GENOME MAPPING OF ANTHRACNOSE RESISTANCE IN SORGHUM GERMPLASM

by

Clara María Cruet Burgos

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

Biology

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS
2018

Approved by:

_____	_____
Hugo E. Cuevas, PhD Member, Graduate Committee	Date
_____	_____
Rafael Montalvo, PhD Member, Graduate Committee	Date
_____	_____
Dimuth Siritunga, PhD President, Graduate Committee	Date
_____	_____
James Beaver, PhD Representative of Graduate Studies	Date
_____	_____
Matias J. Cafaro, PhD Chairperson of the Department	Date

ABSTRACT

Sorghum [*Sorghum bicolor* (L.) Moench] is the fifth most produced cereal in the world with increasing popularity by its ability to grow on marginal lands with minimal input of water and nutrients. Nevertheless, sorghum production is limited by foliar diseases such as anthracnose that can cause yield loses of up to 50% in susceptible varieties. Although several resistance sources have been identified in multiple sources of germplasm, the lack of inheritance studies limits their effective use in sorghum breeding programs. Therefore, three populations of recombinant inbred lines (RILs) derived from the crosses between resistant lines IS18760, QL3 or SC112-14 with the susceptible line PI609251 were evaluated for anthracnose reaction in Texas, Florida, Georgia and Puerto Rico. In parallel, three high density linkage maps based on genotype by sequence analysis were constructed to identify quantitative trait loci (QTL) for resistant response. The IS18760 and QL3 RILs populations showed segregation for anthracnose resistance in Texas and Puerto Rico, and susceptibility against pathotypes from Florida and Georgia, while SC112-14 segregated in the four locations. Genome scan of IS18760 by composite interval mapping detected 4 QTL in Texas and 4 QTL in Puerto Rico, while the scan of QL3 detected 4 QTL in Puerto Rico. Both analyses resulted in the detection of one QTL of 1.63 Mb in chromosome 4 that accounted for 22% and 9% of the phenotypic variation in Texas and Puerto Rico. The genome scan of SC112-14 identified one QTL in chromosome 5 (*Cs-SC112*) that explained 0.51, 0.57, 0.37 and 0.52% of the phenotypic variance in Puerto Rico, Texas, Florida and Georgia, respectively. Subsequently, analysis of 1,500 $F_{2:3}$ segregating individuals reduced *Cs-SC112* locus to an interval of 23 Kb enclosing three candidate genes. Comparative analysis with Plant Resistance Gene Database suggested Sobic.005g172300, which is characterized by the presence of F-box and Ser-Thr kinase domains, is the best candidate gene in the locus. Resequencing analysis of both parental lines identified four point mutations that generate two protein variants. Candidate genes within these loci suggest that resistance response is controlled by multiple defense mechanisms including signaling cascades and transcriptional reprogramming genes. The inheritance information and molecular markers developed in this study can facilitate the use of these resistant sources in breeding programs.

RESUMEN

El sorgo [*Sorghum bicolor* (L.) Moench] es el quinto cereal más producido en el mundo y su popularidad va en aumento por su capacidad de crecer en tierras marginales con un aporte mínimo de agua y nutrientes. Sin embargo, la producción de sorgo se limita por enfermedades foliares como la antracnosis que pueden reducir el rendimiento hasta en un 50% en variedades susceptibles. Aunque se han identificado varias fuentes de resistencia en múltiples fuentes de germoplasma, la falta de estudios de herencia limita su uso efectivo en los programas de mejoramiento. Por consiguiente, se evaluó la respuesta a antracnosis de tres poblaciones de líneas endogámicas recombinantes (LER) derivadas del cruce entre las líneas resistentes IS18760, QL3 y SC112-14 con la línea susceptible PI609251 en Texas, Florida, Georgia y Puerto Rico. Paralelamente, se construyeron tres mapas de ligamiento con alta densidad, basados en el análisis de genotipo por secuenciación, para identificar loci cuantitativos (QTL) para la respuesta de resistencia. Las poblaciones LER de IS18760 y QL3 mostraron segregación para resistencia a la antracnosis en Texas y Puerto Rico, y susceptibilidad frente a patotipos de Florida y Georgia, mientras que SC112-14 segregó en las cuatro localidades. La evaluación del genoma de IS18760 utilizando un mapeo de intervalo compuesto detectó 4 QTL en Texas y 4 QTL en Puerto Rico, mientras que la evaluación de QL3 detectó 4 QTL en Puerto Rico. Ambos análisis resultaron en la detección de un QTL de 1.63 Mb en el cromosoma 4 que explica el 22% y el 9% de la variación fenotípica en Texas y Puerto Rico, respectivamente. En SC112-14 se identificó un QTL en el cromosoma 5 (*Cs-SC112*) que explica 51, 57, 37 y 52% de la varianza fenotípica en Puerto Rico, Texas, Florida y Georgia, respectivamente. Posteriormente, el análisis de 1,500 individuos $F_{2:3}$ redujo el locus *Cs-SC112* a un intervalo de 23 Kb con tres genes candidatos. El análisis comparativo con la base de datos de genes de resistencia de plantas sugirió que el gen Sobic.005g172300, caracterizado por la presencia del dominio F-box y una quinasa Ser-Thr, es el mejor candidato en la región. Un análisis de re-secuenciación de ambas líneas parentales identificó cuatro mutaciones que generan dos variantes de proteínas. Los genes candidatos dentro de estos loci sugieren que la respuesta de resistencia está controlada por múltiples mecanismos de defensa relacionados a la cascada de señalización y reprogramación transcripcional. La información de herencia y los marcadores moleculares desarrollados en este estudio son necesarios para hacer un uso adecuado de estas fuentes de resistencia en programas de mejoramiento.

DEDICATION

To Nancy, Juan and Carla,

Thank you for being my support and strength during my journey. Thank you for never doubting me. Thank you for helping me get up those times I have fallen. Thank you for making our days together so complicatedly simple.

"Life is not measured by the number of breaths we take, but by the moments that take our breath away"- Vicky Corona

ACKNOWLEDGMENTS

I would like to thank Dr. Hugo E. Cuevas for giving me the opportunity to do my research in his laboratory and introduce me to the world of plant breeding. I thank you for being an excellent mentor and having patience to explain, even with drawings, all those things that I did not understand. Your door was always open for questions regarding research as well as for the development of this writing. Thank you for promoting my exposure to conferences and programs that contributed to my development as a person and scientist.

I will always be grateful to Dr. Dimuth Siritunga for being a key player in what I am today as a scientist. When I had not even taken the genetics class, you allowed me to work in your laboratory and introduced me to the world of research with which I am fascinated today. I thank you because when I needed support, you were there for me. Thank you for the commitment you showed with my development as a scientist by accepting me as a master student while helping me to do research in the area I was interested in.

I thank the "yuca sisters" Dianiris, Cristina and Edlin for the unconditional support they have given me during these years. For listening to me when I needed it most, for those coffees that energized me, for those tips that motivated me, for those laughs that made me happy, for those moments that I treasure and take away.

To my laboratory partner for the last year Giseiry, I thank you for your advice and ideas both in the field of research and in my personal life. Thank you for understanding my frustrations and my joys in the laboratory

To Dr. Rafael Montalvo, thank you for accepting my invitation to be a member of my graduate committee and for the support you have given me during my journey.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDIXES

1 INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is a monocot plant belonging to the family Poaceae characterized by its inflorescence (raceme) and grain (caryopsis) in the form of a panicle (Dial, 2012). Carbonized seeds dating 8,000 years BC have been excavated at Nabda Playa archeological site suggesting an early interest and consumption by hunter-gatherers in Africa (Dahlberg & Wasylikowa, 1996; Wendorf et al., 1992). However, its domestication took place approximate 4000-3000 years BC (Dillon et al., 2007) in the Northeast Africa, an area that extends from Ethiopia to surrounding countries. Sorghum dispersion process was parallel to human migration across Africa and other world regions, reaching southern Africa by 10-16th century AD and eastern Africa by 200 AD. During the first millennium BC, it was brought to India and spread throughout Asia by the 13th century (Hagerty, 1941). Captive slaves carried sorghum seeds to the western hemisphere during the 17-18th centuries. This wide range of distribution exposed sorghum through a broad range of environments producing a highly diverse crop that has experienced multiple re-domestication events.

The cultivated sorghum taxa of the world is classified into five major races: Bicolor, Guinea, Caudatum, Kafir and Durra, and the intermediates involving all the pair-wise combinations (Harlan & de Wet, 1972). Each race is differentiated by its inflorescence type and is associated with a re-domestication process to a particular environments (Smith & Frederiksen, 2001; Snowdon, 1936). Bicolor is the most variable and primitive distributed across the African Savannah and Asia. It's characterized by open panicles of medium size with elongated grains, medium in height, tend to tiller profusely and generally has low yield (Dahlberg, 2000). Caudatum, originally from Central Sudan and surrounding areas, is thought to have been derived from a cross between early Bicolor and wild sorghum relative. The panicles of this race are medium to large in size, dense and slightly opened with grains flat on one side and round in the other. This is the most important race for agriculture due to its high yield and superior seed quality. Guinea is mainly distributed in eastern and western Africa, and is characterized by its low yield, long and loose panicles with a flattened, twisting and ovate grain (Dahlberg, 2000). Durra is grown primarily in the drier regions of Arabia and Asia Minor, and on a smaller scale in India, Myanmar, Ethiopia and along the West Nile. It is also thought to originate from a cross between an early Bicolor and wild relative and is adapted to drier conditions in Ethiopia and Eritrea. The race is characterized by its dense and compact panicles that are ovate in shape, with grains medium to large in size and

biconvex in shape. Kafir is the less variable race mainly grown in the southern region of Africa. The panicles are erect, elongated, semi compact, and cylindrical with elliptical grains resulting in a high yielding race (Harlan & de Wet, 1972). The phenotype differences between sorghum races are the result of human selection and the ongoing evolution in their particular environments. As a result, sorghum is now grown in an array of environments, ranging from tropical to temperate regions with a broad spectrum of uses for animal and human consumption.

1.1 Sorghum uses

In agriculture sorghum can be separated into three major types: sweet, forage and grain. Sweet sorghum produces high quantity of soluble sugars in their stalks, and is useful for food, feed and biofuel production. Forage sorghum has a high level of cellulosic biomass with relatively low stalk sugars to be properly digested by ruminants. In addition, it can be used for silage providing the same energy levels of maize. Grain sorghum produces large, starchy seed head on a relatively short stalk and is the most widely produced type. Most grain sorghum (55%) is destined for human consumption while the remainder is used for animal feed and biofuel production. The extensive use of sorghum for different agriculture practices has led to the development of different breeding programs with the objective to satisfy various markets.

Today, more than 500 million people over 30 developing countries rely on sorghum as a dietary staple crop (Wang et al., 2014). In Africa and Asia, sorghum grains are consumed in breads, boiled in porridges, steamed in couscous and in rice-like products, snacks and malt alcoholic and non-alcoholic beverages (Dahlberg et al., 2011). The main grain component is starch followed by protein, non-starch polysaccharides and fat giving an average energetic value of 356 kcal/100g (Dicko et al., 2006). In addition, considerable quantities of vitamins (thiamin, riboflavin, pyridoxine, A, D, E and K) and mineral (phosphorous, potassium, iron and zinc) are present in the grains. Despite its nutritional value, sorghum derived products are scarce in the United States and European markets. However, the recent interest in white sorghum as a gluten free source to substitute wheat from human diet, and the identification of varieties with high concentration of antioxidants has increased its popularity (Awika & Rooney, 2004). In fact, U.S. stakeholders concurred that sorghum consumption and production will increase significantly during the next 10 years.

The interest of sweet and grain sorghum in the renewable energy field have increased in the last 20 years. As a C4 plant, sorghum requires less input of water and nutrients in comparison with corn and wheat (Burrell et al., 2015). In fact, it can have high yield under drought conditions (400-600 mm of rain per year), high altitudes (2300 meters), and is tolerant to waterlogging (Burrell et al., 2015; Dicko et al., 2006). The growing period can be limited to four months compared to the 12-16 months required for sugarcane (Reddy et al., 2005). Moreover, sweet sorghum juicy fermentation efficiency can be up to 90%, while stillage can be use as forage for animals due to being rich in micronutrients and minerals (Haussmann et al., 2002; Reddy et al., 2005). Likewise, the grain fermentable carbohydrates are 23% higher than corn (Rutto et al., 2013). The broad environmental adaptability and low input requirements places sorghum as one of the most promising source for biofuel production. Today, eight ethanol producing plants in United States use 15-20% of the USA grain sorghum production (J. Dahlberg et al., 2011), and multiple breeding and research programs are directed to develop new sweet sorghum cultivars for biofuel production.

