'G 2333'. *Theoretical and Applied Genetics* 96: 87–94

Yunis H and Elad Y (1989) Survival of dicarboximide-resistant strains of *Botrytis cinerea* in plant debris during summer in Israel. *Phytoparasitica* 17: 13–21

Zadoks JC (1967) Types of losses caused by plant diseases. In: Chiarappa L (eds), *FAO papers presented at the symposium on crop losses*, 149-158. Rome: FAO

Zadoks JC and Schein RD (1979) *Epidemiology and plant disease management*. New York: Oxford University Press

Zhang J, Song Q, Cregan PB, Nelson RL, Wang X, Wu J and Jiang G (2015) Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics* 16: 217

Zhao J, Williams CC and Last RL (1998) Induction of *Arabidopsis* tryptophan pathway enzymes and camalexin by amino acid starvation, oxidative stress, and an abiotic elicitor. *The Plant Cell* 10: 359–370

Zhu JK (2016) Abiotic stress signaling and responses in plants. *Cell* 167: 313–324

Zipfel C, Robatzek S, Navarro L, Oakeley EJ, Jones JDG, Felix G and Boller T (2004) Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* 428: 764–767

Zou C, Wang P and Xu Y (2016) Bulked sample analysis in genetics, genomics and crop improvement. *Plant Biotechnology Journal* 14: 1941–1955

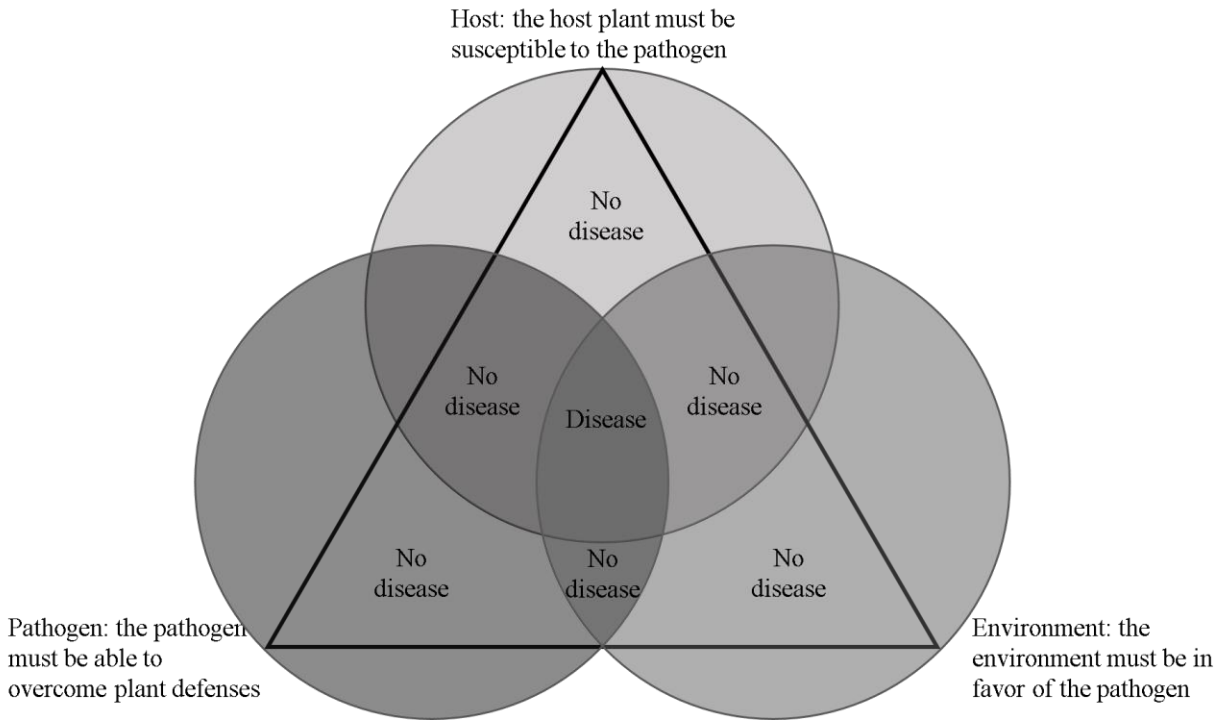


Figure 1.1. The disease triangle that is constituted by host, pathogen and environment. An epidemic will take place only when it meets three requirements: the host plant is susceptible to the pathogen, the pathogen is able to overcome plant defenses, and the environment is in favor of the pathogen.

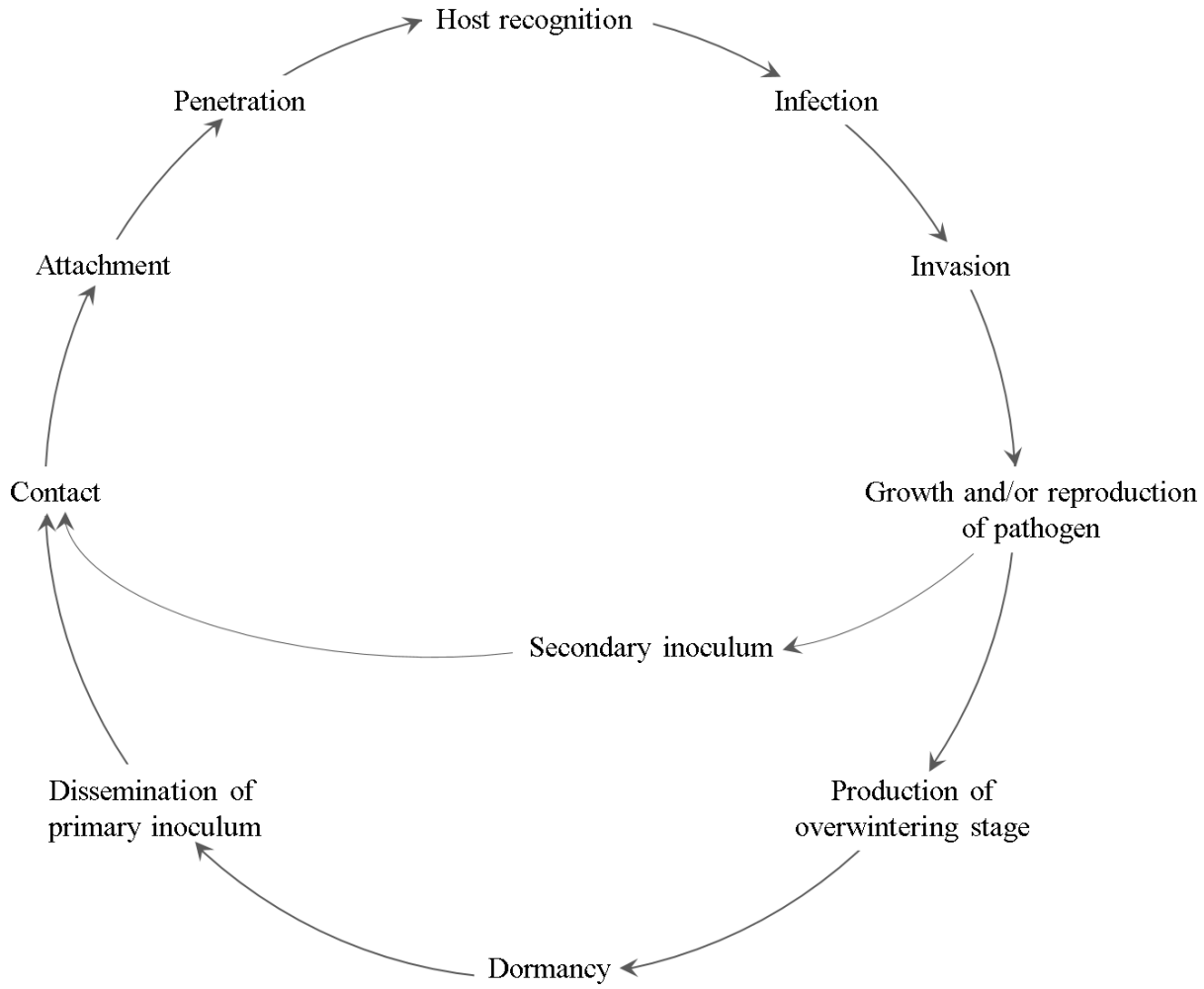


Figure 1.2. Distinct stages in development of a generalized disease cycle.

# CHAPTER 2. INHERITANCE OF SORGHUM ANTHRACNOSE RESISTANCE TO CsGL1

## 2.1 Abstract

Sorghum anthracnose caused by *Colletotrichum sublineolum* is a widely prevalent fungal disease in Africa, Asia and other warm and humid areas. Introduction of host resistance to susceptible cultivars is an effective and recommended method to control the disease. To achieve this goal, we developed a mapping population by a cross of P9830, an inbred line with extreme resistance to strain CsGL1 of the pathogen, and a susceptible line TAM428. The population was phenotyped in the greenhouse, and statistical analysis of this data suggested that the resistance to CsGL1 is controlled by a single gene since the segregation ratio at $F_2$ generation fits Mendel's law of inheritance. Bulked segregant analysis was conducted to locate and map the QTL associated with CsGL1 resistance. The results showed a significant peak (153 SNPs) on the distal end of the long arm of chromosome 8. All 153 SNPs are located in a very narrow range, containing only two candidate genes (*Sobic.008G166400* and *Sobic.008G166550*). *Sobic.008G166400* is the most likely candidate since the majority of SNPs are located within it. Both genes encode leucine-rich repeats (LRR)-containing proteins, which are considered to play an important role in effector recognition and signaling in plant immune response. However, fine-mapping and further exploration in that QTL region is still necessary to confirm the candidate genes or QTL.

## 2.2 Introduction

Sorghum (*Sorghum bicolor* (L.) Moench) is an important grain crop and widely planted in the world. In 2016, total sorghum production in the world was over 57.6 million metric tons, with the African continent being the largest producer at 27.2 million metric tons of produced (FAO 2017). Sorghum was harvested on over 44.7 million hectares of lands in 2016, among which over two

thirds (30.5 million hectares) were in Africa. Sorghum plays a vital role in global food security particularly among the poorest farmers in Nigeria, Ethiopia, Sudan and other developing countries.

Both biotic and abiotic stresses can seriously influence sorghum production. Sorghum anthracnose caused by *Colletotrichum sublineolum* is a fungal disease widely spread in warm and humid areas, especially in West and Central Africa (Marley *et al.*, 2005). The pathogen is able to infect foliage, stalk, panicle and grain, and causes severe reductions in both grain yield and quality (Tesso *et al.*, 2012). Among these different phases of the disease, foliar anthracnose is considered the most common and damaging phase, appearing at 30 to 40 days after seedling emergence (da Costa *et al.*, 2003). Screening techniques for anthracnose were introduced and improved by previous research (Pande *et al.*, 1991; Mehta, 2002; Mehta *et al.*, 2005). Generally, whole sorghum plants are inoculated by spraying with a spore (conidial) suspension of *C. sublineolum* (average spore concentration is usually $10^6$ spore/ml). Inoculating too early can lead to false positives in resistance phenotyping due to unnatural phytoalexin accumulation in seedling leaves (Singh and Boora, 2008).

Several cultural methods managing sorghum anthracnose have been reported. Field sanitation is an effective method to control the disease, as *C. sublineolum* might survive as mycelium, conidia and microsclerotia in crop debris during winter (Marley, 2004). Seed treatment and foliar applied fungicides can also manage anthracnose (Akpa *et al*., 1992), although it is not economical or sustainable for poor farmers in Africa.

Introduction of genetic resistance is considered the most important and effective way to protect sorghum from this devastating disease since it is relatively economical, sustainable and environmental friendly. However, resistance to sorghum anthracnose might be short lived due to the highly variable *C. sublineolum* forms. Breeding for stable and broad host resistance is therefore very difficult. Stacking multiple anthracnose resistance genes into sorghum cultivars can help protect the crop from emerging virulent strains of the pathogen. To accomplish this broad based resistance, resistant sorghum lines and resistance genes or quantitative trait loci (QTL) need to be identified and markers for resistance alleles need to be developed. Towards this end, Coleman and

Stokes (1954) found that resistance to stalk phase (was called stalk red rot at that time) and foliar phase of anthracnose was controlled by two closely linked dominant genes which are 9.57cM apart. Eleven sorghum inbred lines (A 2276-2, IS 3547, IS 8283, IS 9146, IS 9249, IS 18758, SPV 386, PB 8892-2, PS18601-3, PM 20873-1-3 and M 35610) were reported to have stable anthracnose resistance across multiple locations (Burkina Faso, India, Nigeria, Zambia and Zimbabwe) over one to ten years (Pande *et al.*, 1994). A single recessive resistance gene co-segregating with two RAPD (random amplified polymorphic DNA) markers was identified and mapped by bulked segregant analysis (BSA) based on the population whose resistant parent is SC326-6 (Boora *et al*., 1998). Mehta *et al.* (2005) identified at least five unique sources of sorghum anthracnose resistance from 11 germplasm accessions, among which SC748-5 showed the most stable resistance across all five environments and was considered as a valuable and potential resource for future study. A dominant resistance gene *Cg1* was identified and mapped at the end of linkage group LG-05 from SC748-5 by RFLP (restriction fragment length polymorphism) and AFLP (amplified fragment length polymorphism) markers (Perumal *et al.*, 2009). Another recessive gene was mapped on the long arm of chromosome 8 by RAPD and SCAR (sequence characterized amplified regions) markers (Singh *et al.*, 2006).