1.2 Sorghum production

Sorghum is the fifth most produced cereal in the world after maize, wheat, rice and barley (Food and Agriculture Organization, 2015).The five major producing countries are United States (16.2%), Mexico (12.4%), Nigeria (9.9%), Sudan (9.3%), and India (7.9%) covering 55.7% of the total world production. In fact, America and Africa produces 39.2% and 34.5% of the production, respectively, while the other 23.4% is produced by Asia, Europe and Oceania. The majority of sorghum production in U.S. is located in the dryer central and southern regions, however, is expanding to the warm and humid regions of the south. Although 21 states produce sorghum, Kansas and Texas accounted for 72% of the total area planted with 3.3 and 3.1 million acres, respectively. Over the last 50 years the yield per acre has increased 0.53% annually due to the development of new cultivars and the improvement in farming techniques (Food and Agriculture Organization, 2015), however, the rate of yield increase is lower than the observed in other staple crops such as maize and wheat. This production constraint could be associated to the scarcity of varieties with tolerance to biotic and abiotic factors.

Sorghum production in U.S. and around the world has been limited by several abiotic and biotic factors. Although sorghum is more tolerant to abiotic stress such as drought, heat, salinity,

flooding, and heavy metal toxicity compared to other cereal crops, these abiotic stresses can reduce yield up to 50% in some varieties (Thakur et al., 2011). Likewise, insect pests such as stem borers *(Busseola fusca*, *Chilo partellus*, *Sesamia calamistis*), midge (*Contarinia sorghicola*) and sugarcane aphids (*Melanaphis sacchari*), in combination with major diseases like anthracnose (*Colletotrichum sublineola),* downy mildew (*Peronosclerospora sorghi*) and charcoal rot (*Macrophomina phaseolina*) have the potential to reduce the yield up to 100% in highly infected fields (Wang et al., 2014). Therefore, current research efforts are directed to the identification and development of more tolerant and/or resistant cultivars that may possibly retain and increase sorghum yield worldwide.

1.3 Anthracnosis

One of the major constrains in sorghum production is anthracnosis, a fungal disease caused by *Colletotrichum sublineola P*. Henn., in Kabat & Bubal (syn. *Colletotrichum graminicola* (Ces.) G. W. Wilson). The fungus infects all above-ground tissues: stalks, lead, peduncle, panicle and grain and it can develop in live or dead tissues (Néya & Le Normand, 1998). The anthracnose disease appears as a leaf blight, stalk rot and head blight, and the most common. Severe infection on leaves can reduce yield by 50% or more in susceptible cultivars (Pande et al., 1994). The disease was first reported in North Carolina and Texas, U.S. in 1911 and 1912, respectively. Today it is an important disease in southeast states such as Georgia, Texas, Arkansas, and Florida (Pastor-Corrales & Frederiksen, 1980). Likewise, anthracnose has become a disease of economic importance in Latin America, where it is one of the main yield-limiting factors in Brazil, Venezuela, and Guatemala. In addition, the tropical and subtropical regions that are used as winter nursery by seed companies, such as Puerto Rico, have been affected by the disease as well.

Warm and humid environmental conditions are most suitable for the development and spread of anthracnose (Pande et al., 1994). The spores of *C. sublineola* disperse through rain splash and adhere to the surface on any above-ground tissue of the plant where it germinates (Perfect et al., 1999) (Figure 1). The maturation of the apressorioum involves the formation of the penetration pore followed by the infection peg directed by the physical and chemical signals of the plant surfaces. The cuticle is deformed inward meaning initial penetration of cell wall by the fungus involves exertion of mechanical force (Wharton et al., 2001). At 36 hours after inoculation, infection vesicles are observed, and before 48 hours, the emergence of the primary hyphae occurs.

Secondary hyphae begin to emerge and ramifies throughout the leaf sheath and infected epidermal cells collapse (Wharton et al., 2001).

Anthracnose infection in the leaves show small circular discolored spots that can range in size from 2 to 6 mm ($^{1}/_{16}$ to $^{1}/_{4}$ of an inch) in diameter (Lebeau, 1951). The color of the initial laceration varies between cultivars from reddish to blackish, however, the center becomes darker while host cells die and acervuli with prominent setae appear (Figure 2). The infection in leaves occurs from the 3-4 leaf stage to mature plants, but highly susceptible cultivars die before maturity. The infection in the stalk is known as the red rot characterized by a discoloration of the vascular tissue causing drying and shrinking that eventually lead to its breaking (Lebeau, 1951). The greatest yield reductions are observed when the stalk midrib and foliar infections occurs together.

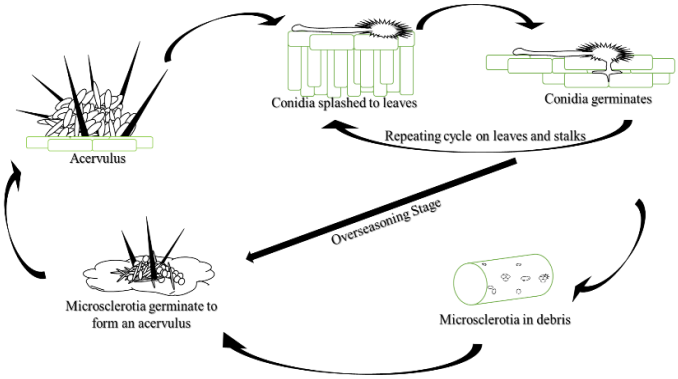

Figure 1. Life cycle of *Colletotrichum sublineola* causal agent of anthracnose disease in sorghum [*Sorghum bicolor* (L.) Moench]

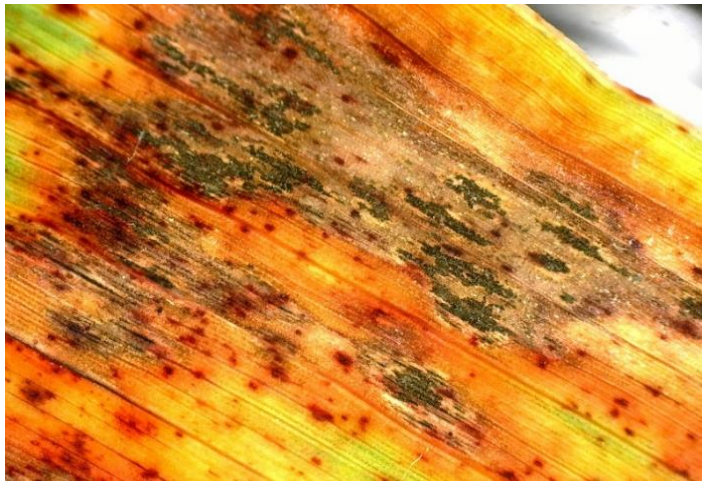

Figure 2. Symptom of anthracnose (*Colletotrichum sublineola*) disease on sorghum [*Sorghum bicolor* (L.) Moench] leaves.

The pathogen population of *C. sublineola* is genetically highly diverse with several pathotypes having been identified based on their virulence (Prom et al., 2012). These pathotypes emerge from different geographical areas around the world (Casela & Frederiksen, 1993; Marley, Thakur et al., 2001; Pande et al., 1991; Valèrio et al., 2005) and/or from particular regions within countries (Ali & Warren, 1987; Cardwell et al., 1989; Moore et al., 2008; Prom et al., 2012).

Recent genetic diversity studies among isolates from Arkansas, Georgia, Texas, and Puerto Rico found that the virulence is not associated with a particular geographic region (Prom et al., 2012). Therefore, the development of new resistant varieties might require multiple resistance genes to cope with the pathogen diversity present in a particular location. Today, pathotypes are defined based on the disease reaction response of eighteen sorghum differential lines (Prom et al., 2012). Unfortunately, the genetic control underlying resistance responses in differential lines is only known for SC748-5 (Perumal et al., 2009); SC112-14 (Cuevas et al., 2014), and SC414-12E (Patil et al., 2017), and no resistant genes have been isolated that can lead to a better understanding of the molecular mechanism.

Multiple practices have been established to achieve an adequate control of the disease (Resende et al., 2013). Crop rotation has been used to prevent losses however, it is not a feasible option for highly susceptible cultivars (Burrell et al., 2015; Crill & Khush, 1982; Silva et al., 2015). A transgenic approach using two genes from *Trichoderma harzianum* (chitinase and chitosanase) could not provide a complete resistant response (Kosambo-Ayoo et al., 2011). In this regard, the genetic compatibility of sorghum with Johnson grass (*Sorghum halapense*), limit the opportunity of using transgenic sorghum plant in open field production. Today, the most economical and environmentally safe control of anthracnose is the use of resistant cultivars. The evaluation of exotic sorghum germplasm accessions from China, Ethiopia, Mali, Mozambique, Sudan, Uganda, Zimbabwe, South Africa, and Burkina Faso has identified highly resistant germplasm (Cuevas & Prom, 2013; Erpelding & Prom, 2004; Erpelding & Prom, 2006; Perumal et al., 2009; Prom et al., 2012). Nevertheless, the lack of inheritance studies and limited knowledge of the genetic relatedness among the majority of these sources of resistance limits their effective use in sorghum breeding programs. In fact, the narrow genetic diversity within germplasm from a particular country suggests the presence of redundant sources (i.e. identical by descent). Today, several commercial hybrids cultivars produced in United States use the sorghum line SC748-5 as the primary source of resistance, but the high genetic diversity and mutation rate of the anthracnose pathogen suggest new virulent pathotypes may arise and resistance collapse. The identification of genes/alleles associated with multiple resistance sources is imperative to pyramid resistant genes, optimize their use and to increase resistance durability.

1.4 Anthracnose inheritance

The genetic control of anthracnose resistance response is not well understood. The first inheritance study suggested that two linked dominant genes provide resistance to different phases of infection (Coleman & Stokes, 1954). Nevertheless, subsequent inheritance studies suggest a recessive gene control resistance in lines SC326-6 (Boora et al., 1997) and G73 (Singh, 2006), while a dominant gene control resistance in lines SC748-5 (Mehta et al., 2005) and SC112-14 (Cuevas et al., 2014). The discrepancy among inheritance studies could be attributed to the existence of multiple resistant genes or response to different pathotypes.

The genome mapping of eight resistant sources in U.S. have resulted in the identification of 11 loci distributed in chromosome 2 (Patil et al., 2017), 4 (Patil et al., 2017), 5 (Cuevas et al., 2014; Patil et al., 2017; Perumal et al., 2009), 6 (Klein et al., 2001), 7 (Felderhoff et al., 2016) and 9 (Biruma et al., 2012; Felderhoff et al., 2016; Patil et al., 2017). The first mapped resistance locus was detected in chromosome 6 by the analysis of RILs derived from the cross of Sureño and RTx430 (Klein et al., 2001). A linkage block of several resistant genes in the distal region of chromosome 5 was associated with the resistant response in SC748-5 [*Cg1*; Burrell et al., (2015); Perumal et al., (2009)]. The resistant locus present in SC112-14 (*Cs-SC112*), which confers resistance against pathotypes from Texas, Arkansas and Puerto Rico, was mapped in proximity to the *Cg1* locus (Cuevas et al., 2014). Comparative genomic analysis determined that *Cs-SC112* and *Cg1* are independent loci. Recently, the resistant locus from SC414-12E was mapped 5 Mb away from the *Cg1* and overlapping with *Cs-SC112 (*Patil et al., 2017). The contrasting transcript profiling among resistant (BS04/BS05) and susceptible (MU07/193D) lines from Uganda associated two loci on chromosome 9 (*Cs1A* and *Cs2A*) with the observed resistance response (Biruma et al., 2012). Genomic analysis revealed that each of these loci are closely linked and both enclose genes with a nucleotide binding-leucine rich repeat (*Cs1B* and *Cs2B*, respectively). The resistant response present in the cultivar "*Bk7*" was genomic dissected into two loci located on chromosome 7 and 9 (Felderhoff et al., 2016). The locus in chromosome 9 (expanded 3.2 Mb) explained majority of the resistant response, was validated in other segregating populations, and characterized by the presence of 36 resistant genes. This locus was also associated with the resistant response present in SC155-14E (Patil et al., 2017). Certainly, these resistance sources

should be further studied to determine the most adequate strategy to its introgression and uses in United States breeding programs.

Association mapping study using the ICRISAT minicore collection and pathotypes from India led to the discovery of eight regions associated with anthracnose resistance located in chromosomes 1, 6, 8 and 10 (Upadhyaya et al., 2013). Moreover, the analysis of RIL population derived from the resistant line 296B and IS18551 identified loci in chromosomes 4, and 6, while in RIL population derived from the resistant line HC136 and susceptible line G73, identified a QTL in chromosome 8. The location of the QTLs were not precise, but are independent from those identified with pathotypes from U.S. This shows that sorghum genome enclose multiple loci for anthracnose resistance , and their exact identification and isolation are needed to develop highly resistance cultivars.

Previous researchers have agreed that anthracnose resistance is under the control of multiple loci that provide protection against particular pathotypes. Today, the defensive molecular mechanism involved in these resistant responses is not well understood. Improving our understanding of host/pathogen relationships against multiple pathotypes and the inheritance from different resistant sources is needed to increase the durability of resistance. This research aims to study the disease reactions in sources of resistance from Sudan (IS18760), Ethiopia (SC112-14) and India (QL3), and to identify candidate genes within the *Cs-SC112* resistant locus. The development of molecular markers tightly linked to resistance loci will facilitate the development of new resistant varieties and the screening of exotic sorghum germplasm based on molecular approaches.

2 OBJECTIVES

1) Genomic dissection of anthracnose resistance response present in sorghum differential lines IS18760 (Sudan), QL3 (India) and SC112-14 (Ethiopia)

2) Identification of candidate genes within the resistant locus *Cs-SC112*

3 MATERIALS AND METHODS

3.1 Recombinant Inbred Lines and Field Experiments

Three biparental recombinant inbred lines (RILs; $F_{3:4}$ and $F_{4:5}$) populations derived from the cross of PI609251 and three anthracnose resistance sources (IS18760, QL3 and SC112-14) were used to study the inheritance of anthracnose resistant response (Figure 3; *unpublished Hugo E. Cuevas, USDA-ARS*). These RILs populations are composed of 135, 97 and 111 lines, respectively, and hereafter referenced according to their unique resistant parent, IS18760, QL3 and SC112-14.

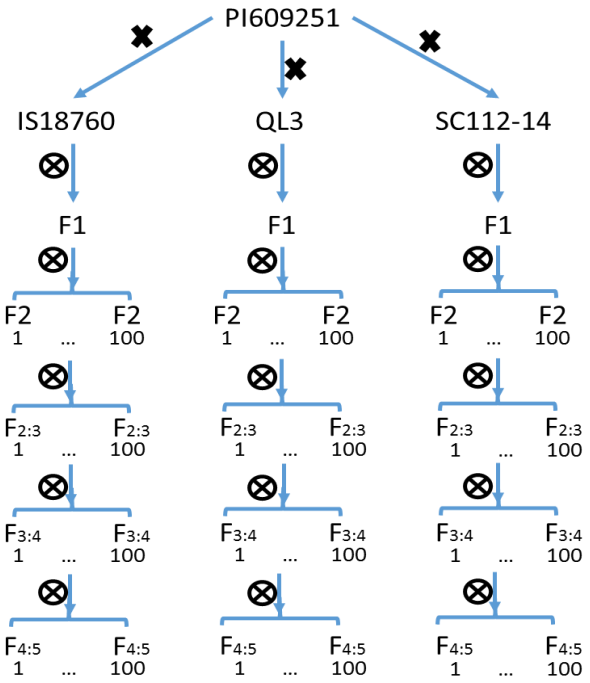

Figure 3. Schematic design for the development of three sets of recombinant inbred lines (RILs) in sorghum [*Sorghum bicolor* (L.) Moench]. Three anthracnose resistant lines (IS18760, QL3 and SC112-14) were crossed to the common susceptible line PI609251 to derived the RILs

The three RILs populations and parental lines were evaluated for disease reaction after inoculation with anthracnose isolates at Texas A&M University Research Farm, College Station, Texas (TX), University of Florida Suwannee Valley Agricultural Extension and Education Center in Live Oak, Florida (FL) and at the University of Georgia College of Agricultural and Environmental Sciences, Tifton, Georgia (GA) during the summer of 2016. The RILs were also evaluated at the experimental farm of the Tropical Agriculture Research Station in Isabela, Puerto Rico (PR) during the fall of 2016. The experimental design at all locations was a randomized complete block design (RCBD) consisting of two blocks with plots of 3.1 m in length and 0.9 m between rows.

3.2 Anthracnose resistant response and flowering time

The inoculation and disease assessment methods were similar to those described by Prom et al. (2009). Initially, isolates of *C. sublineola* pathotypes representing the pathotype diversity

present in each location were used to colonize sorghum seeds. These seeds were then used to inoculate 30-40 day old plants by placing approximate 10 *C. sublineola*-colonized seeds into the leaf whorls. Disease assessment was conducted on mature plants (i.e. 30 days after flowering) based on a scale 1-5 (Table 1). The average of the two replicates was then used to categorize the anthracnose response of the plants into resistant (1-2), moderate resistant (2.5-3) and susceptible (≥3).

Table 1. Disease assessment scale use to evaluate anthracnose (*Colletotrichum sublineola*) resistant response in sorghum [*Sorghum bicolor* (L.) Moench].

Score	Symptoms
1	No symptoms or chlorotic flecks on leaves
2	Hypersensitive reaction on inoculated leaves but not acervuli in the center
3	Lesions on inoculated leaves with acervuli formation and no symptoms observed on other leaves
4	Necrotic lesions with acervuli observed on inoculated and bottom leaves with the infection spreading to middle leaves
5	Most leaves necrotic due to infection including infection on the flag leaf

Scale based on Prom et al. (2009)

3.3 Flowering time

Flowering times were obtained for each plot in Puerto Rico and Florida, defined as the number of days until 50% of the plants within a plot reached anthesis. The average of the two replicates was used to determine flowering time of a RIL at each location. Phenotypic correlation between flowering time and anthracnose resistant response were calculated using PROC CORR in SAS student version 9.4.

3.4 Statistical analysis

Anthracnose resistant response for combined locations was subject to analyses of variance (*ANOVA*) using the mixed linear model:

$$Y = \mu + L + B(L) + R + RxL + \varepsilon$$

and when data for each location was analyzed individually, the *ANOVA* mixed linear model was:

$$Y = \mu + B + R + \varepsilon$$

The prefix *Y* denotes the anthracnose resistant response, μ the common effect, *B* the block effect, *R* refers to the effect of the genotype (RILs), *L* is the effect of the location, *B(L)* is the effect of blocks nested in the locations, *RxL* is the RILs by environment interaction and ε is the error term. The locations were treated as fixed effects, while *B, R, RxL*, and *B(L)* were treated as random effects. The *ANOVA* were performed in SAS student version 9.4 using *PROC MIXED* with method *TYPE 3*.

3.5 Heritability

Estimates of variance components were used to calculate narrow-sense heritability (h^2) for anthracnose resistance response based on progeny-means basis for combined locations using the formula:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{GXL}^2}{l} + \frac{\sigma_e^2}{rl}}$$

and for each location using the formula:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}}$$

The prefix σ_g^2, σ_e^2, σ_{GXL}^2 refers to the genotypic (RILs), error and genotype by location interaction variances, respectively, while *r* and *l* are the number of block and locations, respectively. The standard error of the narrow-sense heritability was calculated as SE (h^2) = $\frac{\sqrt{\sigma_g^2}}{\sigma_P^2}$ where the σ_P^2 corresponds to the phenotypic variance. The phenotypic variance for the combined location is given by $\sigma_g^2 + \frac{\sigma_{GXL}^2}{l} + \frac{\sigma_e^2}{rl}$ and for the individual locations by $\sigma_g^2 + \frac{\sigma_e^2}{r}$.

3.6 Genome Mapping

3.6.1 Genotype by sequence

The DNA was isolated from seedlings using the methods describe by Guillemaut and Maréchal-Drouard (1992) with some modifications, purified with ZR 96 DNA Clean & Concentrator-5 (Zymo Research, Irvine, CA, USA), and quantified with Quantus Fluorometer (Promega, Madison, WI). Subsequently, the DNA were normalized to 5 ng/µL using HPLC water

and subjected to genotyping-by-sequence (GBS) (Elshire et al., 2011). First, 50 ng of genomic DNA and 1.5 ng of adapter (i.e. Ilumina adapter and barcode) were mixed and digested using the ApeKI enzyme (New England BioLabs, Ipswich, MA) at 75^{o}C for 2 hours, and ligated using T4 DNA Ligase (New England BioLabs, Ipswich, MA) at 22^{o}C for 1 hour followed by its denaturalization at 65^{o}C for 30 minutes. Barcode RILs DNA were pooled and cleaned-up using QIAquick PCR Purification columns (QIAGEN, Hilden, DE), and amplified using adapter specific primers. The reactions were purified using QIAquick PCR Purification columns to elute a GBS library. Quality and libraries concentrations were determined by capillary electrophoresis at Weill Cornell Medical College core facility. These GBS libraries were single-end sequenced at the iLab facility in Weill Cornell Medical College using the Ilumina platform High-seq 2500 or Nextseq500 high-output V2 with the 75 cycle kit. Illumina raw sequencing data were process through Tassel 3.0 (Glaubitz et al., 2014), and aligned to sorghum reference genome v.3.1 (Paterson et al., 2009) to call SNPs.

3.6.2 Recombination Linkage Maps

High density recombination linkage maps for each RIL populations were generated using SNPs from GBS. Firstly, SNPs were filtered based on minimum allele frequency (>0.40) and percent of missing (<50%) using Tassel 5.0 (Glaubitz et al., 2014). Since heterozygous genotypes are not informative in RIL linkage maps, they were set as missing, and SNPs filtered again based on percent of missing (<50%). Missing genotypes were imputated with NPUTE (Roberts et al., 2007) analyzing each chromosome and RIL populations independently.

The SNPs of each RIL population were converted to AB format using Tassel 5.0, where A refers to genotype similar to resistant parental lines. The R package OneMap F2/RIL (Margarido et al., 2007) was used to select groups of SNPs linkage with an LOD≥5 and a maximum recombination frequency of 0.20. The groups of linkage SNPs were ordered based on their physical genomic position in each chromosome, and recombination distance estimate based on Kosambi function. To improve the quality of linkage maps for QTL analysis, a Perl script (Nave et al., 2016) was used to systematically select approximately 30 high quality SNPs evenly distributed per chromosome based on genetic distance and recombinant quality information. First, SNPs were removed from a specific map interval if it introduced a double crossing over (DCO) and another high-quality SNP is present within the interval. Since SNPs saturation varies among chromosome

and RIL populations, these scanning chromosome intervals were adjusted from 0.1 to 8 cM in IS18760, 0.5 to 2 cM in QL3 and 0.5 to 4 cM in SC112-14. The recombination distance between this high-quality SNPs were re-estimated in R package OneMap F2/RIL based on Kosambi function. Graphical representations of linkage maps were created using software MapChart (Voorrips, 2002).

3.6.3 QTL mapping

Quantitative trait locus analysis for anthracnose resistant response in each location were performed using Composite Interval Mapping (CIM) (Jansen & Stam, 1994; Zeng, 1994) as implemented in Windows QTL Cartographer (Basten et. al 2004). The standard backward regression method (i.e. Model 6) with walking speed of 0.5 cM were used for the CIM of the three populations. The window size ranged from 5 cM (IS18760) to 10 cM (SC112-14 and QL3) with up to 10 SNPs to control background effects. A QTL was declared significant based on the LOD threshold calculated using 1,000 permutations with an experimental-wise error rate of P=0.05 for each location and RILs population.

3.7 Identification of candidate genes in Cs-SC112 locus

3.7.1 Segregating progenies

The main QTL of the RIL population derived from SC112-14, located in distal region of chromosome 5 (*Cs-SC112*), was fine mapped to identify candidate genes. Two $F_{2:3}$ families segregating within the QTL region were identified based on genotyping analysis of F_2 populations used to develop the RILs (Cuevas et al., unpublished).

3.7.2 KAPS Markers

To delimit the *Cs-SC112* locus a series of SNPs were selected to develop Kompetitive Allele Specific PCR (KASP; LGC genomics) markers (Table 2). The KASP assay PCR reaction involve two allele-specific forward primers, each with a unique tail sequence that correspond with the FRET cassettes labelled with FAM or HEX dye, and one common reverse primer. These three primers reaction allow to determine alleles based on the FAM or HEX fluorescence cassette signal. The reactions fluorescence signals were read with Agilent Mx3005P QPCR System and the

KlusterCaller program (LGC Genomics) was used to assign corresponding genotypes. Reaction conditions for KASP assay were according to recommend concentrations of manufactures, with the addition of 1% v/v PVP, 1% v/v BSA and 15% v/v of GC Enhancer (Supplementary Table A1).

Table 2. Kompetitive Allele Specific PCR (KASP) assay markers designed for the identification of recombinants within the anthracnose resistant locus *Cs-SC112*.