With the advance of biotechnology, genomic analysis has become a more viable approach to dissecting the genetic basis of sorghum disease resistance and its breeding applications. One of the most commonly used techniques is genome-wide association study (GWAS), through which several QTL and candidate genes have been revealed. Adeyanju *et al.* (2015) reported 14 significant SNPs (single nucleotide polymorphisms) and several candidate genes associated with sorghum stalk rot resistance, according to their GWAS on a sorghum association panel. Single nucleotide polymorphisms tightly linked to genes in the canonical biosynthetic (on chromosome 1) and the catabolic (on chromosome 8) pathways were identified to play important roles in managing dhurrin content in sorghum leaves by GWAS (Hayes *et al.*, 2015). Boyles *et al.* (2016) conducted GWAS to identify loci related with grain yield components, such as grain yield per primary panicle, grain number per primary panicle and 1000-grain weight. Cuevas *et al.* 2018

found resistance loci on chromosome 1 and chromosome 5 condition resistance to sorghum anthracnose through different defense mechanisms.

Bulked segregant analysis combined with next generation sequencing (NGS) is another widely used technique in sorghum research. Han *et al.* (2015) reported a QTL controlling water content in sorghum stem on chromosome 6 in a $F_2$ population. Jiao *et al.* (2018) identified a single gene *Sobic.001G228100* which causes an epi-cuticular wax deficient phenotype in two ethyl methane sulfonate treated mutants.

In this study, we developed a recombinant inbred mapping population to identify inheritance of anthracnose resistance and used BSA combined with NGS to locate the resistance loci.

## 2.3 Materials and Methods

### 2.3.1 Plant materials and population development

A mapping population was developed from a cross between a resistant parent, P9830, and a susceptible parent, TAM428. P9830 is highly resistant to a *Colletotrichum sublineolum* strain CsGL1 (Figure 2.1), which is collected from Kansas. P9830 also shows hypersensitive reaction to two other strains, Georgia and Sc29. TAM428 is highly susceptible to CsGL1 (Figure 2.3) and another four anthracnose strains, CsGL2, Georgia, Cs27 and Sc29. The cross between P9830 and TAM428 was made by hand emasculation in the greenhouse. $F_1$ seed, harvested in 2014, was planted in 7.6cm $\times$ 7.6cm square plastic pots (Hummert, Earth City, MO, USA) in the greenhouse shortly after harvest. $F_2$ seed was harvested from each $F_1$ plant and stored separately in 2015. And then the $F_2$ population was individually planted in pressed peat biodegradable Jiffy pots (Hummert, Earth City, MO, USA) in the greenhouse in May 2015. The $F_2$ seedlings in Jiffy pots were later (early June) transplanted to Purdue University Agronomy Center for Research and Education (ACRE) in West Lafayette, Indiana, after phenotyping for resistance to CsGL1 in the greenhouse prior to field transplanting. The population was advanced by single seed descent to the $F_6$ generation based on the timeline in Figure 2.3. There were a total of 618 lines at the $F_2$ generation

and 500 of these were successfully advanced to the $F_6$ generation as true-breeding recombinant inbred lines (RILs).

### 2.3.2 CsGL1 treatment, phenotyping method and Chi-square test

CsGL1 was grown on half-strength potato dextrose agar (1/2 PDA) (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) with 24 hours/day ambient room light for two to three weeks until the whole 100mm × 15mm Petri dish (Fisher Scientific, Hampton, NH, USA) was covered by spores. A spore suspension of CsGL1 was prepared with 0.1% Tween-20 to a concentration of $1\times10^6$ spore/ml. In the fourth week after planting in the greenhouse, all the seedlings were sprayed with CsGL1 spore suspension in misting chambers with 24hrs/day supplemental light and 70% relative humidity for two days. After two days in misting chambers, seedlings were directly moved onto the misting bench in the greenhouse. Disease severity was scored and recorded when the susceptible check TAM428 fully showed the symptoms, usually five to seven days after moving onto the misting bench. Since P9830 is highly resistant to CsGL1, it shows no symptoms or hypersensitive reaction. Expecting therefore either parental type or an intermediate between them among early generation segregants, only three phenotypic categories were applied in this experiment: highly resistant, moderately resistant (resistant but having several small lesions) and susceptible. The whole population was screened for CsGL1 resistance at $F_2$ and $F_3$ generations, and some selected RILs were screened at $F_6$ generation. Candidate homozygous lines (with respect to CsGL1 resistance or susceptibility), not segregating at the $F_3$ generation, were screened twice more at $F_6$ to confirm their individual phenotypes (and assumed genotypes).

Chi-square ($\chi^2$) test was conducted to check the segregation ratio at F2 generation. Both highly resistant and moderate resistant plants were considered as resistant plants. The null hypothesis of this $\chi^2$ test is that the actual numbers of resistant and susceptible lines at $F_2$ generation match the expected numbers of resistant (three quarters of the total plant number) and susceptible (one quarter of the total plant number) lines, respectively. The significance level of this $\chi^2$ test is 0.05.

### 2.3.3 Extraction of genomic DNA, and construction/sequencing of two DNA bulks

Based on the result of phenotyping, 48 resistant RILs and 48 susceptible RILs showing consistent resistance or susceptibility to CsGL1 in all scored generations, were selected for BSA. Healthy leaf samples were taken from $F_6$ RILs before disease screening and stored in -80°C laboratory freezer (VWR International, Radnor, PA, USA). Genomic DNA was extracted individually from $F_6$ samples by DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). DNA quality was checked on 1% agarose gel (BioExpress, Radnor, PA) and the concentration was measured by Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA). DNA from each line was diluted to 10ng/μl and then equally pooled to build the resistant and susceptible bulks, respectively. Two bulks were sent to Purdue Genomics Core Facility and deep sequenced by Illumina HiSeq 2500 (Illumina, San Diego, CA, USA).

### 2.3.4 Genome assembly and bulked segregant analysis

Sequencing data was analyzed by Purdue Research Computer Clusters. Initial sequencing data was checked for read quality by FastQC (Babraham Institute, Cambridge, UK). The sorghum reference genome used in this study is *Sorghum bicolor* v3.0 downloaded from Phytozome. The genome was aligned to this reference genome by command line, -bwa mem, and ambiguously mapped reads were removed by MarkDuplicates (Picard). Two softwares, Poopulation2 and QTL-Seq (Takagi *et al.*, 2013), were used to analyze allele frequency differences and do Fisher's Exact Test. The Manhattan plot was generated by qqman package in R programming.

## 2.4 Results

### 2.4.1 Development of the mapping population

According to Figure 1a, there were no symptoms observed on P9830's leaves after inoculated with one *Colletotrichum sublineolum* strain CsGL1. When checking the leaf sample at 10× and

20× magnification after inoculation, the plant tissue looked clean with no hyphae observed within plant tissues (Figure 2.1b and 2.1c). This result indicates that CsGL1 is not able to infect P9830, *i.e.*, P9830 is immune to CsGL1. P9830 was also tested with other four *Colletotrichum sublineolum* strains, Georgia, CsGL2, Cs27 and Sc29. P9830 showed hypersensitive reaction to Georgia and Sc29 (Figure 2.2b and 2.2d), but was susceptible to CsGL2 and Cs27 (Figure 2.2a and 2.2c).

TAM428 showed a large area of necrosis with numerous fungal fruiting bodies (acervuli) after inoculating with CsGL1 (Figure 2.3a). For some leaves of TAM428, almost the whole leaf was infected and dead. Not only leaves, but also the stem and sheaths were infected and showed lesions. The disease developed rapidly on TAM428. Under optimal conditions, TAM428 usually showed significant symptoms only three to five days after inoculation. Massive fungal hyphae and acervuli were visible at 10× and 20× magnification (Figure 2.3b and 2.3c). TAM428 was also clearly susceptible to the other four *C. sublineolum* strains, Georgia, CsGL2, Cs27 and Sc29 (Figure 2.4). Although P9830 is also susceptible to CsGL2 and Cs27, TAM428 looks more susceptible as evidenced by larger areas of lesions on its leaves.

### 2.4.2 Phenotyping of the population and inheritance of resistance

In the $F_2$ population, 137 resistant, 323 moderately resistant and 149 susceptible plants were observed (Table 2.2). Combining the resistant and moderately resistant segregants together (460), the ratio of this phenotypic class to the susceptible class (149) fits the expected Mendelian 3:1 segregation pattern ($\chi^2$ equals 0.2566). This result suggests that resistance to CsGL1 is a qualitative trait and controlled by a single gene. The whole $F_3$ population (sum of all progeny, with one plant advanced without intended selection from each $F_2$) was also screened with CsGL1, and the result is listed in Table 2.3. The total number of $F_3$ individuals screened was 580, less than those scored in the $F_2$ (609). When compared the $F_3$ generation's phenotypic data with that of $F_2$ generation's, we found that the number of susceptible lines reduced significantly, which is 30 lower than the $F_2$'s. Reasons for this reduction in susceptible lines in $F_3$ generation could be: (1). some susceptible

lines were highly infected by the pathogen and the necrosis was expanded to almost all areas in all leaves. Those highly susceptible lines died in a short time after inoculation, even before transplanting. (2). some lines died after transplanting probably due to the damage on their roots. Since the susceptible lines had more severe symptoms and damage caused by CsGL1, they were weaker than the resistant lines after inoculation, and therefore had a higher probability to die in the process of transplanting. Because the phenotypic data was recorded before transplanting and every $F_2$ plant would be advanced by single seed descent, $F_2$ lines died after transplanting would still have $F_2$ phenotypic data but could not produce any $F_3$ seed. And the correlation between $F_2$'s and $F_3$'s phenotypic data is 0.8574, which means the relationship between $F_2$'s and $F_3$'s phenotypic data is strong.

156 $F_6$ lines showing most extreme phenotypes (either resistant or susceptible) at both $F_2$ and $F_3$ generations were selected and screened with CsGL1 twice to confirm their phenotypes before BSA. 60 lines showed consistent resistance to CsGL1 all through the $F_2$, $F_3$ and $F_6$ generations, and 48 lines showed consistent susceptibility through all three generations. And therefore, 48 resistant lines were randomly selected to build the resistant bulk and all 48 susceptible lines were selected to build the susceptible bulk for BSA.

### 2.4.3 Bulked segregant analysis

According to the Manhattan plot generated by Poopulation 2 and R programming (Figure 2.5), we can observe three peaks above the threshold: one is on chromosome 2, one on chromosome 6, and one on chromosome 8. Based on the BSA results, there are four SNPs above the threshold located on chromosome 2, one significant SNP located on chromosome 6 and 153 SNPs above threshold on chromosome 8. All the four significant SNPs on chromosome 2 are within a 40kbp region. The SNPs on chromosome 8 cluster are more tightly within a 14kbp region, which is very narrow compared to the whole chromosome (over 62Mbp). Results from QTL-Seq also showed that there is a significant QTL ($\alpha$=0.01) associated with CsGL1 resistance on chromosome 8 (Figure 2.6). However, according to the result of QTL-Seq, there are no significant QTL (either

α=0.01 or α=0.05) observed on chromosome 2 or chromosome 6 (Figure 2.7 and 2.8).

All the SNPs were checked for their specific location based on the sorghum reference genome (*Sorghum bicolor* v3.0). All the significant SNPs on chromosome 2 and chromosome 6 are not located within any candidate genes. However, on chromosome 8, 92 out of 153 significant SNPs are intergenic. And the rest of 61 SNPs are located within only two candidate genes: 58 of SNPs are within *Sobic.008G166400* and three SNPs are within *Sobic.008G166550*. These two candidate genes are located near each other on the distal end of the long arm of chromosome 8. Furthermore, the identity between these two candidate genes is 94.1%, which is considered to be very high.

## 2.5 Discussion

According to the extremely resistant (immune) phenotype showed by P9830, we initially made the hypothesis that according to the gene-for-gene theory (Flor, 1971), to trigger and build successful resistance to anthracnose pathogen CsGL1 in P9830, there is a resistance gene, whose product is able to recognize the product of a specific avirulence gene from the pathogen. And this gene could be identified and mapped by a mapping population. At the same time, we found that a sorghum inbred line, TAM428, is susceptible to all the anthracnose strains (CsGL1, CsGL2, Georgia, Sc29 and Cs27) we have, so we decided to build the mapping population by crossing these two inbred lines. Due to the immunity symptom showed by P9830, we initially thought the disease resistance should be controlled by a single or very few genes. To test this hypothesis, the (P9830×TAM428)$F_2$ population was screened and $\chi^2$ analysis showed that the segregation ratio fit the Mendelian 3:1, implying that CsGL1 resistance is inherited through dominant alleles of a single gene, or very few genes which are tightly linked.