Marker Name	Position
SNP_Sb005g022920	64784456
SNP_Sb005g55256930	64840543
SNP_Sb005g172300	65194327
SNP_Sb005g55625107	65208713
SNP_Sb005g56057061	65640662

3.7.3 Delimitation of Cs-SC112

A total of 1,500 segregating individuals were germinated in the greenhouse and DNA from seedlings were isolated as described previously. To identify recombination events seedlings were genotyped with KASP markers that flank the *Cs-SC112* locus (Table 2). Later, recombinant individuals were genotyped with additional KASP markers, or amplicon directed sequencing that target SNPs within the *Cs-SC112* locus. Recombinants were transplanted to 5 L pots up to approximately 5th leaf stage, afterward, were transplanted to the experimental farm of the Tropical Agriculture Research Station in Isabela, PR to determine disease reaction to anthracnose. Plant inoculation and evaluations were similar to those previously described herein (Prom et al. (2009).

3.7.4 Candidate gene sequence

Partial sequence of candidate gene was performed to the parental lines and/or recombinant progenies to identify protein variants. Gene sequence for Sobic005g172300 was obtained from Phytozome (https://phytozome.jgi.doe.gov) and primers designed with IDT's PrimerQuest© (https://www.idtdna.com/primerquest). DNA amplification was completed using Taq Polymerase (Promega, Madison, WI) with the manufacturer recommended conditions and an annealing temperature of 60^0C. Amplicons were verified through agarose electrophoresis and cleaned with ExoSAP-IT PCR Product Cleanup Reagent (Thermo Fisher Scientific, Waltham, MA). Sequencing was based on BigDye terminator chemistry at SeqWright Genomic Service (Houston,

TX) and sequence chromatograms examined using Sequencher v 4.1 (Gene Codes, Ann Arbor, MI).

4 RESULTS

4.1 Phenotyping and Statistical Analysis

The disease reactions of the RIL populations and parental lines varied among locations (Table 3). The parental lines IS18760 and QL3 were moderately resistant (score 2.5) and resistant (score 2.0), respectively, in Puerto Rico and Texas, but, susceptible in Florida and Georgia. The parental line SC112-14 had a resistant (score 2) across the four locations. The RIL populations derived from IS18760 and QL3 segregated for disease reaction to anthracnose isolates in Puerto Rico and Texas, while the RIL population derived from SC112-14 segregated at the four locations.

Analysis of variance for each RIL population showed that some lines have different disease reactions to anthracnose isolates across locations (i.e. genotype x environment interaction; Table 4). Differences among blocks were observed for the QL3 RIL population in Texas (Table 5 and 6). Narrow sense heritability across locations were estimated to be 0.66, 0.39 and 0.90 for IS18760, QL3 and SC112-14, respectively. In Puerto Rico, narrow sense heritability of IS18760 and QL3 RILs were estimated to be 0.73 and 0.43, respectively, while in Texas were estimated to be 0.43 and 0.69, respectively. In SC112-14 RILs populations narrow sense heritability varied from 0.75 in Texas, 0.82 in Puerto Rico and 0.87 in Florida and Georgia, respectively. Anthracnose readings and flowering time were not correlated at any location.

The IS18760 and QL3 RILs population exhibited a skewed distribution to susceptibility in Puerto Rico and normal distribution in Texas. In Puerto Rico, 75% of the IS18760 RILs were susceptible, 20% moderately resistant and 5% resistant, while in Texas 68% were moderately resistant, 24 % resistant and 8% susceptible (Table 3, Supplementary Table A2). Anthracnose readings averaged 3.94 and 2.73 in Puerto Rico and Texas, respectively. Eleven RILs that were susceptible in Puerto Rico were resistant in Texas, while five were resistant at both locations. In Puerto Rico, only 2% of the QL3 RILs were resistant whereas 12% were moderately resistant and 86% were susceptible. In Texas 29% of the RILs were resistant, 63% moderately resistant and 8% susceptible (Table 3, Supplementary Table A3). Anthracnose readings averaged 4.19 and 2.71 in Puerto Rico and Texas, respectively. Two RILs were resistant in both locations, while fourteen that were susceptible in Puerto Rico were resistant in Texas.

The SC112-14 RILs population exhibited a bimodal distribution (resistance or susceptible) in Puerto Rico, Florida and Georgia, and a shift towards resistance in Texas (Table 3,

Supplementary Table A4). In Florida, Georgia and Puerto Rico, 48%, 34%, and 30% of the RILs were resistant, respectively, while 50% 44%, and 61%, were susceptible, respectively. In Texas, 42% and 44% classified as resistant and moderately resistant, respectively, while 14% were susceptible. Anthracnose resistant response averaged 3.60, 2.62, 3.23, and 3.14 for Puerto Rico, Texas, Georgia and Florida, respectively. Eleven RILs were resistant (readings <2.0) at all locations, while another 10 RILs were resistant or moderately resistant across locations.

Table 3. Anthracnose (*Colletotrichum sublineola*) resistant response of three sets of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred lines (RILs) and parental lines evaluated during 2016 in Texas, Florida, Georgia, and Puerto Rico. The RILs derived from resistant lines IS18760, QL3, and SC112-14 were crossed with the common susceptible line PI609251.

Location	Parents (disease reaction)		RILs (number of lines)		
	IS18760	PI609251	Resistant[1]	Moderate Resistant[1]	Susceptible[1]
Puerto Rico	MR	S	6	23	90
Texas	MR	MR	24	69	8
Florida	S	S	-	-	-
Georgia	-	S	-	-	-
	QL3	PI609251	Resistant	Moderate Resistant	Susceptible
Puerto Rico	R	S	2	10	71
Texas	R	MR	22	48	6
Florida	S	S	-	-	-
Georgia	S	S	-	-	-
	SC112-14	PI609251	Resistant	Moderate Resistant	Susceptible
Puerto Rico	R	S	21	6	43
Texas	R	MR	31	32	10
Florida	R	S	29	11	30
Georgia	R	S	24	16	31

[1]Anthracnose resistant response according to Prom et al., (2009) where score 1.0-2.0 were resistant (R), 2.5-3.0 moderate resistant (MR), and 3.0-5.0 susceptible (S).

Table 4. Analysis of variance and narrow sense heritability (h^2) of anthracnose (*Colletotrichum sublineola*) disease reaction of three sets of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred lines (RILs) evaluated during 2016 in Texas, Florida, Georgia, and Puerto Rico. The RILs derived from resistant lines IS18760, QL3 and SC112-14 crossed with the common susceptible line PI609251.

	IS18760[1]			QL3[1]			SC112-14[2]		
Source of variance	Degree of Freedom	Mean Square		Degree of Freedom	Mean Square		Degree of Freedom	Mean Square	
Location (L)	1	118.48	***	1	114.88	**	3	21.26	***
Blocks(L)	2	0.37		2	1.35	*	4	0.27	
Genotype (G)	119	1.27	***	89	0.81	*	78	6.30	***
GXL	99	0.46	**	67	0.54	*	208	0.72	***
Residual	204	0.30		113	0.37		264	0.35	
h^2		0.66			0.39			0.90	

[1] Analysis of variance for anthracnose disease reaction in Puerto Rico and Texas.

[2] Analysis of variance for anthracnose disease reaction in Puerto Rico, Florida, Georgia and Texas.

*, **, and *** refers to significant at $p \leq 0.05$, 0.01, and 0.001, respectively.

Table 5. Analysis of variance and narrow sense heritability (h^2) of anthracnose (*Colletotrichum sublineola*) disease reactions of two sets of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred lines (RILs) evaluated during the summer 2016 in Puerto Rico and Texas. The RILs derived from anthracnose resistant lines IS18760 and QL3 crossed with the common susceptible line PI609251.

	IS18760				**QL3**			
	Puerto Rico		Texas		Puerto Rico		Texas	
Source of variance	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square
Blocks	1	0.02	1	0.72	1	2.01	1	0.69 *
Genotype	118	1.27 ***	100	0.45 **	81	0.99 *	75	0.46 ***
Residual	105	0.35	99	0.25	55	0.61	58	0.15
h^2		0.73		0.45		0.43		0.69

*, **, and *** refers to significant at $p \leq 0.05$, 0.01 and 0.001, respectively.

Table 6. Analysis of variance and narrow sense heritability (h^2) of anthracnose (*Colletotrichum sublineola*) disease reactions of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred lines (RILs) evaluated during 2016 in Puerto Rico, Texas, Florida and Georgia. The RILs derived from anthracnose resistant line SC112-14 crossed with the common susceptible line PI609251.

	Puerto Rico		Texas		Florida		Georgia	
Source of variance	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square	Degree of Freedom	Mean Square
Blocks	1	0.52	1	0.03	1	0.45	1	0.06
Genotype	74	2.82 ***	72	0.85 ***	70	2.70 ***	70	2.65 ***
Residual	61	0.55	63	0.22	70	0.31	70	0.33
h^2		0.82		0.75		0.87		0.87

*, **, and *** refers to significant at $p \leq 0.05$, 0.01 and 0.001, respectively

4.2 Genotype-by-sequence

Genotype-by-sequence analysis for the RILs populations resulted in the identification of 60,828 (IS18760), 5,101 (QL3) and 1,763 (SC112-14) SNPs with a minor allele frequency (MAF) higher than 0.01 (Table 7). These SNPs were reduced to 1,358 (IS18760), 1,036 (QL3 and 959 (SC112-14), after filter based on MAF < 0.40 and less than 50% missing (considering heterozygous as missing). The frequency of heterozygosity was 5.55%, 2.75% and 18.21% in RILs populations IS18760, QL3 and SC112-14, respectively.

Table 7. Genome distribution of single nucleotide polymorphism (SNPs) obtained by the genotype-by-sequence analysis of three sets of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred lines (RILs). The RILs derived from anthracnose resistant lines IS18760, QL3 and SC112-14 crossed with the common susceptible line PI609251.

	MAF[1] ≥ 0.01			MAF[1] ≥ 0.4; Missing[2] ≤50%		
Chr.[3]	IS18760	QL3	SC112-14	IS18760	QL3	SC112-14
1	9,316	707	243	142	88	165
2	7,981	669	173	213	119	85
3	8,430	491	197	233	97	91
4	6,620	489	193	154	110	126
5	4,694	480	204	147	138	89
6	5,563	420	178	111	91	96
7	4,272	302	104	57	43	61
8	3,716	371	130	74	125	70
9	4,887	362	149	122	90	89
10	5,349	810	193	105	135	87
	60,828	5,101	1,763	1,358	1,036	959

[1]Minor Allele Frequency, [2] Heterozygous were considered as missing, [3]Chromosome.

4.3 Recombination Linkage Maps

IS18760: The linkage analysis of 1,355 SNPs grouped a total of 1,358 SNPs (i.e. 3 not linked) which were employed to construct a highly saturated linkage map. This map expands 2,629 cM with an average distance of 2.10 cM (Table 8, Supplementary Figure A1 A), and enclose 1,101 recombination events with an average bin size of 0.62Mb. The largest gap was 23.99 cM located in the top distal region of chromosome 10. A subset of 302 high quality SNPs selected based on recombination distance and quality (Supplementary Table A5), produced a map that expand 1,522

cM with an average distance of 5.17 cM. The gap in the top distal region of chromosome 10 increased to 40.49 cM (Table 9, Supplementary Figure A1 B). The map covers approximately 91.97% of the genome.

QL3: The linkage analysis of 1,036 SNPs grouped a total of 1,008 SNPs which were used to construct a highly saturated linkage map. This map expands 3,404.60 cM with an average distance of 3.60 cM (Table 8, Supplementary Figure A2 A), and enclose 881 recombination events with an average bin size of 0.76Mb. The largest gap was 33.55 cM located in chromosome 1. A subset of 221 SNPs (Supplementary Table A5), selected based on recombination distance and quality, generate a map that expands 1,280.03 cM with an average distance of 7.20 cM. The largest gap was of 69.48 cM located in chromosome 3 (Table 9, Supplementary Figure A2 B). The map covers approximately 88.62% of the genome being chromosome 2 the least coverage with 83.73%.

SC112-14: The linkage analysis of 959 SNPs grouped a total of 952 SNPs which were used to construct a highly saturated linkage map. This map expands 1,939.31 cM with an average distance of 2.09 cM (Table 8; Supplementary Figure A3 A), and enclose 743 recombination events with an average bin size of 0.91 Mb. The largest gap was 21.14 cM located in chromosome 8. A subset of 299 markers (Supplementary Table A5), selected based on the recombination distance and quality, generated a linkage map that expands 1,301.22 cM with an average distance of 4.41 cM (Table 9, Supplementary Figure A3 B). The largest gap was 33.8 cM located in chromosome 1. This map covers approximately 91.32 % of the genome.

Table 8. Number of single nucleotide polymorphisms (SNPs) per chromosome, total map distance, and SNPs average for three recombination linkage maps obtained by the analysis of three sets of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred lines (RILs). The RILs derived from anthracnose resistant lines IS18760, QL3 and SC112-14 were crossed with the common susceptible line PI 609251.

	IS18760			QL3			SC112-14		
Chr.[1]	SNP	cM[2]	Avg.[3]	SNP	Length	Avg.	SNP	Length	Avg.
1	142	327.06	2.32	84	329.14	3.97	164	335.85	2.06
2	213	327.31	1.54	116	358.87	3.12	83	182.85	2.23
3	233	376.41	1.62	97	370.85	3.86	91	182.24	2.02
4	154	309.40	2.02	110	332.12	3.05	125	227.80	1.84
5	147	208.89	1.43	137	429.39	3.16	89	161.37	1.83
6	111	270.59	2.46	91	389.23	4.32	96	202.88	2.14
7	57	131.09	2.34	40	219.01	3.31	61	165.01	2.75
8	74	223.26	3.06	117	294.09	2.54	70	156.54	2.27
9	122	218.58	1.81	90	347.48	3.90	88	161.40	1.86
10	101	236.67	2.34	126	334.42	2.68	85	163.37	1.94
	1,355	2,629.30	2.10	1,008	3,404.60	3.60	952	1,939.30	2.09

[1] Chr. refers to chromosome, [2] cM refers to the total recombination distance in centimorgans according to Kosambi, [3] Avg. refers to the average distance between SNPs.

Table 9. Number of single nucleotide polymorphisms (SNPs) per chromosome, total map distance, and average of SNPs for three recombination linkage maps obtained by a subset of SNPs selected based on recombination information for three sets of recombinant inbred lines (RILs) in sorghum [*Sorghum bicolor* (L.) Moench]. The RILs derived from three anthracnose resistant lines (IS18760, QL3, SC112-14) were crossed with the common susceptible line PI 609251.

	IS18760			QL3			SC112-14		
Chr.	SNP	Length	Avg.	SNP	Length	Avg.	SNP	Length	Avg.
1	32	200.36	6.46	21	186.76	9.34	38	197.78	5.10
2	36	174.01	4.97	25	115.22	4.80	26	118.29	4.73
3	31	203.24	6.77	23	163.44	7.43	32	126.85	4.09
4	30	161.69	5.58	24	98.35	4.28	35	148.19	4.36
5	31	131.82	4.35	31	153.22	5.21	35	116.87	3.44
6	29	133.17	4.76	23	128.30	5.83	29	113.66	4.06
7	27	98.88	3.80	8	73.49	10.50	23	124.87	5.68
8	28	138.36	5.12	30	108.72	3.75	28	116.37	4.31
9	30	129.58	4.47	16	144.31	9.62	33	112.16	3.51
10	29	151.42	5.41	29	108.22	11.18	27	126.18	4.85
	303	1,522.50	5.17	230	1,280.00	7.20	306	1,301.20	4.41

[1] Chr. refers to chromosome, [2] cM refers to the total recombination distance in centimorgan according to Kosambi, [3] Avg. refers to the average distance between SNPs.

4.4 Composite interval mapping

The composite interval mapping analysis of the three RIL populations identified 13 QTL for anthracnose resistance distributed in chromosome 1, 3, 4, 5, 6, 7, 8, 9 and 10 (Table 10). Eight QTL were identified in the IS18760 RIL population, while 4 and 1 were detected in the QL3 and SC112-14 RIL populations, respectively.

The eight QTL from IS18760 were located in chromosome 3, 4, 5, 6, 7, 9 and 10 (Supplementary Figure A4). The QTL in chromosome 4 (qPR-TX4-55), 6 (qPR6-16), 7 (qPR7-77) and 9 (qPR9-51) explained 22, 11, 10 and 9% of the phenotypic variation observed in Puerto Rico, while QTL in chromosome 3 (qTX3-58), 4 (qTX4-15 and qPR-TX4-55), 5(qTX5-63) and 10 (qTX10-12) explained 5, 10, 9, 12 and 10% of the phenotypic variation observed in Texas. Additive effects ranged from -0.43 (qPR-TX4-55) to 0.27 (qPR9-51) in Puerto Rico and from -0.21 (qTX10-12) to 0.17 (qTX5-63) in Texas. The QTL qPR-TX4-55 was detected in both

locations with additive effects of -0.43 and -0.17 in Puerto Rico and Texas, respectively. This QTL expands 1.6 Mb (54.09-55.72 Mb) and encloses 183 putative genes.

Four QTL from QL3 were identified in Puerto Rico. These QTL are located in chromosome 1, 4, 5 and 8 (Supplementary Figure A5) and explained 10% (qPR1-73), 25% (qPR4-53), 11% (qPR5-70) and 13% (PR8-53) of the observed phenotypic variation. The largest QTL is located in chromosome 4 (qPR4-53) with an additive effect of -0.42. This QTL overlaps with the QTL (qPR-TX4-55). detected in IS18760 (Figure 4). The analysis of anthracnose resistance response observed in Texas could not detected QTL.

The analysis of SC112-14 detected a single QTL (LOD ≥12) for the four locations (PR, TX, FL, and GA) located at the distal region of chromosome 5 (Figure 5; Supplementary Figure A6). This QTL explained from 57% (TX) to 37% (FL) of the phenotypic variation with an additive effect ranging from -0.51 (TX) to -0.90 (PR) (Figure 5). This 3.1 Mb locus confirmed the *Cs-SC112* locus (Cuevas et al. 2014), and overlap with the QTL detected in IS18760 (qTX5-63) which have positive additive effect (0.17) in Texas.

Table 10. Quantitative trait loci (QTL) associated with anthracnose (*Colletotrichum sublineola)* disease reactions in three sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred line (RIL) population. The RILs derived from anthracnose resistant lines IS18760, QL3 and SC112-14 were crossed with the common susceptible line PI609251, and evaluated at Puerto Rico (PR), Florida (FL), Georgia (GA) and Texas (TX) during 2016.

RILs	Location	Chr.[1]	QTL	Linked SNP[2]	Interval (cM)[3]	Genomic Region (Mb)[4]	LOD[5]		R2[6]	Add.[7]
IS18760	PR	4	qPR-TX4-55	Sb004g_55716346	6.78	54.09-55.72	9.10	***	0.22	-0.43
	PR	6	qPR6-16	Sb005g_1683577	9.52	0.38-1.92	5.27	***	0.11	-0.29
	PR	7	qPR7-77	Sb007g_777177	10.9	0.78-7.55	4.86	***	0.10	-0.35
	PR	9	qPR9-51	Sb009g_51062242	4.46	50.73-51.06	3.52	*	0.09	0.27
	TX	3	qTX3-58	Sb003g_58468765	0.57	58.47-60.20	3.28	*	0.05	0.12
	TX	4	qTX4-15	Sb004g_15341110	7.53	10.91-33.67	5.47	***	0.10	-0.17
	TX	4	qPR-TX4-55	Sb004g_55716346	16.78	54.09-55.72	5.12	**	0.09	-0.17
	TX	5	qTX5-63	Sb005g_63684995	12.57	62.91-65.53	6.02	***	0.12	0.17
	TX	10	qTX10-12	Sb010g_12810290	9.84	11.00-12.94	5.41	***	0.10	-0.21
QL3	PR	1	qPR1-73	Sb001g_ 73355610	16.71	59.46-70.16	3.68	*	0.10	-0.26
	PR	4	qPR4-53	Sb004g_ 53280964	24.18	53.28-57.70	6.83	***	0.25	-0.42
	PR	5	qPR5-70	Sb005g_ 70802720	1.46	70.28-70.80	3.60	*	0.11	0.71
	PR	8	qPR8-53	Sb008g_ 53829072	11.00	6.38-56.72	4.68	**	0.13	0.36
SC112-14	PR	5	qPR-TX-FL-GA5-64	Sb005g_ 64840542	23.16	62.51-65.52	15.87	***	0.51	-0.90
	TX	5	qPR-TX-FL-GA5-64	Sb005g_ 64840542	14.58	62.51-64.84	17.84	***	0.57	-0.51
	FL	5	qPR-TX-FL-GA5-64	Sb005g_ 64829235	19.66	62.51-65.52	12.28	***	0.37	-0.75
	GA	5	qPR-TX-FL-GA5-64	Sb005g_ 64829235	23.16	62.51-65.52	21.64	***	0.52	-0.86

[1]Chromosome, [2] Single nucleotide polymorphisms, [3] centimorgan distance according to Kosambi, [4]Mega base pairs, [5]Logarithm of the odds, [6] Phenotypic variability explained by the QTL, [7] Additive effect of the QTL whereas negative value indicates the allele is derived from the resistant lines IS18760, QL3 or SC112-14.

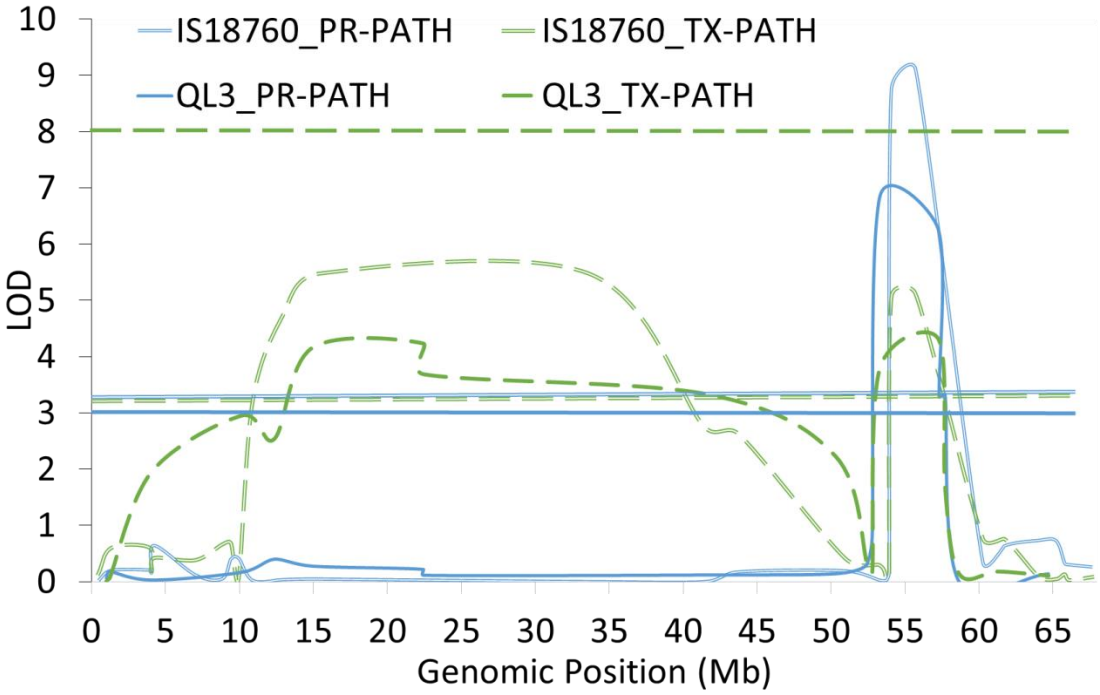

Figure 4. Composite interval mapping in chromosome 4 for anthracnose (*Colletotrichum sublineola*) disease reactions in sorghum [*Sorghum bicolor* (L.) Moench] using two recombinant inbred lines (RILs) populations derived from the resistant lines IS18760 and QL3 crossed with the susceptible line PI609251. The RILs were evaluated for anthracnose resistant response in Texas and Puerto Rico during the summer and fall of 2016, respectively.

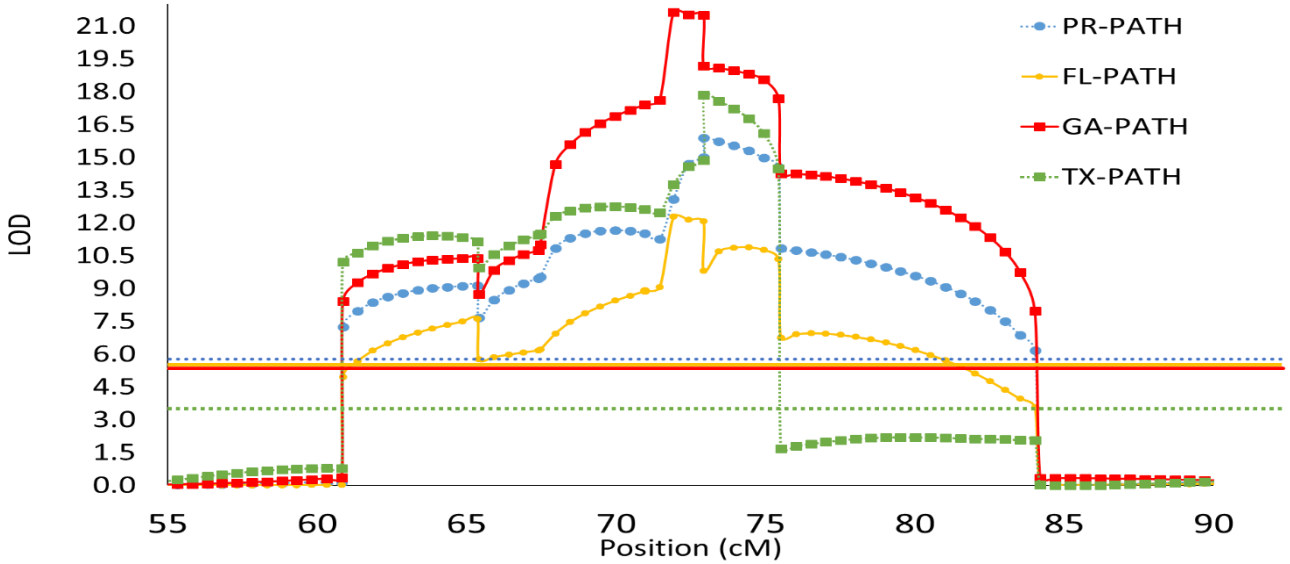

Figure 5. Composite interval mapping in chromosome 5 for anthracnose (*Colletotrichum sublineola*) resistant response in sorghum [*Sorghum bicolor* (L.) Moench] using a recombinant inbred line (RILs) population derived from the resistant line SC112-14 crossed with the susceptible line PI609251. The RILs were evaluated for anthracnose resistant response in Puerto Rico (PR), Texas (TX), Florida (FL) and Georgia (GA) during the summer and fall of 2016, respectively.