Although it has been demonstrated that BSA combined with NGS (NGS in this dissertation only refers to second-generation sequencing technology, not including third-generation sequencing technology) worked for QTL mapping in sorghum and other crops at the $F_2$ generation (Luan *et al.*, 2012; Takagi *et al.*, 2013; Han *et al.*, 2015), we chose to use RILs because BSA with RILs increases the precision of QTL localization (Magwene *et al.*, 2011). Phenotyping the mapping population in different generations also increases the confidence of the phenotypic data, which is

critical to BSA. Because we were able to cycle through multiple generations annually by using off-season nurseries, advancing a mapping population from $F_2$ generation to $F_6$ generation was not too time-consuming. Furthermore, a recombinant inbred population could be used for other purposes.

In this study, all the significant SNPs located on chromosome 8 were found to be in a 14kbp range, which seems very narrow. Furthermore, when all significant SNPs were checked one by one on chromosome 8, most of them were located in only two candidate genes (the others were not located in genes). Using a recombinant inbred mapping population can lead to more precise and narrow QTL regions, compared to BSA with a $F_2$ population. However, fine-mapping is still needed to confirm the candidate genes. Since the peak on chromosome 8 is very prominent, fine-mapping would be focused on the region around those significant SNPs. In addition, BSA combined with NGS is able to accelerate fine-mapping, because it can effectively provide polymorphic SNPs and DNA markers (Trick *et al.*, 2012; Han *et al.*, 2015). Polymerase chain reaction (PCR) based DNA markers, especially indel (insertion or deletion) markers designed for fine-mapping based on the NGS data are described in the next chapter.

Both candidate genes, *Sobic.008G166400* and *Sobic.008G166550*, encode leucine-rich repeats (LRR)-containing proteins, and the identity between these two genes is 94.1%, which implies that they might have redundant function. LRR-containing proteins play an important role in plant defense responses, especially in effector recognition and signaling (Belkhadir *et al*., 2004; Jones and Dangl, 2006; Bent, 2007). LRR-containing proteins can result in effector-triggered immunity often (but not necessarily) followed by a hypersensitive response (HR), a type of programmed cell death (Dodds and Rathjen, 2010; Katagiri and Tsuda, 2010; Choi *et al.*, 2011). In this study, HR was not observed on P9830 after inoculation. It is possible that some resistance genes confer extreme resistance to diseases without showing HR cell death (Yu *et al*., 1998; Bendahmane *et al*., 1999; Clough *et al.*, 2000), although it is not very common. More details about plant disease resistance mechanisms are discussed in the next chapter after fine-mapping confirmed that candidate genes for CsGL1 resistance are LRR-containing protein encoding genes.

The peak was mapped physically at approximately 60Mbp, and all genes in the region between 59Mbp and 61Mbp were checked whether they encode LRR-containing proteins. There are total of 196 genes in this region, according to the sorghum reference genome (*Sorghum bicolor* v3.1.1) on Phytozome. Among these 196 genes, five of them (*Sobic.008G166400, Sobic.008G166550, Sobic.008G167300, Sobic.008G167500,* and *Sobic.008G172900*) encode LRR-containing proteins, one (*Sobic.008G174966*) encodes proteins similar to LRR-containing protein family, and one (*Sobic.008G175032*) encodes proteins weakly similar to LRR-containing protein family. All of seven genes are possibly involved in disease resistance, and therefore one objective for fine-mapping in the next chapter is to isolate candidate genes for CsGL1 resistance from this region.

BSA combined with NGS is a very effective and efficient tool to detect QTL and candidate genes, especially for qualitative traits and quantitative traits dominated by some genes with major effects. However, it is possible that NGA-assisted BSA might not be able to detect the 'real' candidate genes, due to the technical limitation of NGS. NGS techniques (including Illumina sequencing used in this study) generate millions of short sequence reads, which are usually no longer than 800bp (Liu *et al.*, 2012; Luo *et al.*, 2012; Buermans and den Dunnen, 2014). This requires researchers to align and assemble the short sequence reads to the reference genome, which could be very challenging if the genome is highly repetitive (Treangen and Salzberg, 2012). In addition to this, it is possible that disease resistance genes are absent from the reference genome, if the reference line is susceptible to the disease. In this study, the reference sorghum line BTx623 is susceptible to the anthracnose strain CsGL1. In this case, although some of the significant SNPs on chromosome 8 were located on two candidate genes, there is still a chance that real candidate genes are absent from the reference genome. If 'real' candidate resistance genes are absent from the reference genome and have high identity with those two genes mapped on chromosome 8, it is possible that all sequence reads for 'real' candidate genes were aligned to those two genes, misleading us to wrong candidate genes. Third-generation sequencing (TGS) could be a potential way to solve this problem. In contrast to NGS, TGS generates much longer read length averaging between 5kbp to 15kbp (Bleidorn, 2016; Lee *et al.*, 2016). One of advantages of TGS and long

reads is that it enables and improves genome assembly without the aid of reference genome (Kingsford *et al*., 2010; Koren and Phillippy, 2015). As a result, taking advantage of TGS might be a good method to solve this potential problem. Although TGS technique was not applied in this project, it could be done in the future.

## 2.6 References

Adeyanju A, Little C, Yu J and Tesso T (2015) Genome-wide association study on resistance to stalk rot diseases in grain sorghum. *G3 Genes|Genomes|Genetics* 5: 1165–1175

Akpa AD., Gana SD and Alabi O (1992) Foliar fungicidal spray for the control of some common foliar diseases of sorghum in Nigeria. *Nigerian Journal of Plant Protection* 14: 92-96

Belkhadir Y, Subramaniam R and Dangl JL (2004) Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Current Opinion in Plant Biology* 7: 391–399

Bendahmane A, Kanyuka K and Baulcombe DC (1999) The *Rx* gene from potato controls separate virus resistance and cell death responses. *The Plant Cell* 11: 781–791

Bent AF (2007) Plant disease resistance genes: Function meets structure. *The Plant Cell*, 8: 1757–1771

Bleidorn C (2016) Third generation sequencing: Technology and its potential impact on evolutionary biodiversity research. *Systematics and Biodiversity* 14: 1–8

Boora KS, Frederiksen R and Magill C (1998) DNA-based markers for a recessive gene conferring anthrancnose resistance in sorghum. *Crop Science* 38: 1708–1709

Boyles RE, Cooper EA, Myers MT, Brenton Z, Rauh BL, Morris GP and Kresovich S (2016) Genome-wide association studies of grain yield components in diverse sorghum germplasm. *The Plant Genome* 9: doi: 10.3835/plantgenome2015.09.0091

Buermans HPJ and den Dunnen JT (2014) Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta* 1842: 1932–1941

Choi HW, Kim YJ and Hwang BK (2011) The hypersensitive induced reaction and leucine-rich repeat proteins regulate plant cell death associated with disease and plant immunity. *Molecular Plant-Microbe Interactions* 24: 68–78

Clough SJ, Fengler KA, Yu I, Lippok B, Smith RK and Bent AF (2000) The *Arabidopsis dnd1* "defense, no death" gene encodes a mutated cyclic nucleotide-gated ion channel. *Proceedings of the National Academy of Sciences* 97: 9323–9328

Coleman OH and Stokes IE (1954) The inheritance of resistance to stalk red rot in sorghum. *Agronomy Journal* 46: 61–63

da Costa RV, Casela CR, Zambolim L and Ferreira AS (2003) The sorghum anthracnose. *Fitopatologia Brasileira* 28: 345–354

Cuevas HE, Prom LK, Cooper EA, Knoll JE and Ni X (2018) Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in the U.S. sorghum association panel. *The Plant Genome* 11: doi: 10.3835/plantgenome2017.11.0099

Dodds PN and Rathjen JP (2010) Plant immunity: Towards an integrated view of plant–pathogen interactions. *Nature Reviews Genetics* 11: 539–548

FAO (2017) Food and Agriculture Organization of the United Nations, Statistics Division. http://www.fao.org/faostat/en/#data/QC (Cited 20 February 2019)

Flor HH (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology* 9: 275–296

Han Y, Lv P, Hou S, Li S, Ji G, Ma X, Du R and Liu G (2015) Combining next generation sequencing with bulked segregant analysis to fine map a stem moisture locus in sorghum (*Sorghum bicolor* L. Moench). *PLoS ONE* 10: e0127065

Hayes CM, Burow GB, Brown PJ, Thurber C, Xin Z and Burke JJ (2015) Natural variation in synthesis and catabolism genes influences dhurrin content in sorghum. *The Plant Genome* 8: doi: 10.3835/plantgenome2014.09.0048

Jiao Y, Burow G, Gladman N, Acosta-Martinez V, Chen J, Burke J, Ware D and Xin Z (2018) Efficient identification of causal mutations through sequencing of bulked $F_2$ from two allelic bloomless mutants of *Sorghum bicolor*. *Frontiers in Plant Science* 8: 2267

Jones JDG and Dangl JL (2006) The plant immune system. *Nature* 444: 323–329

Katagiri F and Tsuda K (2010) Understanding the plant immune system. *Molecular Plant-Microbe Interactions* 23: 1531–1536

Kingsford C, Schatz MC and Pop M (2010) Assembly complexity of prokaryotic genomes using short reads. *BMC Bioinformatics* 11: 21

Koren S and Phillippy AM (2015) One chromosome, one contig: Complete microbial genomes from long-read sequencing and assembly. *Current Opinion in Microbiology* 23: 110–120

Lee H, Gurtowski J, Yoo S, Nattestad M, Marcus S, Goodwin S, McCombie WR and Schatz MC (2016) Third-generation sequencing and the future of genomics. *BioRx*volume: doi: http://dx.doi.org/10.1101/048603

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L and Law M (2012) Comparison of next-generation sequencing systems. *Journal of Biomedicine and Biotechnology* 2012: 251364

Luan J, Wang F, Li Y, Zhang B and Zhang J (2012) Mapping quantitative trait loci conferring resistance to rice black-streaked virus in maize (*Zea mays* L.). *Theoretical and Applied Genetics* 125: 781–791

Luo C, Tsementzi D, Kyrpides N, Read T, Konstantinidis KT (2012) Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *PLoS ONE* 7: e30087

Magwene PM, Willis JH and Kelly JK (2011) The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology* 7: e1002255

Marley PS (2004) Effects of integrating host plant resistance with time of planting or fungicides on anthracnose and grain mould and yield of sorghum (*Sorghum bicolor*) in the Nigerian northern Guinea Savanna. *Journal of Agricultural Science* 142: 345–350

Marley PS, Diourte M, Neya A and Rattunde FW (2005) Sorghum anthracnose and sustainable management strategies in West and Central Africa. *Journal of Sustainable Agriculture* 25: 43–56

Mehta PJ (2002) Genetic anthracnose resistance genes in sorghum: characterization and molecular markers. PhD dissertation, Texas A&M University, 91 pp

Mehta PJ, Wiltse CC, Rooney WL, Collins SD, Frederiksen RA, Hess DE, Chisi M and TeBeest DO (2005) Classification and inheritance of genetic resistance to anthracnose in sorghum. *Field Crops Research* 93: 1–9

Pande S, Mughogho LK, Bandyopadhyay R and Karunakar RI (1991) Variation in pathogenicity and cultural characteristics of sorghum isolates of *Colletotrichum graminicola* in India. *Plant Disease* 75: 778-783

Pande S, Thakur RP, Bandyopadhyay R and Reddy BVS (1994) Development of screening methods and identification of stable resistance to anthracnose in sorghum. *Field Crops Research* 38: 157–166

Perumal R, Menz MA, Metha PJ, Katile S, Gutierrez-Rojas LA, Klein RR, Prom LK, Schlueter JA, Rooney WL and Magill CW (2009) Molecular mapping of *Cg1*, a gene for resistance to anthracnose (*Colletotrichum sublineolum*) in sorghum. *Euphytica* 165: 597–606

Singh M and Boora KS (2008) Molecular characterization of anthracnose resistance gene in sorghum. In: Rao GP, Wagner C, Singh RK and Sharma ML (eds) *Plant Genomics and Bioinformatics*. Stadium Press LLC, Houston, Texas, USA, 373-388

Singh M, Chaudhary K, Singal HR, Magill CW and Boora KS (2006) Identification and characterization of RAPD and SCAR markers linked to anthracnose resistance gene in sorghum [*Sorghum bicolor* (L.) Moench]. *Euphytica* 149: 179–187

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S, Innan H, Cano LM, Kamoun S and Terauchi R (2013) QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *The Plant Journal* 74: 174–183

Tesso T, Perumal R, Little CR, Adeyanju A, Radwan GL, Prom LK and Magill CW (2012) Sorghum pathology and biotechnology - A fungal disease perspective : Part II . Anthracnose , stalk rot and downy mildew. *The European Journal of Plant Science and Biotechnology* 6: 31–44

Treangen TJ and Salzberg SL (2012) Repetitive DNA and next-generation sequencing: Computational challenges and solutions. *Nature Reviews Genetics* 13: 36–46

Trick M, Adamski NM, Mugford SG, Jiang C, Febrer M and Uauy C (2012) Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biology* 12: 14

Yu I, Parker J and Bent AF (1998) Gene-for-gene disease resistance without the hypersensitive response in *Arabidopsis dnd1* mutant. *Proceedings of the National Academy of Sciences* 95: 7819–7824