4.5 Identification of candidate genes in Cs-SC112 locus

The genotyping analysis of 1,500 $F_{2:3}$ progenies with the markers SNP_Sb005g55256930 and SNP_Sb005g56057061 identified 221 and 250 plants homozygous for SC112-14 and PI609251 alleles, respectively, and 576 heterozygous. A total of 116 progenies (7%) were putative recombinants of which 55 and 61 had the PI609251 and SC112-14 allele, respectively. Recombinants with PI609251 allele were evaluated with KASP markers SNP_Sb005g022920 and SNP_Sb005g56057061 to confirm recombination events and select plants for field evaluation. A total of seven recombinants were confirmed, of which 3 were resistant ($F_{2:3}$-#15 176, 467, and 649; score <2.0), three moderate resistance ($F_{2:3}$-#15 105, 667 and 1040; score > 2 and < 3) and one susceptible ($F_{2:3}$-#15 375; score > 3) (Table 11).

Table 11. Genotyping and anthracnose (*Colletotrichum sublineola*) resistant response of seven sorghum [*Sorghum bicolor* (L.) Moench] recombinants within resistant locus *Cs-SC112* at chromosome 5. Recombinants identified in $F_{2:3}$ families derived from the cross of anthracnose resistant line SC112-14 with suscetible line PI609251 and evaluated in Puerto Rico during the fall of 2016

Family	ID	M1[1]	M2[2]	M3[3]	M4[4]	Anthracnose Response[5]
F2:3- #15	105	H[6]	H	H	B	3
F2:3- #15	176	H	-	H	B	2
F2:3- #15	375	H	B	B	B	3.5
F2:3- #15	467	B	H	H	H	2
F2:3- #15	649	H	H	H	B	2
F2:3- #15	667	B	B	B	H	3
F2:3- #15	1040	B	B	B	H	3

[1] KASP SNP_Sb005g022920; [2] Insertion/Deletion in_Sb005g172100;

[3] KASP SNP_Sb005g172300; [4] KASP SNP_Sb005g56057061; [5] Anthracnose response according to Prom et al., (2009), where 1-2 are resistant, 2-3 moderate resistant, and 3-5 susceptible; [6] H, B and - refers to heterozygous, homozygous for PI609251, and missing, respectively

4.6 Delimitation of the Cs-SC112

The seven recombinant individuals were genotyped with three additional markers to delimit the recombination events within *Cs-SC112* locus. Two markers (INDEL_Sb005g172100 and SNP_Sb005g172300) were polymorphic and divided the locus into 3 genomic segments of 386, 23 and 446 Kb, respectively. Both markers co-segregated perfectly with anthracnose resistant

response indicating resistant gene should be located within or in proximity to the 23 Kb genomic interval.

4.7 Candidate gene analysis

The 23 Kb region encloses three annotated genes: Sobic.005g172100, Sobic.005g172200 and Sobic.005G172300 [Phytozome 12; (https://phytozome.jgi.doe.gov)] that characterize by the presence of protein domains Ankyrin repeats, DUF295 and F-box, respectively, which have been previously associated with disease resistance. Due to the absence of markers near to this 23 Kb region, we expanded the candidate gene analysis to a 100 Kb window (i.e. ~40 Kb down and upstream of the 23 Kb region) (Figure 6). Eight additional annotated genes with protein domains previously associated with disease resistance such as PAP-Fibrillin, WD40, G-beta and ubiquitin-protein ligase were identified. To discern among these eleven candidate genes, they were compared to the Plant Resistance Gene Database (https://prgdb.crg.eu) where Sobic.005g172300 showed similarity with to a putative Ser-Thr kinase resistant gene from barley (*Hordeum vulgare*). Partial resequencing analysis (~59%) of Sobic.005G172300 in parental lines SC112-14 and PI609251, and the susceptible reference genome of BTx623 detected four SNPs at the 76, 478, 584 and 905 protein coding position. Three SNPs were not synonymous encoding 2 proteins variants that divided both resistant and susceptible genotypes (Table 12).

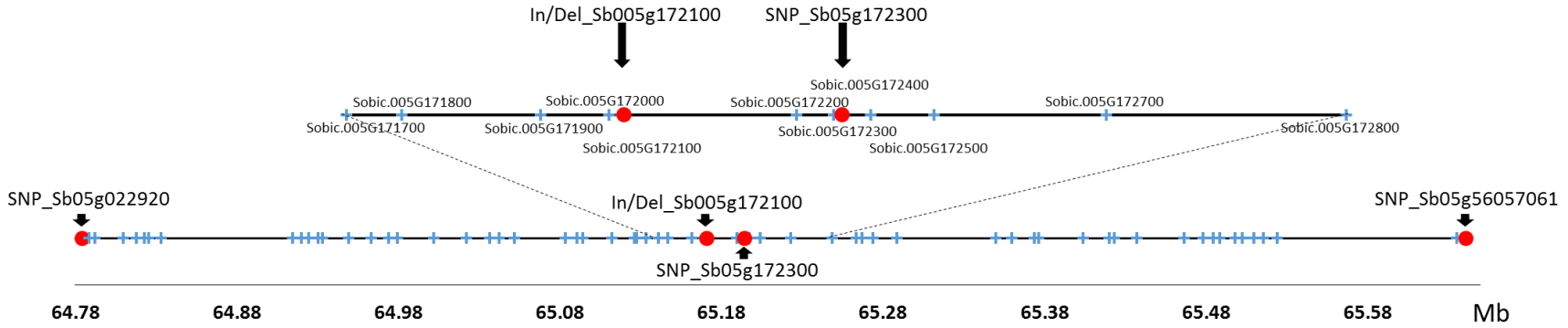

Figure 6. Fine mapping of anthracnose (*Colletotrichum sublineola*) resistance locus *Cs-SC112* in sorghum [*Sorghum bicolor* (L.) Moench] chromosome 5. The bottom panel represent the *Cs-SC112* genomic region, the position of Kompetitive Allele Specific PCR (KASP) and insertion/deletion markers (red dots) used to delimited recombination events, and the genes enclosed within the locus (blue crosses). The top panel represents the delimited 23 Kb region and ~40 Kb down and upstream (i.e. 100 Kb region) that enclose 11 candidate genes.

Table 12. Genomic and protein sequence analysis of sorghum [*Sorghum bicolor* (L.) Moench] gene Sobic.005G172300.1 putative associated with anthracnose (*Colletotrichum sublineola*) resistant response. Resequencing analysis based on resistant line SC112-14, and susceptible lines PI609251 and BTx623.

Anthracnose Response	Line	Coding Position			
		376[1]	478[1]	584[1]	905[1]
Susceptible	BTx 623	T	C	C	A
Susceptible	PI609251	T	C	C	A
Resistant	SC112-14	C	G	T	G
		Amino Acid[2]			
Susceptible Protein Variant		R	C	R	T
Resistant Protein Variant		R	W	C	A

[1] Single nucleotide polymorphism (SNP) position based on cDNA sequence of Sobic.005G172300.1 according to annotation in Phytozome 12; [2] Amino acid encoded by the codon containing the SNP

5 DISCUSSION

Anthracnose is one of the most detrimental diseases for sorghum production in tropical and subtropical regions. The utilization of multiple resistance sources is projected to be the most reliable alternative to control the disease. Therefore, the genetic basis of three sources of anthracnose resistance from Sudan, India and Ethiopia was studied across four locations (Puerto Rico, Texas, Georgia and Florida) to provide the information needed to produce genomics tools for use in sorghum breeding programs.

5.1 Quantitative Trait Loci

The identification of germplasm providing a broad range of resistance is necessary for the development of new sorghum varieties and hybrids. The line SC112-14 had resistant responses in Puerto Rico, Texas, Florida and Georgia, while a previous study based on greenhouse evaluation showed resistant against particular pathotypes from Texas, Georgia, Arkansas and Puerto Rico (Cuevas et al., 2014). However, susceptible responses have also observed in greenhouse evaluations against pathotypes from College Station (AMP 157) and Wharton (AMP 234), Texas (Prom et al., 2012). Differences in the resistant response among greenhouse and field evaluations could be attributed to differences in irrigation (i.e. humidity), population diversity and pressure of pathogen. In the experiment in Texas, field inoculations were performed in only a few plants per plot hence, there is the possibility that frequencies of the race AMP157 were too low to observe a susceptible response. Certainly, the broader resistant response of SC112-14 make it as important source for breeding programs together with resistance lines SC748-5 (Burrell et al., 2015) and SC155-14E (Patil et al., 2017). The resistant locus present in SC748-5 and SC155-14E, were mapped in chromosome 5 and 9, respectively, being different from the locus of SC112-14. Therefore, these three resistance sources should be pyramided into elite sorghum varieties to provide resistance against multiple pathotypes. Certainly, the development of a single variety with high genetic combining ability (e.g. RTx430) and multiple resistant genes will be a valuable germplasm tool for national sorghum breeding programs, which can then be strategically used to increase resistant durability.

Comparative genomic analysis among inheritance studies could be used to confirm location of specific QTL. The QTL in chromosome 4 contributing to anthracnose resistance

response in IS18760 and QL3 overlap with the previously identified locus in lines SC414-12E and SC155-14E (Patil et al., 2017). This shared genomic region expands 1.63 Mb (54.09-55.72 Mb) enclosing 183 putative genes not associated with R gene families. Likewise, the QTL in chromosome 6 associated with the resistance present in IS18760 overlaps with the QTL identified in the line 296B (Mohan et al., 2010). This shared genomic region expands 1.54 Mb and enclose over 100 genes. Reducing the large size of both genomic regions is necessary to validate these QTL and identify candidate resistance genes. The development of different sets of near isogenic lines (NILs) could be the most adequate approach to dissect these QTLs given that the background noise is diminished allowing to precisely study the physiological role of the resistant response. In fact in wheat, NILs were used to dissect seed dormancy QTL into two candidate genes (Barrero et al., 2015). The genetic markers developed and identified in this study provide the necessary tools for the development of NILs. One alternative approach that could accelerate the NILs development is derived them from RILs that are heterozygous for a particular QTL. For instance, the IS18760 RILs population have three lines (IS18760-RIL-007, IS18760-RIL-027 and IS18760-RIL-046) heterozygous for the qPR-TX4-55 QTL and homozygous for >85% of the genome. Therefore, the progenies derived from the self-pollination of these lines will be segregating for the QTL, and could lead to the development of NILs in few generations of marker assisted selection while avoiding the traditional multiple backcrossing system.

5.2 Candidate genes

Resistance genes (R-genes) most commonly encode intracellular proteins containing nucleotide binding leucine rich repeat domains (NB-LRR) or extracellular receptor with LRR domains. In the current gene-for-gene hypothesis, these domains recognize the pathogen elicitors triggering an immune response that stop the disease progression (Bent & Mackey, 2007). The genes present in *Cs-SC112* locus delimited here do not belong to an R gene family. The best candidate gene is characterized by a F-box and Ser-Thr kinase domains. The definition of R-genes has been lately modified due to the identification and characterization of numerous resistance genes with other functional domains. The F-Box domain is a protein-protein interactor involved in protein ubiquitination and degradation that control crucial processes such as embryogenesis, hormonal response, senescence and pathogen resistance (Xu et al., 2009). For instance, the gene SON1 in *Arabidopsis thaliana* has a F-box domain and play an important role in ubiquitination of

proteins for subsequent degradation resulting in an increase systemic resistance to *Pseudomonas syringae* pv. *tomato* (Delaney, 2002). Today, many plant resistance genes are characterized by the presence of intracellular protein kinase that recognize pathogen avirulence proteins (*Avr*) and induced plant signaling cascades (Bent & Mackey, 2007). Certainly, the results suggest that *Cs-SC112* locus interact with pathogens avirulence proteins (*Avr* genes) leading to the activation of the plant immune signaling cascade. Transcriptional analysis at different stages of plant development under anthracnose infection could give insights on the resistance response mechanism. Likewise, complementation by transgenic approaches will be useful to validate the candidate gene.