Table 2.1. Timeline of developing the mapping population crossed by P9830 and BTx623 during year 2014 to 2016.

| Generation of the population | Timeline |
|---|---|
| Parents (P9830 and TAM428) and $F_1$ | Phenotyped by Dr. Fuyou Fu in the greenhouse at Lilly Hall in 2014 |
| $F_2$ population | Screened in the greenhouse at Lilly Hall in June 2015 and transplanted to ACRE |
| $F_3$ seed | $F_3$ seeds harvested in ACRE in October 2015 |
| $F_4$ seed | Harvested in Puerto Rico in March 2016 |
| $F_5$ seed | Harvested in Puerto Rico in June 2016 |
| $F_6$ seed | Harvested in ACRE in October 2016 |

Table 2.2. Score for CsGL1 resistance in the $F_2$ generation. Phenotypes are indicated as highly resistant (R), moderately resistant (M), and susceptible (S). The *p*-value of Chi-square test for total screened $F_2$ plants is 0.2566.

| $F_2$ subpopulations* | R | M | S | Total number of plants |
|---|---|---|---|---|
| 1 | 3 | 29 | 13 | 45 |
| 2 | 9 | 26 | 11 | 46 |
| 5 | 5 | 37 | 4 | 46 |
| 8 | 5 | 8 | 5 | 18 |
| 9 | 5 | 30 | 11 | 46 |
| 10 | 28 | 8 | 10 | 46 |
| 12 | 13 | 19 | 13 | 45 |
| 13 | 5 | 29 | 11 | 45 |
| 14 | 13 | 18 | 14 | 45 |
| 16 | 5 | 26 | 15 | 46 |
| 17 | 18 | 18 | 9 | 45 |
| 18 | 15 | 16 | 14 | 45 |
| 19 | 10 | 28 | 8 | 46 |
| 20 | 3 | 31 | 11 | 45 |
| Total | 137 | 323 | 149 | 609 |

* Each $F_2$ subpopulation was developed from a single $F_1$ plant. $F_2$ seeds were harvested from single $F_1$ plant, and stored separately. Every $F_2$ seed from the same $F_1$ plant became a single line in the $F_2$ subpopulation.

Table 2.3. Phenotypic data from $F_3$ lineages. Phenotypes indicate the number of screened plants in each class, resistant (R), segregating, and susceptible (S). Chi-square test was not done for the $F_3$ generation since the number of susceptible lines reduced significantly from $F_2$ to $F_3$.

| $F_3$ subpopulation | R | Segregating | S |
|---|---|---|---|
| 1 | 3 | 25 | 14 |
| 2 | 8 | 27 | 8 |
| 5 | 4 | 35 | 4 |
| 8 | 7 | 7 | 1 |
| 9 | 12 | 22 | 11 |
| 10 | 25 | 9 | 9 |
| 12 | 14 | 18 | 11 |
| 13 | 3 | 29 | 9 |
| 14 | 14 | 18 | 12 |
| 16 | 6 | 26 | 13 |
| 17 | 23 | 15 | 7 |
| 18 | 16 | 23 | 4 |
| 19 | 16 | 24 | 6 |
| 20 | 4 | 28 | 10 |
| Total | 155 | 306 | 119 |


Figure 2.1. Phenotype of P9830 after inoculated with CsGL1. (a) Overall phenotype observed by the naked eye; (b) Leaf sample under a microscope at 10× magnification; (c) Leaf sample under a microscope at 20× magnification.

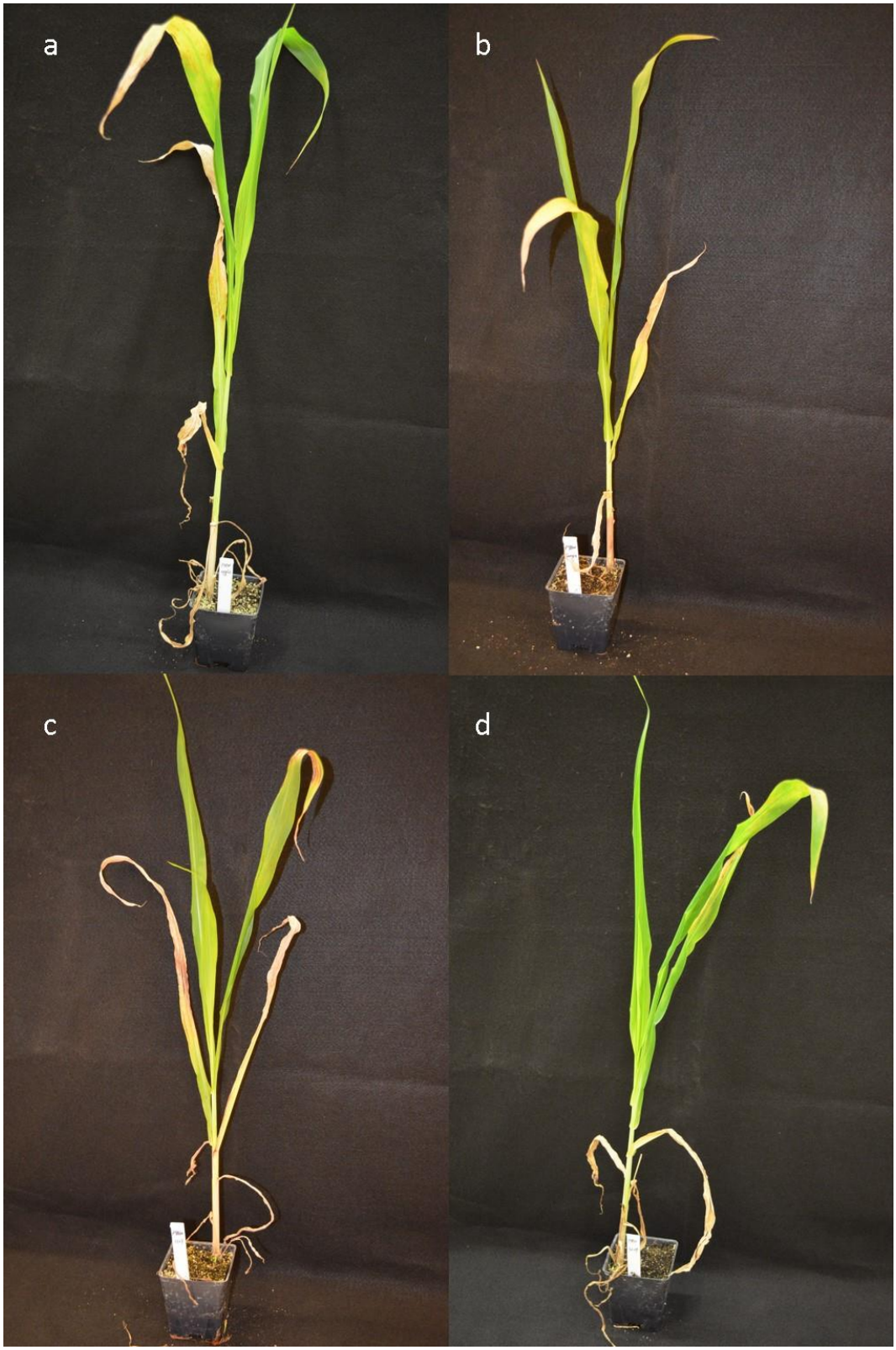


Figure 2.2. Phenotypes of P9830 after inoculating with (a) CsGL2; (b) Georgia; (c) Cs27; (d) Sc29.

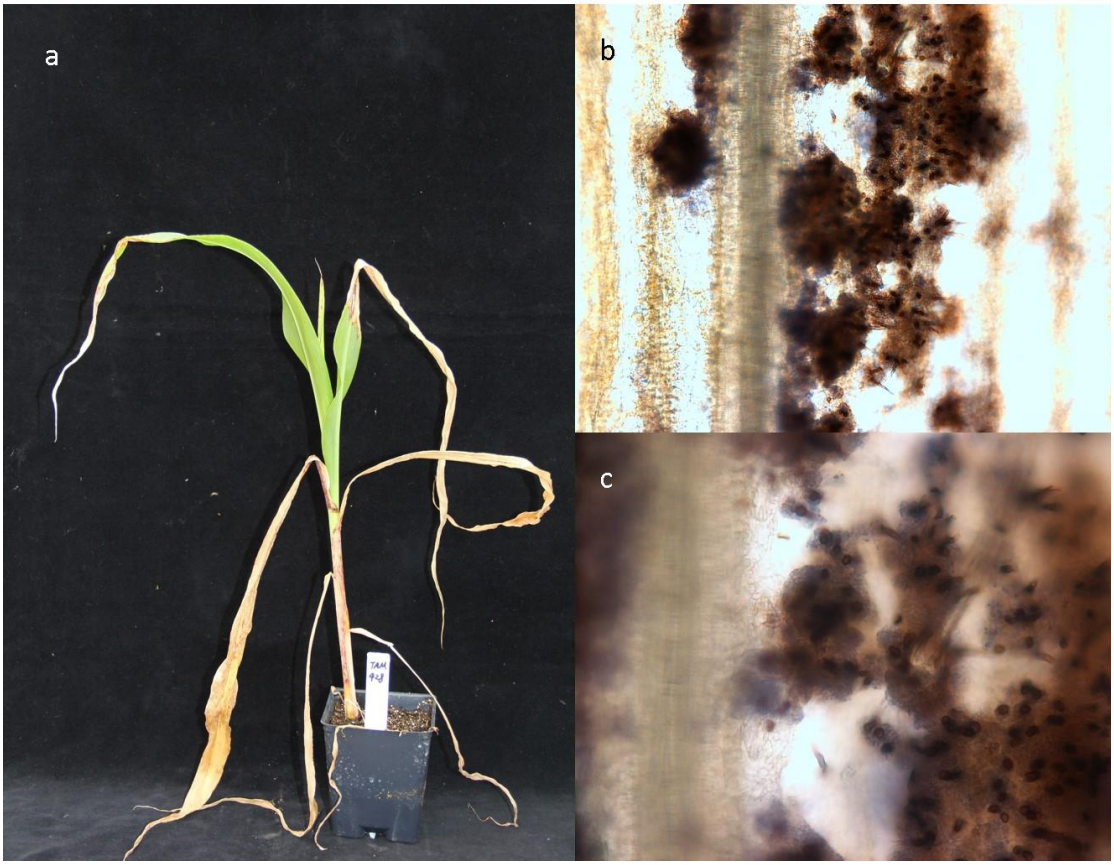


Figure 2.3. Phenotype of TAM428 after inoculating with CsGL1. (a) Overall phenotype observed by the naked eye; (b) Leaf sample under a microscope at 10× magnification; (c) Leaf sample under a microscope at 20× magnification.


Figure 2.4. Phenotypes of TAM428 after inoculated with (a) CsGL2; (b) Georgia; (c) Cs27; (d) Sc29.

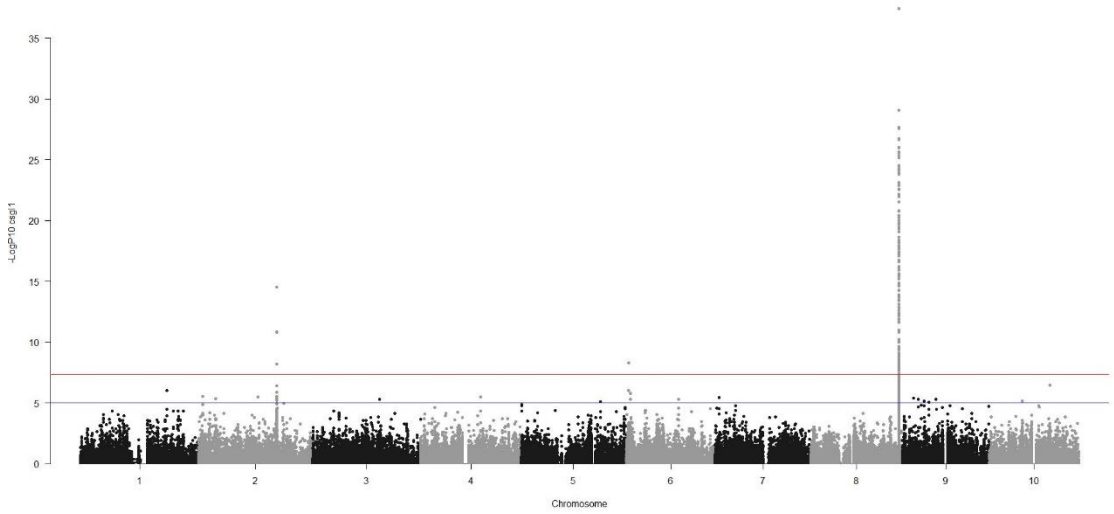


Figure 2.5. Manhattan plot of the BSA results. There are three peaks above the threshold (upper) line ($\alpha$=0.01). The most significant peak occurs on chromosome 8.

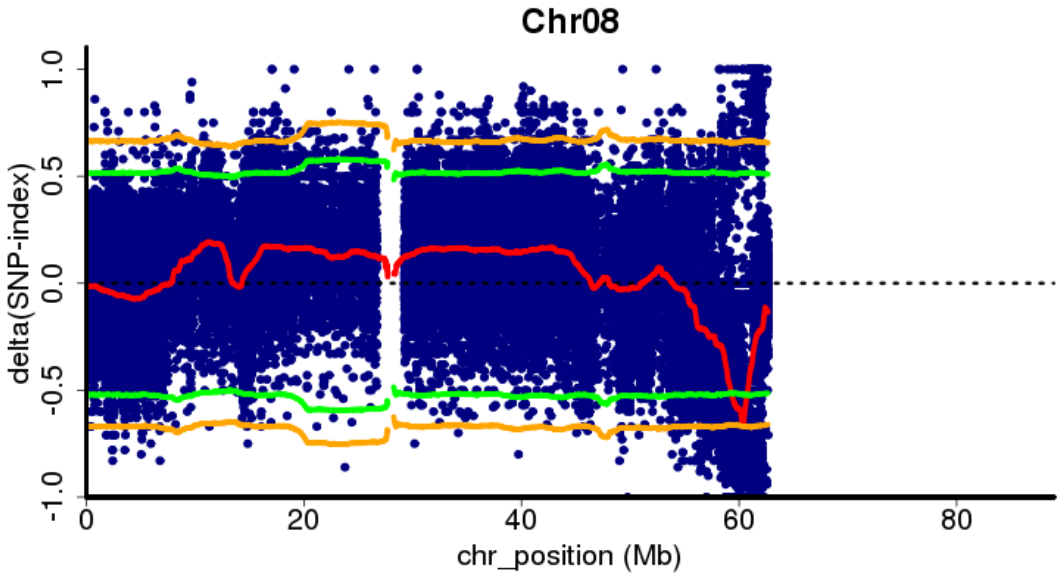


Figure 2.6. Definition of SNP-index and Δ(SNP-index) was interpreted by Takagi et al. (2013). The orange and green lines equal the criteria with α=0.01 and α=0.05, respectively. A very significant curve hits 99% confidence of a QTL associated with CsGL1 at approximately 60Mbp on chromosome 8.

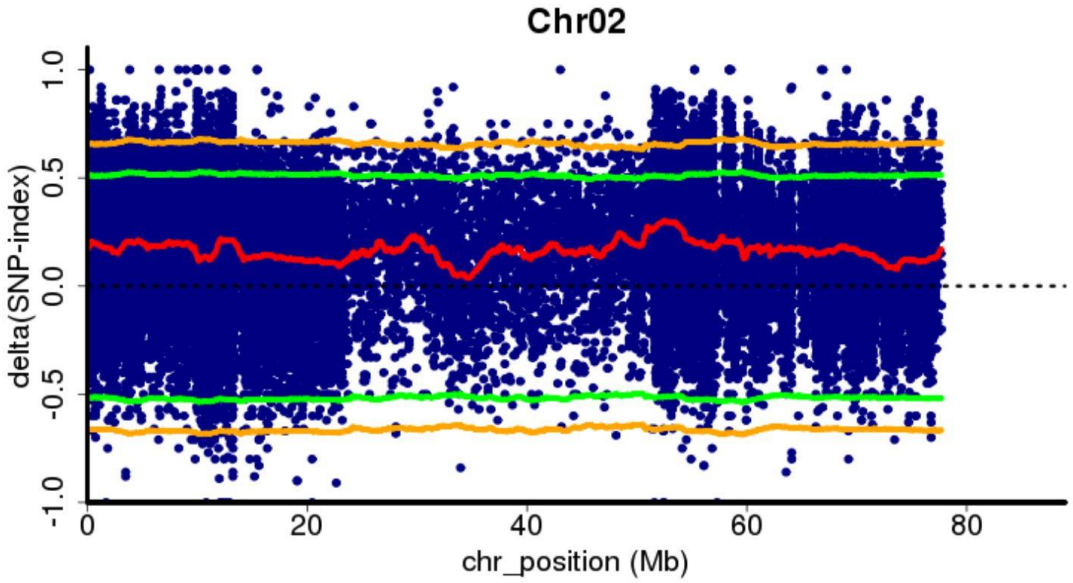


Figure 2.7. Results generated by QTL-seq. Although there is a peak on chromosome 2 on the Manhattan plot, no significant loci were observed based on QTL-seq.

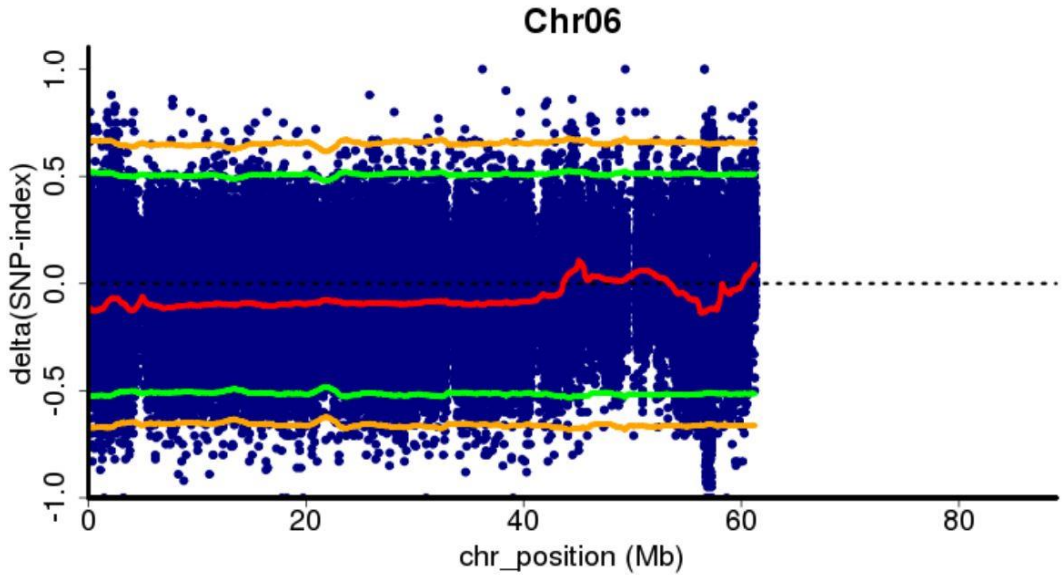


Figure 2.8. Results generated by QTL-seq. Although there is a peak on chromosome 6 on the Manhattan plot, no significant loci were observed based on QTL-seq.

# CHAPTER 3. FINE MAPPING AND FURTHER EXPLORATION OF THE CANDIDATE RESISTANCE LOCUS

## 3.1 Abstract

The candidate CsGL1 resistance locus was mapped on the distal end of the long arm of chromosome 8 by BSA in Chapter 2. In this chapter, we used the sequencing data to design primers, fine mapped the candidate locus, and did qPCR analysis for the candidate gene *Sobic.008G166400*. Based on indel and SNP markers, we constructed a linkage map between positions 59,451,828 and 60,473,671 on chromosome 8, found recombination within this region, and found that the SNP marker *08.60001890_TG* (inside the candidate gene *Sobic.008G166400*) co-segregates with the resistance. The SNP marker *08.60001890_TG*, together with another two indel markers (*Del_6014* and *In_5986*) are closely linked to the resistance allele and can be used for marker-assisted selection in the future. According to the qPCR analysis, the expression level of *Sobic.008G166400* in a resistance RIL 17-12 remained unchanged after *C. sublineolum* inoculation (among 0, 12, 24, 36 and 48 hpi). The expression of *Sobic.008G166400*, however, was induced in the susceptible RIL 13-31 at 12, 24 and 36 hpi after infection. We concluded that *Sobic.008G166400* might be the first candidate resistance gene against CsGL1, and *Sobic.008G166550* might be a duplicated gene of *Sobic.008G166400* and another candidate gene.

## 3.2 Introduction

Anthracnose caused by fungal agents *Colletotrichum* spp. is a common disease in sorghum, maize, sugarcane, strawberry, banana, mango and many other plant species (Cannon *et al.*, 2012). Some researchers consider it as one of the top ten fungal plant pathogens based on its significant scientific/economic importance (Dean *et al.*, 2012). Research on *Colletotrichum* spp. and their hosts has been widely conducted.

One important resistance gene (R gene) against *C. sublineolum* in sorghum, *Cg1*, is a dominant gene at the distal end of chromosome 5 conferring stable resistance to a *C. sublineolum* isolate 430BB-85, which is common in Texas (Perumal *et al.*, 2009). Two PCR-based markers, *Xtxp549* and *Xtxa6227*, were identified to be tightly linked to *Cg1* and considered useful for marker-assisted selection. The physical position of *Cg1*, however, still seems unclear and the resistance gene has not been cloned yet. Burrell *et al.* (2015) reported several candidate resistance genes and five gene families associated with disease resistance in the region between 59.97 to 60.77Mbp on chromosome 5. Among these genes, two of them are NB-LRR genes (Burrell *et al.*, 2015), which are suspected to play an important role in effector recognition and signaling in plant defense (Jones and Dangl, 2006). Based on some GWAS results, three candidate genes (*Sobic.005G172300*, *Sobic.005G182400* and *Sobic.005G228400*) on chromosome 5 were also reported by Cuevas *et al.* (2018), involving a F-box domain, NB-LRR domain and Oryzalide A biosynthesis. On chromosome 9, two NB-LRR genes, *Cs1A* (*Sb09g027470*) and *Cs2A* (*Sb09g004240*), and their tightly linked duplicated genes, *Cs1B* (*Sb09g027520*) and *Cs2B* (*Sb09g004210*), were reported to confer resistance to a *C. sublineolum* isolate from Uganda (Biruma *et al.*, 2012). According to a *Brome mosaic virus*-based virus-induced gene silencing (VIGS) system and real-time quantitative PCR (qPCR) analysis, silencing of either *Cs1A* or *Cs2A* led to susceptible phenotypes of BS04/05, which was resistant to the anthracnose isolate. Down-regulation of other related genes encoding a lipid transfer protein, a zinc finger-like transcription factor and a cell death related protein also caused a crash of resistance (without hypersensitive response (HR) to the pathogen) to the anthracnose pathogen (Biruma *et al.*, 2012). Patil *et al.* (2017) also reported 27 candidate genes related to plant disease resistance on chromosome 9, involving programmed cell death, protein ubiquitination, oxidative stress response, NB-ARC (the core nucleotide-binding fold in NB-LRR proteins) and other immunity response related functions. Although several candidate resistance genes against anthracnose pathogens were revealed by many researchers, functions of those genes seldom have been deeply studied in sorghum. Expression levels of those candidate genes in

different sorghum organs or at different stages are also not known. Studies on these aspects is a potential direction for future anthracnose resistance researches in sorghum.

Anthracnose crown rot caused by *Colletotrichum acutatum* (major causal agent), *C. fragariae* and *C. gloeosporioides* (less frequent) is a common destructive fungal disease leading to significant yield loss in strawberry (Mertely *et al*., 2009), and therefore could be a reference for anthracnose resistance studies in sorghum. On this pathogen, Amil-Ruiz *et al.* (2016) reported that accumulation of two phytohormones, salicylic acid (SA) and jasmonic acid (JA) was detected after *C. acutatum* infection, without induction of some important SA and JA pathways related plant defense genes (e.g. *FaPR1-1*, *FaLOX2*, *FaJAR1*, *FaPDF1*, and *FaGST1*). Their results indicated that the pathogen uses an unclear strategy to overcome the strawberry immunity. Encinas-Villarejo *et al.* (2009) first isolated a *WRKY* gene *FaWRKY1* in strawberry and reported it has an important role in regulating resistance to *C. acutatum* infection. Differential expression of *FaWRKY1* in different tissues of the same strawberry variety was also observed, which is in agreement with the point that the resistance genes against *C. acutatum* could be genotype and tissue dependent (Casado-Diaz *et al.*, 2006). They also made a comparison of the function of *WRKY1* gene between strawberry and *Arabidopsis*, and estimated that this resistance gene might involve different mechanisms of defense responses in both plant species. Higuera *et al.* (2019) found the strawberry *FaWRKY1* transcription factor acts as a negative regulator in *C. acutatum* resistance in strawberry, which contrasts with the positive role of *AtWRKY1* in response to the bacterial pathogen *Pseudomonas syringae*.

In this chapter, we undertook fine mapping for the candidate resistance locus to CsGL1 detected by BSA in Chapter 2. And we also checked the expression level of the candidate gene on chromosome 8.

## 3.3 Materials and Methods

### 3.3.1 PCR-based primers design

Insertion-deletion (indel) markers were designed based on the genomic sequence of P9830 (resistance parent), TAM428 (susceptible parent), resistant and susceptible DNA bulks within 1.2Mbp range (59.4Mbp to 60.6Mbp) for fine mapping. A total of 20 indels over 10bp were designed by Primer3 (Koressaar and Remm, 2007; Untergasser *et al*., 2012) and synthesized by IDT (Integrated DNA Technologies, Coralville, IA, USA). All the indels used for fine mapping are listed in Table 3.1.