5.3 Breeding anthracnose resistance

The introgression of anthracnose resistance genes into elite germplasm is determined by the source of resistant and the targeted region of sorghum production. The high heritability and monogenic inheritance of SC112-14 suggests that introgression based on backcrossing at a single environment is appropriated to achieve new varieties with broader resistant response. Definitely, the two nearest markers to *Cs-SC112* locus (INDEL_Sb005g172100 and SNP_Sb005g172300) provide the genomic tools necessary to accelerate its introgression with limited field phenotypic evaluation. In contrast, the introgression of resistance from IS18760 and QL3 into elite germplasm requires multi-location experiments that account for its low heritability and high GxL interaction. Marker assisted recurrent selection (MARS) (Ribaut & Delannay, 2010) leading to the accumulation of resistance genes from different chromosomes is a strategy that could be effective owing to this complex inheritance. For instance, one large F_2 population (>300 individuals) derived from these lines (i.e. IS18760 or QL3) is genotyped to select the individuals with the highest recombination events and anthracnose resistance QTL. Selected individuals are advanced until the generation of $F_{2:4}$ families/lines based on single seed descent (SSD) method. These $F_{2:4}$ families/lines are evaluated in multi-location trials, and superior families/lines are selected based on agronomical performance and resistance responses. The selections are strategically intercrossed to undergo a new recombination cycle, and the process is repeated up to the identification of families/lines with durable resistance and superior agronomical performance. This selection approach has been successful for the introgression of disease resistance for rice blast

(*Magnaporthe oryzae*) and powdery mildew (*Blumeria graminis f. sp. tritici, Bgt*) in rice (Miah et al., 2007) and wheat (Xu et al., 2017) cultivars, respectively.

6 CONCLUSION

Anthracnose resistance in sorghum is a complex trait where its expression and inheritance is affected by multiple genetic and environmental factors. The resistant response observed in SC112-14 is controlled by one locus (*Cs-SC112*) on chromosome five that provides resistance against pathotypes from Puerto Rico, Texas, Florida and Georgia. In contrast, the resistance response observed in IS18760 and QL3 is controlled by multiple genes that provide resistance against pathotypes from Puerto Rico and Texas. Despite the high number of QTL identified in both populations, the common QTL present in chromosome four indicate this genomic region enclose important anthracnose resistance genes. The delimitation of *Cs-SC112* locus to a 23kb region with three candidate genes showed that anthracnose resistant response is instead controlled by genes involved in signaling cascades and transcriptional reprograming rather than recognition of pathotype-associated molecular patterns. Molecular markers developed herein are effective for marker assisted selection directed to increase resistant durability.

.

7 LITERATURE CITED

Ali, M. E. K., & Warren, H. L. (1987). Physiological Races of *Colletotrichum graminicola* on Sorghum. *Plant Disease*, *71*(5), 402–404. doi:10.1094/PD-71-0402

Awika, J. M., & Rooney, L. W. (2004). Sorghum phytochemicals and their potential impact on human health. *Phytochemistry*, *65*(9), 1199–1221. http://doi.org/10.1016/j.phytochem.2004.04.001

Barrero, J. M., Cavanagh, C., Verbyla, K. L., Tibbits, J. F. G., Verbyla, A. P., Huang, B. E., Rosewarne, G. M., Stephen, S., Wang, P., Whan, A., Rigault, P., Hayden, M. J., Gubler, F. (2015). Transcriptomic analysis of wheat near-isogenic lines identifies PM19-A1 and A2 as candidates for a major dormancy QTL. *Genome Biology*, *16*(1), 93. http://doi.org/10.1186/s13059-015-0665-6

Basten, C. J., Weir, B. S., & Zeng, Z.-B. (2004). QTL cartographer, Version 1.17. *Department of Statistics, North Carolina State University, Raleigh, NC.* Retrieved from http://statgen.ncsu.edu/qtlcart/WQTLCart.htm

Bent, A. F., & Mackey, D. (2007). Elicitors, Effectors, and *R* Genes: The New Paradigm and a Lifetime Supply of Questions. *Annual Review of Phytopathology*, *45*(1), 399–436. http://doi.org/10.1146/annurev.phyto.45.062806.094427

Biruma, M., Martin, T., Fridborg, I., Okori, P., & Dixelius, C. (2012). Two loci in sorghum with NB-LRR encoding genes confer resistance to *Colletotrichum sublineolum*. *Theoretical and Applied Genetics*, *124*(6), 1005–1015. http://doi.org/10.1007/s00122-011-1764-8.

Boora, K. S., Frederiksen, R. A., & Magill, C. W. (1997). DNA-Based Markers for a Recessive Gene Conferring Anthracnose Resistance in Sorghum. *Crop Science*, *38*(6), 1708–1709.

Burrell, M., Sharma, A., Patil, N. Y., Collins, S. D., Anderson, W. F., Rooney, W. L., & Klein, P. E. (2015). Sequencing of an anthracnose-resistant sorghum genotype and mapping of a major QTL reveal strong candidate genes for anthracnose resistance. *Crop Science*, *55*(2), 790–799. http://doi.org/10.2135/cropsci2014.06.0430

Cardwell, K., Hepperly, P., & Frederiksen, R. A. (1989). Pathotypes of *Colletotrichum graminicola* and Seed Transmission of Sorghum Anthracnose. *Plant Disease*, *73*(3), 255–257.

Casela, C. R., & Frederiksen, R. A. (1993). Survival of *Colletotrichum graminicola* Sclerotia in Sorghum Stalk Residues. *Plant Disease*, *77*(88), 825–827.

Coleman, O. H., & Stokes, I. E. (1954). The inheritance of resistance to stalk red rot in sorghum. *Agronomy Journal*, *46*(2), 61–63.

Crill, J. P., & Khush, G. S. (1982). Effective and stable control of rice blast with monogenic resistance. *Evolution of the Gene Rotation Concept for Rice Blast Control*, *933*, 87–102.

Cuevas, H. E., & Prom, L. K. (2013). Assessment of molecular diversity and population structure of the Ethiopian sorghum [*Sorghum bicolor* (L.) Moench] germplasm collection maintained by the USDA-ARS National Plant Germplasm System using SSR markers. *Genetic Resources and Crop Evolution*, *60*, 1817–1830. http://doi.org/10.1007/s10722-013-9956-5

Cuevas, H. E., Prom, L. K., & Erpelding, J. E. (2014). Inheritance and molecular mapping of anthracnose resistance genes present in sorghum line SC112-14. *Molecular Breeding*, *34*, 1943–1953. http://doi.org/10.1007/s11032-014-0151-y

Dahlberg, J. A. (2000). Classification and Characterization of Sorghum. In *Sorghum Origin, History, Technology, and Production* (pp. 99–130).

Dahlberg, J. A., & Wasylikowa, K. (1996). Image and statistical analyses of early sorghum remains (8000 B.P.) from the Nabta Playa archaeological site in the Western Desert, southern Egypt. *Vegetation Histrory and Archaeobotany*, *5*, 293–299.

Dahlberg, J., Berenji, J., Sikora, V., & Dragana, L. (2011). Assessing sorghum [*Sorghum bicolor* (L) Moench] germplasm for new traits: Food, fuels & unique uses. *Maydica*, *56*(1750), 85–92.

Dial, H. I. (2012). Plant guide for sorghum (Sorghum bicolor L.). *USDA-Natural Resources Conservation Service*, 5. Retrieved from http://www.nrcs.usda.gov/

Dicko, M. H., Gruppen, H., Traoré, A. S., Voragen, A. G. J., & Van Berkel, W. J. H. (2006). Sorghum grain as human food in Africa: relevance of content of starch and amylase activities. *African Journal of Biotechnology*, *5*(5), 384–395. Retrieved from http://www.academicjournals.org/AJB

Dillon, S. L., Shapter, F. M., Henry, R. J., Cordeiro, G., Izquierdo, L., & Lee, L. S. (2007). Domestication to crop improvement: Genetic resources for Sorghum and Saccharum (Andropogoneae). *Annals of Botany*, *100*, 975–989. http://doi.org/10.1093/aob/mcm192

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), 1–10. http://doi.org/10.1371/journal.pone.0019379

Erpelding, J. E., & Prom, L. K. (2004). Evaluation of Malian Sorghum Germplasm for Resistance Against Anthracnose. *Plant Pathology Journal*, *3*(2), 65–71. http://doi.org/10.3923/ppj.2004.65.71

Erpelding, J. E., & Prom, L. K. (2006). Variation for Anthracnose Resistance within the Sorghum Germplasm Collection from Mozambique, Africa. *Plant Pathology Journal*, *5*(1), 28–34.

Felderhoff, T. J., McIntyre, L. M., Saballos, A., & Vermerris, W. (2016). Using Genotyping by Sequencing to Map Two Novel Anthracnose Resistance Loci in Sorghum bicolor. *G3: Genes Genomes G*, *6*(7), 1935–1946. http://doi.org/10.1534/g3.116.030510

Food and Agriculture Organization. (2015). FAOSTAT Statistics Database.

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*, *9*(2), 1–11. http://doi.org/10.1371/journal.pone.0090346

Guillemaut, P., & Maréchal-Drouard, L. (1992). Isolation of plant DNA: A fast, inexpensive, and reliable method. *Plant Molecular Biology Reporter*, *10*(1), 60–65. http://doi.org/10.1007/BF02669265

Hagerty, M. (1941). Comments on Writings Concerning Chinese Sorghum. *Harvard Journal of Asiatic Studies*, 234–263.

Harlan, J. R., & de Wet, J. M. J. (1972). A Simplified Classification of Cultivated Sorghum. *Crop Science*, *12*(2), 172–176.

Haussmann, B., Mahalakshmi, V., Reddy, B., Seetharama, N., Hash, C., & Geiger, H. (2002). QTL mapping of stay-green in two sorghum recombinant inbred populations. *Theoretical and Applied Genetics*, *106*, 133–142.

Jansen, R. C., & Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics*, *136*(4), 1447–1455.

Kim, H. S., & Delaney, T. P. (2002). Arabidopsis SON1 is an F-box protein that regulates a novel induced defense response independent of both salicylic acid and systemic acquired resistance. *The Plant Cell*, *14*(7), 1469–1482. http://doi.org/10.1105/tpc.001867

Klein, R. R., Rodriguez-Herrera, R., Schlueter, J. A., Klein, P. E., Yu, Z. H., & Rooney, W. L. (2001). Identification of genomic regions that affect grain-mould incidence and other traits of agronomic importance in sorghum. *Theoretical and Applied Genetics*, *102*(2–3), 307–319. http://doi.org/10.1007/s001220051647

Kosambo-Ayoo, L. M., Bader, M., Loerz, H., & Becker, D. (2011). Transgenic sorghum (*Sorghum bicolor* L. Moench) developed by transformation with chitinase and chitosanase genes from Trichoderma harzianum expresses tolerance to anthracnose. *African Journal of Biotechnology*, *10*(19), 3659–3670. http://doi.org/10.5897/AJB10.1530

Lebeau, F. J. (1951). *Anthracnose and Red Rot of Sorghum*. Washington, D.C.