Single-nucleotide polymorphism (SNP) markers were also designed (Tanha *et al*., 2015) based on the genomic sequence of P9830 (resistance parent), TAM428 (susceptible parent), resistant and susceptible DNA bulks for co-segregation analysis and fine mapping. Briefly, two pairs of primers (one pair of inner and one pair of outer primers) with same annealing temperature for each SNP were designed by Primer1.The PCR products of inner primers for different nucleotides at the same position on chromosome would show at least 20 bp polymorphism on agarose gel electrophoresis. A total of 18 SNPs (as tetra primers) were designed by Primer3 and synthesized by IDT. All the SNPs used in this experiment are listed in Table 3.2.

### 3.3.2 Co-segregation analysis and construction of linkage map

A total of 360 recombinant inbred lines of the mapping population (P9830 × TAM428) at the $F_6$ generation were randomly selected to do co-segregation analysis and fine mapping. These lines were planted and phenotyped with CsGL1 in the same way as described in Chapter 2 in the greenhouse in Lilly Hall, West Lafayette, Indiana. Healthy leaf samples (with diameter approximately 5mm) were taken from $F_6$ lines by a hand-held paper punch before disease screening and stored in 96-well PCR plate (Life Science Products, Chestertown, MD, USA). Genomic DNA for PCR was extracted by a high-throughput extraction method reported by Xin *et al.* (2003). PCR-based SNP markers were used to screen these selected lines for co-segregation analysis. Indel

markers were used to genotype 360 recombinant inbred lines of the mapping population (P9830 × TAM428) at the $F_6$ generation. A linkage map was constructed based on the phenotypic and genotypic data by JoinMap v4.0 (Kyazma, Wageningen, Netherlands).

### 3.3.3 DNA preparation, wide-range PCR and wide-seq for the candidate genes

The resistant parent P9830 and the susceptible parent TAM428 were planted in the greenhouse and leaf samples were taken at the end of the third week. Genomic DNA was extracted from P9830 and TAM428 leaf tissues with the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Primers to amplify the two candidate genes, *Sobic.008G166400* and *Sobic.008G166550,* were designed by Primer3. Several different pairs of primers and their combinations were tested, and primers that successfully amplified these two genes are listed in Table 3.3. The Phusion Green Hot Start II High-Fidelity PCR kit (Thermo Fisher Scientific, Waltham, MA, USA) was used for wide-range PCR. The quality of wide-range PCR products was checked on a 1% agarose gel (BioExpress, Radnor, PA) and the concentration was measured by Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA). After quality check, PCR products were sent to Purdue Genomics Core Facility for WideSeq (next generation sequencing) analysis.

WideSeq (deep sequence) data was analyzed and assembled by Purdue Research Computer Clusters in the same way as described in Chapter 2. The sorghum reference genome used here is *Sorghum bicolor* v3.0 downloaded from Phytozome. The genome was aligned to the reference genome by the command line, -bwa mem, and ambiguously mapped reads were removed by MarkDuplicates (Picard).

### 3.3.4 RNA preparation and cDNA library construction

To measure the expression level of the candidate genes, a resistance RIL 17-12 (having the same highly resistant phenotype as P9830) and a susceptible RIL 13-31 (having a highly susceptible phenotype like TAM428) were planted separately in 7.6cm × 7.6cm square plastic pots (Hummert, Earth City, MO, USA) in the greenhouse. Both 17-12 and 13-31 were divided into

check and treatment groups, respectively. Seedlings in the treatment group were inoculated with spore suspension of CsGL1 (a concentration of $1\times10^6$-$3\times10^6$ spore/ml with 0.1% Tween-20) in the fourth week after planting. Seedlings in the check group were inoculated with double-distilled water (with 0.1% Tween-20) at the same time when the treatment group was inoculated. Both groups were placed in misting chambers with 24 hours/day light and 70% relative humidity for two days after sprayed with inoculum. After two days in misting chambers, all seedlings were randomly placed onto the misting bench in the greenhouse. Leaf tissue samples were collected at 0, 12, 24, 36 and 48 hours post inoculation (hpi) from 17-12 and 13-31 in both groups with three biological replications. Leaf samples were stored in -80°C laboratory freezer until RNA was extracted.

Total RNA in leaf samples was isolated using TRI Reagent (Molecular Research Center, Cincinnati, OH, USA). The concentration and quality of RNA were analyzed and checked by Nanodrop 2000 spectrophotometer and 1% agarose gel. RNA with good quality and quantity was treated with DNase I (New England Biolabs, Ipswich, MA, USA) and checked for concentration and quality again. DNA-free RNA equivalent to 2μg RNA/reaction was used for SuperScript IV Reverse Transcriptase cDNA synthesis system (Invitrogen, Carlsbad, CA, USA) based on the manufacturer's instructions.

### 3.3.5 Real-time quantitative PCR

For candidate gene expression analysis, real-time quantitative PCR (qPCR) was performed using CFX Connect Real-Time System (Bio-Rad Laboratories, Hercules, CA, USA) with iTaq Universal SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA, USA) according to the manufacturer's recommendations. Primers for the amplification of gene transcripts were designed by Primer3 and listed in Table 3.4. The relative values for candidate genes were normalized using *SbActin* (*Sobic. 3001G112600*).

## 3.4 Results

### 3.4.1 Co-segregation analysis and fine mapping

Based on the genotypic data of SNP markers and phenotypic data, we found that two SNPs (*08.60001890_TG* and *08.60002109_GC*) co-segregated with the resistance phenotype. The SNP *08.60001890_TG* is located at the position 60,001,890 and *08.60002109_GC* is located at the position 60,002,109 on chromosome 8. Resistant lines showed T and G bases at 60,001,890 and 60,002,109 on chromosome 8 while susceptible lines (and reference genome) showed G and C at those two positions, respectively. Both of SNPs locate to the 5' untranslated region (5' UTR) of one candidate gene *Sobic.008G166400* and therefore do not directly cause any structure change in the predicted protein product.

Using seven deletion markers, three insertion markers and one SNP marker, we built a linkage map between the region 59,451,828 and 60,473,671 on chromosome 8 (Figure 3.1). The SNP marker *08.60001890_TG* (inside the candidate gene *Sobic.008G166400*) co-segregates with the resistance trait and two indel markers are closely linked to it. A 30bp deletion at the position of 60,147,358 on chromosome 8 was mapped within approximately 1.5 cM of the anthracnose resistance locus, and a 22bp insertion at the position of 59,860,139 was mapped within approximately 1.2 cM.

### 3.4.2 Expression level of the candidate gene

The phenotypes of the resistant RIL 17-12 (same phenotype with P9830) and the susceptible RIL 13-31 (same phenotype with TAM428) at 0, 24 and 48 hpi are shown in Figure 3.2. Under optimal conditions, necrotic lesions started showing up on leaves of the susceptible line at 24 hpi (Figure 3.2b). At 48 hpi, lesions on the second leaf of 13-31 were very evident, as on TAM428. For 17-12 and P9830, there were no disease symptoms or necrotic lesions observed on leaves.

The relative expression levels of *Sobic.008G166400* were similar at 0 hpi in both 17-12 and 13-31 (Figure 3.3) whether they were inoculated by the spore suspension medium or double-

distilled water. Expression differences were assessed over a period of 48 hours. At 12 hpi, the relative expression level in the 13-31 plant inoculated by the pathogen appeared to be induced while the others remained unchanged or only slightly changed (not significant at $\alpha$ =0.05). At 24 hpi, the relative expression level in the pathogen infected 13-31 plant is approximately twice of the others (significant at $\alpha$ =0.05), which remained close to the 0 hpi expression levels. At 36 hpi, the relative expression of *Sobic.008G166400* in the treated 13-31 plant was reduced compared with that of 24 hpi and was not significant at $\alpha$ =0.05 compared with the other three sets. At 48 hpi, the relative expression of *Sobic.008G166400* in the treated 13-31 plant was still higher than the rest, but not as significantly as levels measured at 24 hpi.

## 3.5 Discussion

According to the linkage map generated by indel and SNP markers, we found recombination in the region between 59,451,828 and 60,473,671 on chromosome 8. This result increased the confidence that the candidate locus with two candidate genes (especially for *Sobic.008G166400*) is related with the resistance to CsGL1 in P9830. When we checked the genomic and amino acid sequences of these two candidate genes, we found that these two genes are possibly duplicated genes with similar functions. Both *Sobic.008G166400* and *Sobic.008G166550* encode LRR proteins. The percent sequence identity of genomic sequences between these two genes is 94.1% based on Kalign (The European Bioinformatics Institute, Hinxton, UK) and the identity of peptide sequences is 73.05% based on NCBI (National Center for Biotechnology Information) BLAST (basic local alignment search tool), which is considered to be nearly identical. The size of *Sobic.008G166400* based on reference genome (*Sorghum bicolor* v3.1.1) on Phytozome is 6892bp, while the size of *Sobic.008G166550* is 3658bp. In this case, *Sobic.008G166550* might encode a truncated protein with similar function of the protein encoded by *Sobic.008G166400*. *Sobic.008G166400* contains only one big exon (4668 bp) according to Phytozome, while *Sobic.008G166550* has three (total of 3468 bp). According to the BLAST result on NCBI, we found that *Sobic.008G166400* and *Sobic.008G166550* have the closest relationship. There is

another LRR gene *Sobic.008G167500* on the distal end of the long arm of chromosome 8 (approximately 150Kbp from *Sobic.008G166400*) that may be an ortholog of *Sobic.008G166400*. *Sobic.006G223101* encoding a "weakly similar to NBS-LRR disease resistance protein" homologue could be another potential homolog of *Sobic.008G166400*, based on the protein sequence BLAST result on NCBI. The identity of the protein sequence between *Sobic.008G166400* and *Sobic.006G223101*, however, is only 47.03%, which is significantly lower than that between *Sobic.008G166400* and *Sobic.008G167500* (73.05%). The closest paralogs in other plant species is in *Setaria* spp., according to the NCBI BLAST results. *Sevir.3G403200* in *Setaria viridis* encodes a LRR protein with over 67% identity to *Sobic.008G166400*. *Sevir.3G414200* in *Setaria viridis* and *Seita.3G396100* in *Setaria italica* (foxtail millet) are another two potential candidate orthologs of *Sobic.008G166400* with over 63% identity. For *Sobic.008G166400* (also called *Sb08g021230* in some papers and databases), it might also be a homolog copy of a wheat rust resistance gene *Lr1* (Upadhyaya *et al.*, 2013).

Although we designed several primers in the 3' UTR and coding regions in *Sobic.008G166550*, the pre-test showed that all of them were unable to amplify efficiently. Primers in the coding region of *Sobic.008G166400* were also ineffective for qPCR. Since we had a limited quantity of RNA, we just used the primers in 3' UTR in *Sobic.008G166400* to do the qPCR analysis. Based on the qPCR result, we found that the relative expression levels of *Sobic.008G166400* almost stayed unchanged in 17-12 at 0, 12, 24, 36 and 48 hpi in both the control and treatment groups. This is not surprising for a resistance gene. Resistance genes can be up-regulated, down-regulated or unchanged after infection. For example, Casado-Diaz *et al.* (2006) reported that the expression level of a strawberry gene *D111ACI01* (encoding a hypersensitive induced reaction protein) didn't show a significant change after *C. actatum* inoculation. Furthermore, Guidarelli *et al.* (2011) and Fang *et al.* (2012) also reported several disease response related genes in strawberry were either up-regulated or down-regulated after the infection with a *Colletotrichum* spp. Based on data on the Morokoshi sorghum transcriptome database (no data was found for *Sobic.008G166550*), the expression level of *Sobic.008G166400*, however, was induced (though not significantly) at 12 hpi

by *Bipolaris sorghicola*, the causal agent of fungal disease target leaf spot (Yazawa *et al.*, 2013). This might indicate that *Sobic.008G166400* might be involved in different disease response mechanisms to different sorghum diseases. To validate this assumption, however, further experiments may need to be conducted. The expression of *Sobic.008G166400* in the susceptible RIL 13-31 seemed to be strictly induced by the pathogen. Similar observations have been made in strawberry studies. Higuera *et al.* (2019) observed significant induction of *FaWRKY1*, a member of WRKY family which is considered to play an important role in plant defense to biotic and abiotic stresses, in susceptible strawberry fruits after the inoculation of *C. acutatum*. Based on all the results above, it was thought that *Sobic.008G166400* might be the first candidate resistance gene, and *Sobic.008G166550* might be a duplicated gene of *Sobic.008G166400* and therefore another candidate gene.

As mentioned in Chapter 2, the resistant parent P9830 showed HR to another two *C. sublineolum* isolates, Georgia and Sc29. Since we have already designed PCR-based markers (the SNP marker *08.60001890_TG* and two indel markers closely linked to the SNP) co-segregate with the candidate resistance locus, one potential future work is to phenotype the RIL population (or part of the population) with these two strains and genotype the population with the markers to check whether the resistance to Georgia and Sc29 is also contributed by this candidate locus on chromosome 8. Since the resistance to CsGL1 (highly resistant without obvious symptoms) is different from that to Georgia and Sc29 (HR), these two different types of resistance are possibly controlled or regulated by different mechanisms. If those markers also co-segregate with the resistance to Georgia and/or Sc29, it may indicate that the candidate resistance locus contributes to several different mechanisms of plant disease defense and confers broader resistance to anthracnose strains (or even other sorghum pathogens) than we expected initially. If so, however, this candidate locus is only one of the candidate loci conferring resistance to Georgia and/or Sc29, since the rest of genome is not screened and analyzed with respect to those two isolates. In this situation, more gene function work needs to be done to discover and confirm this candidate locus (will be described below) since it may play a role in different plant defense mechanisms. In

addition, *Sobic.008G166400* might also be a homolog copy of a wheat rust resistance gene *Lr1*, and might confer resistance (together with other genes) against rust (Upadhyaya *et al.*, 2013). Based on the RNA-seq analysis, *Sobic.008G166400* might also be involved in the interaction between sorghum and *B. sorghicola* and regulate resistance to target leaf spot (Yazawa *et al.*, 2013). The disease resistance range conferred by *Sobic.008G166400* might be much wider than we expected, but this needs to be confirmed through further study.

As mentioned in Chapter 2, BSA combined with NGS is not able to eliminate the possibility that the candidate genes we found were not the 'real' genes if the reference genome lacks them, and therefore they would be missed when aligning resequenced reads from the resistant parent and RIL pools. Fine-mapping based on the PCR-based indel markers and qPCR analysis cannot solve this potential problem either, although they could provide more confidence that the loci we found on chromosome 8 is related with sorghum anthracnose resistance to CsGL1. To fix this potential risk, third-generation sequencing for the two parents, P9830 and TAM428, and the two DNA bulks (both resistant and susceptible bulks) is a preferable method.

Cloning of candidate genes can be a potential future experiment, since it is a necessary process for future transgenic studies. Genome editing techniques such as clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system have been successfully used in sorghum studies (Jiang *et al.*, 2013), although these reports are rare. With the advance of transformation and tissue culture protocols, sorghum transformation technique is more robust and has a transformation frequency approximately 20% or even higher (Liu and Godwin, 2012; Wu *et al.*, 2014; Lowe *et al.*, 2016). Combining these technologies, and doing transgenic analysis (CRISPR-Cas9) for the candidate resistance genes could be a more solid and convincing way to verify candidate genes' involvement in anthracnose resistance. Gene silencing technology such as VIGS is another way to confirm the candidate genes. With the discovery of post-transcriptional gene silencing, researchers started using virus as vectors to knock down gene expression (van Kammen, 1997; Ruiz *et al*., 1998; Baulcombe, 1999). Comparing with CRISPR-Cas9 or other transgenic systems, VIGS is more time saving and much easier to conduct and manage. The limitation of gene silencing is that

it can only do knockout analysis (probably just knock down) but no over-expression analysis (Burch-Smith *et al.*, 2004). Transgenic analysis with both knockout and over-expression is more recommended since it is more complete and complementary to this study than gene silencing.

## 3.6 References

Amil-Ruiz F, Garrido-Gala J, Gadea J, Blanco-Portales R, Munoz-Merida A, Trelles O, de los Santos B, Arroyo FT, Aguado-Puig A, Romero F, Mercado J, Pliego-Alfaro F, Munoz-Blanco J and Caballero JL (2016) Partial activation of SA- and JA-defensive pathways in strawberry upon *Colletotrichum acutatum* interaction. *Frontiers in Plant Science* 7: article 1036

Baulcombe DC (1999) Fast forward genetics based on virus-induced gene silencing. *Current Opinion in Plant Biology* 2: 109–113

Biruma M, Martin T, Fridborg I, Okori P and Dixelius C (2012) Two loci in sorghum with NB-LRR encoding genes confer resistance to *Colletotrichum sublineolum*. *Theoretical and Applied Genetics* 124: 1005–1015

Burch-Smith TM, Anderson JC, Martin GB and Dinesh-Kumar SP (2004) Applications and advantages of virus-induced gene silencing for gene function studies in plants. *The Plant Journal* 39: 734–746

Burrell AM, Sharma A, Patil NY, Collins SD, Anderson WF, Rooney WL and Klein PE (2015) Sequencing of an anthracnose-resistant sorghum genotype and mapping of a major QTL reveal strong candidate genes for anthracnose resistance. *Crop Science* 55: 790–799

Cannon PF, Damm U, Johnston PR and Weir BS (2012) *Colletotrichum*-current status and future directions. *Studies in Mycology* 73: 181–213

Casado-Diaz A, Encinas-Villarejo S, de los Santos B, Schiliro E, Yubero-Serrano E, Amil-Ruiz F, Pocovi MI, Pliego-Alfaro F, Dorado G, Rey M, Romero F, Munoz-Blanco J and Caballero J (2006) Analysis of strawberry genes differentially expressed in response to *Colletotrichum* infection. *Physiologia Plantarum* 128: 633–650

Cuevas HE, Prom LK, Cooper EA, Knoll JE and Ni X (2018) Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in the U.S. sorghum association panel. *The Plant Genome* 11: 170099

Dean R, van Kan JAL, Pretorius ZA, Hammond-Kosack KE, di Pierto A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J and Foster GD (2012) The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology* 13: 414–430

Encinas-Villarejo S, Maldonado AM, Amil-Ruiz F, de los Santos B, Romero F, Pliego-Alfaro F, Munoz-Blanco J and Caballero CL (2009) Evidence for a positive regulatory role of strawberry (*Fragaria* × *ananassa*) Fa WRKY1 and *Arabidopsis* At WRKY75 proteins in resistance. *Journal of Experimental Botany* 60: 3043–3065

Fang X, Chen W, Xin Y, Zhang H, Yan C, Yu H, Liu H, Xiao W, Wang S, Zheng G, Liu H, Jin L, Ma H and Ruan S (2012) Proteomic analysis of strawberry leaves infected with *Colletotrichum fragariae*. *Journal of Proteomics* 75: 4074–4090

Guidarelli M, Carbone F, Mourgues F, Perrotta G, Rosati C, Bertolini P and Baraldi E (2011) *Colletotrichum acutatum* interactions with unripe and ripe strawberry fruits and differential responses at histological and transcriptional levels. *Plant Pathology* 60: 685–697

Higuera JJ, Garrido-Gala J, Lekhbou A, Arjona-Girona I, Amil-Ruiz F, Mercado JA, Pliego-Alfaro F, Munoz-Blanco J, Lopez-Herrera CJ and Caballero JL (2019) The strawberry FaWRKY1 transcription factor negatively regulates resistance to *Colletotrichum acutatum* in fruit upon infection. *Frontiers in Plant Science* 10: article 480

Jiang W, Zhou H, Bi H, Fromm M, Yang B and Weeks DP (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Research* 41: e188

Jones JDG and Dangl JL (2006) The plant immune system. *Nature* 444: 323–329

van Kammen A (1997) Virus-induced gene silencing in infected and transgenic plants. *Trends in Plant Science* 2: 409–411

Koressaar T and Remm M (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23:1289-1291

Liu G and Godwin ID (2012) Highly efficient sorghum transformation. *Plant Cell Reports* 31: 999–1007

Lowe K, Wu E, Wang N, Hoerster G, Hastings C, Cho M, Scelonge C, Lenderts B, Chamberlin M, Cushatt J, Wang L, Ryan L, Khan T, Chow-Yiu J, Hua W, Yu M, Banh J, Bao Z, Brink K, Iko E, Rudrappa B, Shamseer PM, Bruce W, Newman L, Shen B, Zheng P, Bidney D, Falco C, Register J, Zhao Z, Xu D, Jones T and Gordon-Kamm W (2016) Morphogenic regulators *Baby boom* and *Wuschel* improve monocot transformation. *The Plant cell* 28: 1998–2015

Mertely JC, Forcelini BB and Peres NA (2009) Anthracnose fruit rot of strawberry. Plant Pathology Department, UF/IFAS Extension: PP-207. University of Flordia, IFAS, EDIS, Gainesville

Patil NY, Klein RR, Williams CL, Collins SD, Knoll JE, Burrell AM, Anderson WF, Rooney WL and Klein PE (2017) Quantitative trait loci associated with anthracnose resistance in sorghum. *Crop Science* 57: 1–14

Perumal R, Menz MA, Metha PJ, Katile S, Gutierrez-Rojas LA, Klein RR, Klein PE, Prom LK, Schlueter JA, Rooney WL and Magill CW (2009) Molecular mapping of *Cg1*, a gene for resistance to anthracnose (*Colletotrichum sublineolum*) in sorghum. *Euphytica* 165: 597–606

Ruiz MT, Voinnet O and Baulcombe DC (1998) Initiation and maintenance of virus-induced gene silencing. *The Plant Cell* 10: 937–946

Tanha HM, Naeini MM, Rahgozar S, Rasa SMM and Vallian S (2015) Modified tetra-primer ARMS PCR as a single-nucleotide polymorphism genotyping tool. *Genetic Testing and Molecular Biomarkers* 19: 1-6

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M and Rozen SG (2012) Primer3 - new capabilities and interfaces. *Nucleic Acids Research* 40: e115

Upadhyaya HD, Wang Y, Sharma R and Sharma S (2013) SNP markers linked to leaf rust and grain mold resistance in sorghum. *Molecular Breeding* 32: 451–462

Wu E, Lenderts B, Glassman K, Berezowska-Kaniewska M, Christensen H, Asmus T, Zhen S, Chu U, Cho M and Zhao Z (2014) Optimized *Agrobacterium*-mediated sorghum transformation protocol and molecular data of transgenic sorghum plants. *In Vitro Cellular and Developmental Biology - Plant* 50: 9–18

Xin Z, Velten JP, Oliver MJ Burke JJ (2003) High-throughput DNA extraction method suitable for PCR *BioTechniques* 34: 820–826

Yazawa T, Kawahigashi H, Matsumoto T and Mizuno H (2013) Simultaneous transcriptome analysis of sorghum and *Bipolaris sorghicola* by ssing RNA-seq in combination with *de novo* transcriptome assembly. *PLoS ONE* 8: e62460

Table 3.1. Twenty indels used for fine mapping. The products of these indels have at least ten bp polymorphism between resistant and susceptible individuals.

| Name | Forward primer | Reverse primer |
|---|---|---|
| Del_5945 | GAGAAGGTGGTGGCCAAGTA | TTCCTCCGTGAGAAGCAGTT |
| Del_5947 | GCCTAGCGCCTAGAACTTTG | GGATGGGTAGGATCGCATAG |
| Del_5951 | TGGGCGGTTGTTTTATTACAG | GACCATAAATGCTACAGCTATGAA |
| Del_5952 | AGGTTGGGTGGTTTGTGTGT | GGGCGTCGTGCTAGTTTTAG |
| In_5955 | GTTGCATCGCCAAAAAGATT | GCCATAATGTGTGCAAAAACC |
| Del_5963 | CACGAGAGCGAGGTCTTAGG | GGCGGTTCTTTTAGCCAACT |
| Del_5963_2 | CCATGATTCTCCAGCAGCTC | GGGTGTAGACGAGTGGGAAA |
| Del_5968 | CGTGCCTTGAGTCGTATCCT | CTCTCGAGGTTGCGAAGTTT |
| Del_5971 | AGAGGGACGACCTCTTCTCG | GATGCTCGTTCCTGGTGGT |
| Del_5974 | CTAGGGCACCAGCTAGCAAC | GCTGCTTCATCGCCTTGTAT |
| In_5986 | GGAAACCTATTTTGGGTTTGG | TTGTGGGTCGCTCTCTCTCT |
| Del_6014 | AAGTGCACCAAATCCCACTC | TTTGGAGCAAAATCATGACATC |
| Del_6018 | TGTTGCTGCCTACATGGGTA | CGAGATGCCATCACATTCAC |
| Del_6023 | GCATTGTGAGCAGGTTGAGA | GCTTAGGTGGGCAACAAGAC |
| Del_6030 | GCAAAGTGGTTGTCCGTGTA | CACAAATTCTGGTGAGGAACA |
| Del_6036 | TGACCTTCAAAACAGAACATCAA | GCAGTCCGTGTTCATCGTAA |
| Del_6046 | CCATGTGCTTAGGGCTTGTT | TCCCCATCACATCGAATCTT |
| In_6047 | GTTCTTCCGCGGTCTCAACT | ACGGTCGTTCAGAACAGGTC |
| Del_6048 | TCGCGAGACAAATCTTTTGA | CGTACGTGGTTTTGTGAACG |
| Del_6054 | GGAATACAATGGCACCGTGA | GCTGTATCTACGCCGAGGAG |

Table 3.2. Eighteen single-nucleotide polymorphism markers used for fine mapping. Since these SNPs are tetra primers, primers end with O are outer primers (checks) for the SNPs.

| Name | Forward primer | Reverse primer |
|---|---|---|
| 08.59998388_GC | TTTCCCTCTATTGTCAAATCATATTGAG | GTTTGTGAAAAAAGATTTTTCAGACG |
| 08.59998388_O | CCTTAGGGATGATAGAATTTGTAAGCA | TAAATTCATTTCCTAGGCTCATGAAG |
| 08.59998404_CT | ATCATATTCAGGACTGAAAAATCTGTC | ATGTCTTGATAAATTTGTTTGTGCAA |
| 08.59998404_O | AGTTAGGCTATGAAAATCTAGCACATTC | AAATGTAAATTCATTTCCTAGGCTCAT |
| 08.59998416_TA | GAAAAATCTTTTTTCACAAACGAT | CGGAAGTTCATGTCTTGATAGAT |
| 08.59998416_O | ATCCTCAACGTGTTCAGTCC | TCAACATTTGGCTTGTTGTC |
| 08.59998643_CT | TCAAGACATGAACTTCCGATGCC | AAGGCTCTATGCACCATCAATATAGCAA |
| 08.59998643_O | TTGGCACCTCTCCATCCTCAA | GGAGTTGTCCTTTCAACATTTGGCT |
| 08.60000869_TC | GAAACATCTTGACGCACGATTT | GCAGGATGTTGTAAAGAAATTTCG |
| 08.60000869_O | ATTCATTAGAAGGTCCCTCCCA | TTGCTCATCTTTGCCATCTGA |
| 08.60000923_CT | TTGGAGGCATCCACTCTATTAACTC | ACAGATAATGAACCACCCAAGTTTAA |
| 08.60000923_O | TATACTTTGAGAGGTGGTGACAGGA | CAAGGCAAAGCTCAACTCCA |
| 08.60001041_GC | CAGATGGCAAAGATGAGCATGG | TCCACACTGTCGGTAAATGCTACG |
| 08.60001041_O | ATCCTGCATTCTTTTGGAGGCA | GCCCTGAATGTGCGTCACAC |
| 08.60001049_AT | CATCAGGATCAGATGCCA | GCTTCCCTTGCTCATCATA |
| 08.60001049_O | TTTGGAGGCATCCACTCTA | GCAAGCAGATTTGGTTACCT |
| 08.60001104_GC | TGGAGTTGAGCTTTGCCTTCTG | GGCTAATCAAGGGGTCAACACAG |
| 08.60001104_O | AAGGTTTCTCTGCGGTGAATTGTT | GACGAGTTGGACGGCACCTT |
| 08.60001114_CT | TTGCCTTGTGTCTGTTGCCC | GGAGGCCAAGGCTAATCAATGA |
| 08.60001114_O | TTTCTCTGCGGTGAATTGTTGTG | ACTTCCGCATCCAGGACGAG |
| 08.60001118_TG | TTGCCTTGTGTCTGTTGACCCTTT | CGGAGGCCAAGGCTAACCC |
| 08.60001118_O | TGCGGTGAATTGTTGTGCATG | TCCGCATCCAGGACGAGTTG |
| 08.60001890_TG | GAAACCACAAACCACCATTAATT | CAGTATTGGTTATTAGTGGCATTC |
| 08.60001890_O | CGAGGAGATGTGTGCTTGT | GAGTTTGCAAGCATCGGT |

Table 3.2. (cont.)

| Name | Forward primer | Reverse primer |
|---|---|---|
| 08.60002086_AT | ACCTCAAGCAAGATAACTTGAGAA | CTCAGCTTCTGATTCTCCCA |
| 08.60002086_O | TCGAGCCAGAGGAGAAAAG | TCAGATCCTGCACGTTTCTC |
| 08.60002109_GC | GCGAGAATCAGAAGCTGAGATTG | TGACTCAACAGATGTTCTGACACG |
| 08.60002109_O | CTGATGCTACTTCTGCTGCTCG | TTGAAACAACTTGGGTGGACAG |
| 08.60002405_AG | CAGGGAGAGATGAAACCAACA | GCTTATTAGTGGCACTCGATCAC |
| 08.60002405_O | TGATCTGATCTGTCCACCCA | AACTCTGTCGCCGATTGTGT |
| 08.60002555_CT | TGCCCCAAGTTGCTTTGCGTC | TCGCCGATTGTGTGCGCA |
| 08.60002555_O | CTTGCTTCAGCGTGCAGGGAGA | AGCTCGTGAACGGTGGCGG |
| 08.60002559_GC | CAAGTTGCTTTGCTTCGTGG | ACTCTGTCGCCGATTGTCTG |
| 08.60002559_O | AATTACTTGCTTCAGCGTGCAG | TCAGGATAGCTCGTGAACGG |
| 08.60009020_AT | AGGTTTATTTAGAATAAATAAAAATACACA | TCTTTCCGCCCTTGAATTTA |
| 08.60009020_O | CTCCACTAGAAACATTACCAAAGAT | GACGCCACTGCTAATGTAGTC |

Table 3.3. Long-range PCR primers used to amplify the two candidate genes, *Sobic.008G166400* and *Sobic.008G166550* for subsequent wide-seq.

| Name | Forward primer | Reverse primer | Product size |
|---|---|---|---|
| Sobic.008G166400_WR | CGGGACCAGAGACTCAGGAT | GGCCACCCGCTAAATATTATGAC | 8Kbp |
| Sobic.008G166550_WR | TACGCACTAGGGGAAACCAC | GCACAGCAGATGGAAACAAA | 4Kbp |

Table 3.4. Primers used for amplifying the candidate gene *Sobic.008G166400* transcripts. 166400.3'UTR locates at the 3’ UTR of the gene, and 166400.coding locates at the coding area of *Sobic.008G166400.*

| Name | Forward primer | Reverse primer |
|---|---|---|
| 166400.3'UTR | GCCTCGTCTGTGATCTGTATATG | CGGCGGAAAGAGTAGTCATTAT |
| 166400.coding | CCATTCATCGTAGCCTGATCTT | GGTTGAAGCAGGAACGAAATG |

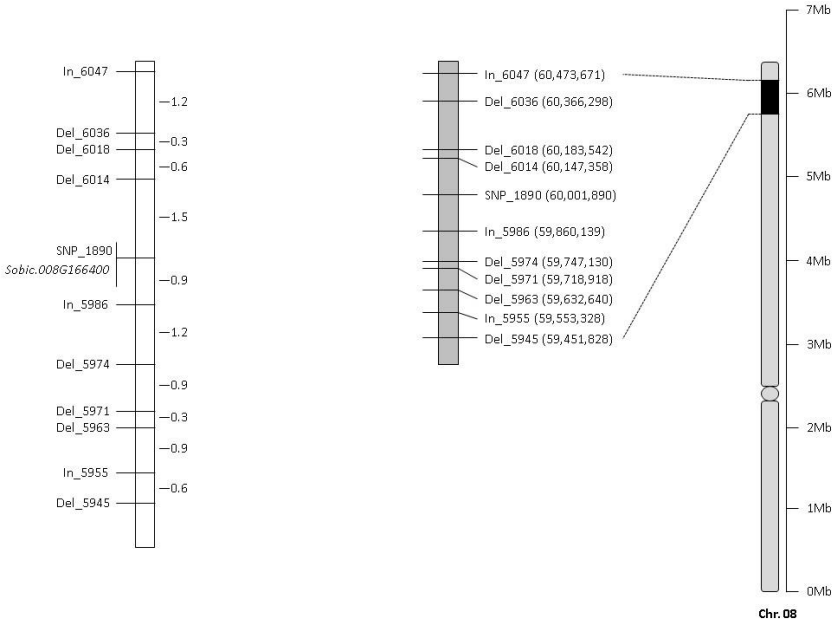


Figure 3.1. Linkage map between positions 59,451,828 and 60,473,671 on chromosome 8 (Phytozome *Sorghum bicolor* v3.1.1) was developed by indel and SNP markers. The left panel shows the genetic distance of the markers with physical intervals indicated on the right.

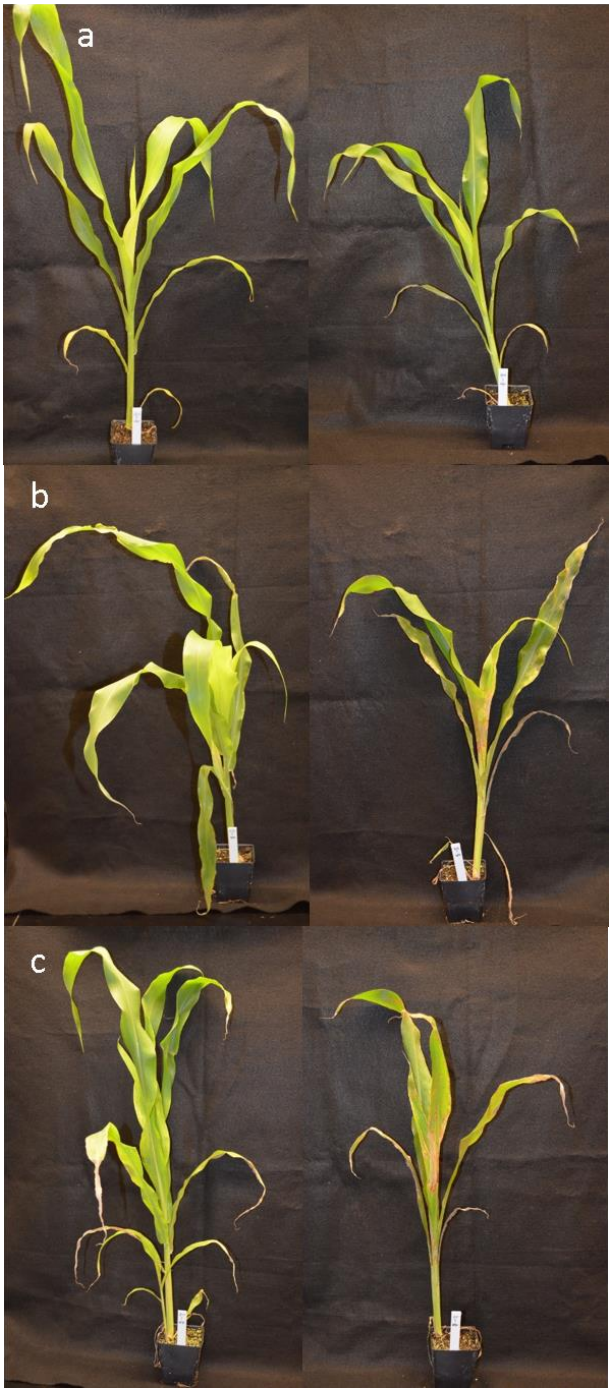


Figure 3.2. Phenotypes of two RILs 17-12 (resistant, left) and 13-31 (susceptible, right) after inoculation at: (a). 0 hpi; (b). 24 hpi; (c). 48 hpi.

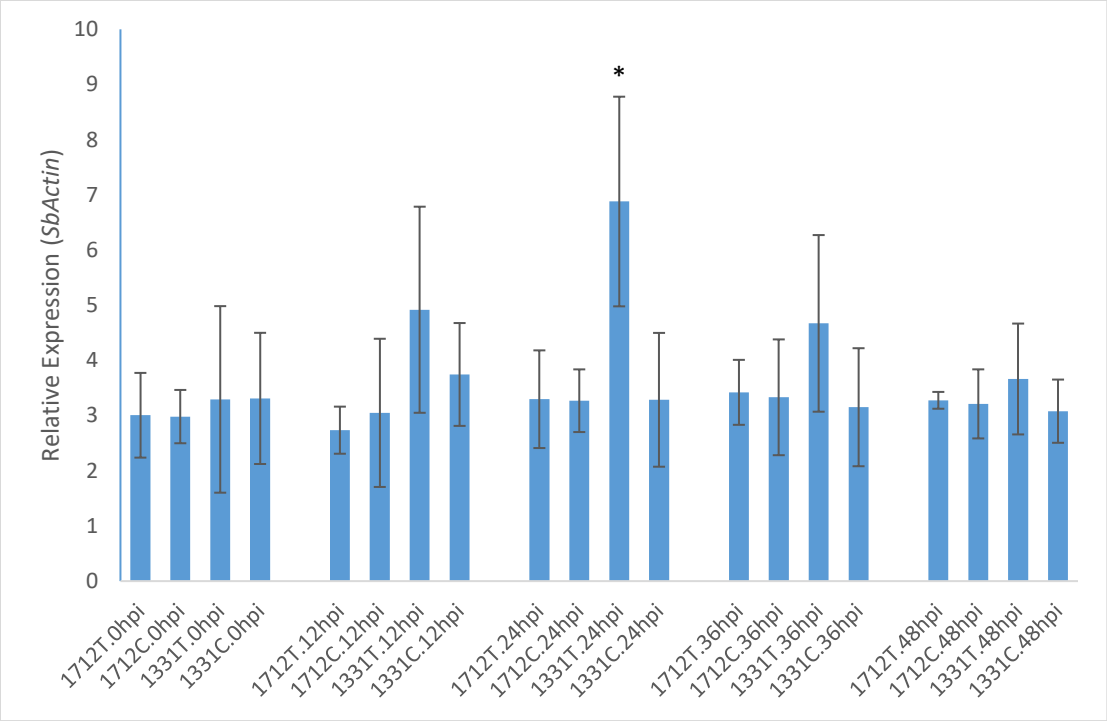


Figure 3.3. The relative expression (to *SbActin*) of the candidate gene, *Sobic.008G166400*, in 13-31 (susceptible) and 17-12 (resistant) at 0, 12, 24, 36 and 48 hpi. T means CsGL1 treatment, and C means check (no pathogen inoculum but double-distilled water). Error bars equal one standard deviation.