Margarido, G. R. A., Souza, A. P., & Garcia, A. A. F. (2007). OneMap: Software for genetic mapping in outcrossing species. *Hereditas*, *144*(3), 78–79. http://doi.org/10.1111/j.2007.0018-0661.02000.x

Marley, P. S., Thakur, R. P., & Ajayi, O. (2001). Variation among foliar isolates of Colletotrichum sublineolum of sorghum in Nigeria. *Field Crops Research*, *69*, 133–142. http://doi.org/10.1016/S0378-4290(00)00128-3

Mehta, P. J., Wiltse, C. C., Rooney, W. L., Collins, S. D., Frederiksen, R. A., Hess, D. E., … Tebeest, D. O. (2005). Classification and inheritance of genetic resistance to anthracnose in sorghum. *Field Crops Research*, *93*, 1–9. http://doi.org/10.1016/j.fcr.2004.09.001

Miah, G., Rafii, M. Y., Ismail, M. R., Puteh, A. B., Rahim, H. A., & Latif, M. A. (2007). Marker-assisted introgression of broad-spectrum blast resistance genes into the cultivated MR219 rice variety Gous. *Journal of the Science of Food and Agriculture*, *97*(9), 2810–2818. http://doi.org/10.1002/j

Moore, J., Ditmore, M., & TeBeest, D. (2008). Pathotypes of Colletotrichum sublineolum in Arkansas. *Plant Disease*, *92*(10), 1415–1420. http://doi.org/10.1094/PDIS-92-10-1415

Murali Mohan, S., Madhusudhana, R., Mathur, K., Chakravarthi, D. V. N., Rathore, S., Nagaraja Reddy, R., Satish, K., Srinivas, G., Sarada M. N., Seetharama, N. (2010). Identification of quantitative trait loci associated with resistance to foliar diseases in sorghum [Sorghum bicolor (L.) Moench]. *Euphytica*, *176*(2), 199–211. http://doi.org/10.1007/s10681-010-0224-x

Nave, M., Avni, R., Ben-Zvi, B., Hale, I., & Distelfeld, A. (2016). QTLs for uniform grain dimensions and germination selected during wheat domestication are co-located on

chromosome 4B. *Theoretical and Applied Genetics*, *129*(7), 1303–1315. http://doi.org/10.1007/s00122-016-2704-4

Néya, A., & Le Normand, M. (1998). Responses of sorghum genotypes to leaf anthracnose (Colletotrichum graminicola) under field conditions in Burkina Faso. *Crop Protection*, *17*(1), 47–53. http://doi.org/10.1016/S0261-2194(98)80012-4

Pande, S., Mughogho, L. K., Bandyopadhyay, R., & Karunakar, R. I. (1991). Variation in Pathogenicity and Cultural Characteristics of Sorghum Isolates of Colletotrichum graminicola in India. *Plant Disease*, *75*(8), 778–783.

Pande, S., Thakur, R., Karunakar, R., Bandyopadhyay, R., & Reddy, B. (1994). Development of screening methods and identification of stable resistance to anthracnose in sorghum. *Field Crops Research*, *38*, 157–166. http://doi.org/10.1016/0378-4290(94)90087-6

Pastor-Corrales, M. A., & Frederiksen, R. A. (1980). Sorghum anthracnose. In Sorghum diseases: a world review. In *Proceedings of International Workshop on Sorghum Diseases* (pp. 289–291).

Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H., Wang, X., Wicker, T., Bharti, A. K., Chapman, J., Feltus, F. A., Gowik, U., Grigoriev, I. V., Lyons, E., Maher, C. A., Martis, M., Narechania, A., Otillar, R. P., Penning, B. W., Salamov, A. A., Wang, Y., Zhang, L., Carpita, N. C., Freeling, M., Gingle, A. R., Hash, C. T., Keller, B., Klein, P., Kresovich, S., McCann, M. C., Ming, R., Peterson, D. G., Mehboob-ur-Rahman, Ware, D., Westhoff, P., Mayer, K. F. X., Messing, J., Rokhsar, D. S. (2009). The Sorghum bicolor genome and the diversification of grasses. *Nature*, *457*(7229), 551–556. http://doi.org/10.1038/nature07723

Patil, N. Y., Klein, R. R., Williams, C. L., Collins, S. D., Knoll, J. E., Burrell, A. M., Anderson, W. F., Rooney, W. L., Klein, P. E. (2017). Quantitative trait loci associated with anthracnose resistance in sorghum. *Crop Science*, *57*(2), 877–890. http://doi.org/10.2135/cropsci2016.09.0793

Perfect, S. E., Hughes, H. B., O 'connell, R. J., & Green, J. R. (1999). Colletotrichum: A Model Genus for Studies on Pathology and Fungal–Plant Interactions. *Fungal Genetics and Biology*, *27*, 186–198.

Perumal, R., Menz, M. A., Mehta, P. J., Katile, S., Gutierrez-Rojas, L. A., Klein, R. R., Klein, P. E., Prom, L. K., Schlueter, J. A., Rooney, W. L., Magill, C. W. (2009). Molecular mapping of Cg1, a gene for resistance to anthracnose (*Colletotrichum sublineolum*) in sorghum. *Euphytica*, *165*, 597–606. http://doi.org/10.1007/s10681-008-9791-5

Prom, L. K., Erpelding, J., Perumal, R., Isakeit, T., & Cuevas, H. (2012). Response of Sorghum Accessions from Four African Countries against Colletotrichum sublineolum, Causal Agent of Sorghum Anthracnose. *American Journal of Plant Sciences*, *3*, 125–129. http://doi.org/10.4236/ajps.2012.31014

Prom, L. K., Perumal, R., Erattaimuthu, S. R., Little, C. R., No, E. G., Erpelding, J. E., Rooney, W. L., Odvody, G. N., Magill, C. W. (2012). Genetic diversity and pathotype determination of Colletotrichum sublineolum isolates causing anthracnose in sorghum. *European Journal of Plant Pathology*, *133*, 671–685. http://doi.org/10.1007/s10658-012-9946-z

Prom, L. K., Perumal, R., Erpelding, J., Isakeit, T., Montes-Garcia, N., & Magill, C. W. (2009). A Pictorial Technique for Mass Screening of Sorghum Germplasm for Anthracnose (Colletotrichum sublineolum) Resistance. *The Open Agriculture Journal*, *3*, 20–25.

Reddy, B. V. S., Ramesh, S., & Reddy, P. S. (2005). Sweet Sorghum – A potential alternate raw material forbBio-ethanol and bio- energy. *SAT eJournal*, *1*(1), 1–8.

Resende, R. S., Rodrigues, F. Á., Costa, R. V., & Silva, D. D. (2013). Silicon and Fungicide Effects on Anthracnose in Moderately Resistant and Susceptible Sorghum Lines. *Journal of Phytopathology*, *161*(1), 11–17. http://doi.org/10.1111/jph.12020

Ribaut, J.-M., de Vicente, M., & Delannay, X. (2010). Molecular breeding in developing countries: challenges and perspectives. *Current Opinion in Plant Biology*, *13*(2), 213–218. http://doi.org/10.1016/j.pbi.2009.12.011

Roberts, A., McMillan, L., Wang, W., Parker, J., Rusyn, I., & Threadgill, D. (2007). Inferring missing genotypes in large SNP panels using fast nearest-neighbor searches over sliding windows. *Bioinformatics*, *23*(13), 401–i407.

Rutto, L. K., Xu, Y., Brandt, M., Ren, S., & Kering, M. K. (2013). Juice, Ethanol, and Grain Yield Potential of Five Sweet Sorghum (Sorghum bicolor [L.] Moench) Cultivars. *Journal of Sustainable Bioenergy Systems*, *3*, 113–118. http://doi.org/10.4236/jsbs.2013.32016

Silva, D. D., Costa, R. V., Cota, L. V., Figueiredo, J. E. F., Casela, C. R., & Lanza, F. E. (2015). Genotype rotation for leaf anthracnose disease management insorghum. *Crop Protection*, *67*, 145–150. http://doi.org/10.1016/j.cropro.2014.10.007

Singh, P. (2006). Using an action research framework to explore assessment: A South African case study. *Systemic Practice and Action Research*, *19*(2), 179–187. http://doi.org/10.1007/s10681-005-9065-4

Smith, C. W., & Frederiksen, R. A. (2001). Sorghum. Origin, History, Technology, and Production. *Wiley Series in Crop Science*.

Snowden, J. D. (1936). The cultivated races of sorghum, pp. 224

Thakur, J. S., Prinja, S., Garg, C., Mendis, S., & Menabde, N. (2011). Social and Economic Implications of Noncommunicable diseases in India. *Indian Journal of Community Medicine*, *36*(Suppl 1), S13-22. http://doi.org/10.4103/0970

Upadhyaya, H. D., Wang, Y.-H., Sharma, R., & Sharma, S. (2013). Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *126*(6), 1649–57. http://doi.org/10.1007/s00122-013-2081-1

Valèrio, H. M., Rèsende, M. A., Weikert-Oliveira, R. C. B., & Casela, C. R. (2005). Virulence and molecular diversity in Colletotrichum graminicola from Brazil. *Mycopathologia*, (159), 449–459. http://doi.org/10.1007/s11046-005-0373-y

Voorrips, R. E. (2002). MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *The Journal of Heredity*, *93*(1), 77–78.

Wang, Y. H., Upadhyaya, H. D., & Kole, C. (2014). Sorghum production for diversified uses.

Wendorf, F., Close, A. E., Schild, R., Wasylikowa, K., Housley, R. A., Harlan, J. R., & Krolik, H.

(1992). Saharan exploitation of plants 8,000 years BP. *Nature*, *359*, 722–724.

Wharton, P. S., Julian, A. M., & O'Connell, R. J. (2001). Ultrastructure of the Infection of *Sorghum bicolor* by *Colletotrichum sublineolum*. *Phytopathology*, *91*(2), 149–158. http://doi.org/10.1094/PHYTO.2001.91.2.149

Xu, G., Ma, H., Nei, M., & Kong, H. (2009). Evolution of F-box genes in plants: Different modes of sequence divergence and their relationships with functional diversification. *Proceedings of the National Academy of Sciences*, *106*(3), 835–840. http://doi.org/10.1073/pnas.0812043106

Xu, H., Cao, Y., Xu, Y., Ma, P., Ma, F., Song, L., Li, L., An, D. (2017). Marker-Assisted Development and Evaluation of Near-Isogenic Lines for Broad-Spectrum Powdery Mildew Resistance Gene Pm2b Introgressed into Different Genetic Backgrounds of Wheat. *Frontiers in Plant Science*, *8*(July), 1–9. http://doi.org/10.3389/fpls.2017.01322

Zeng, Z.-B. (1994). Precision Mapping of Quantitative Trait Loci. *Genetics*, *136*, 1457–1468.

8 APPENDIX

Table A1. Polymerase Chain Reaction (PCR) cycle employed with the Kompetitive Allele Specific PCR (KASP) markers used to genotype 1,500 $F_{2:3}$ sorghum (*Sorghum bicolor* (L.) Moench) progenies derived from the cross of the anthracnose (*Colletotrichum sublineola*) resistant line SC112-14 and the susceptible line PI609251.

SNP_Sb005g022920			**SNP_Sb005g56057061**		
Temperature	Time	Repetition	Temperature	Time	Repetition
94^0C	15:00	1	94^0C	15:00	1
94^0C	0:20	10	94^0C	0:20	10
[1]68-62^0C	1:00		[2]65-57^0C	1:00	
94^0C	0:20	50	94^0C	0:20	40
62^0C	1:00		57^0C	1:00	

[1]The annealing temperature was decreased 0.6 ^{0}C each cycle starting from 68^0C down to 62^0C.

[2]The annealing temperature was decreased 0.8 ^{0}C each cycle starting from 62^0C down to 57^0C.

Table A2. Anthracnose (*Colletotrichum sublineola*) disease reaction of sorghum (*Sorghum bicolor* (L.) Moench) recombinant inbred lines (RILs) derived from the cross of resistant line IS18760 and susceptible line PI609251 evaluated in Puerto Rico (PR) and Texas (TX) during 2016.

RIL ID	PR	TX	RIL ID	PR	TX	RIL ID	PR	TX
RIL-001	5	3.5	RIL-048	.	2.5	RIL-099	3.5	3
RIL-002	5	.	RIL-049	5	.	RIL-100	.	.
RIL-003	3	2	RIL-050	4	2	RIL-101	3.5	2.5
RIL-004	3	2	RIL-051	3.5	2	RIL-102	.	.
RIL-005	3.5	2.5	RIL-052	4.5	2.5	RIL-103	2	2
RIL-007	3	2	RIL-053	3.5	3.5	RIL-105	3	2.5
RIL-008	4	.	RIL-054	4.5	2	RIL-106	.	.
RIL-009	3	2	RIL-055	4	3	RIL-107	3	2.5
RIL-010	3	2	RIL-056	5	3	RIL-108	5	.
RIL-012	4.5	2.5	RIL-058	3	2	RIL-110	5	.
RIL-014	3.5	2.5	RIL-059	5	.	RIL-111	3	.
RIL-015	3.5	3	RIL-060	5	.	RIL-112	4	2.5
RIL-016	4	3.5	RIL-061	5	.	RIL-113	4	3
RIL-017	3.5	2.5	RIL-063	3	3	RIL-114	3	3
RIL-018	4	2.5	RIL-065	4.5	.	RIL-115	4	2
RIL-019	3.5	2	RIL-066	5	.	RIL-116	5	.
RIL-020	5	3	RIL-067	2	2	RIL-117	4.5	.
RIL-021	4	3	RIL-068	3.5	2.5	RIL-118	5	.
RIL-022	2	2	RIL-069	4.5	2.5	RIL-119	.	.
RIL-023	2	2	RIL-070	4	2	RIL-121	.	.
RIL-024	4	3	RIL-071	3	2	RIL-122	.	.
RIL-025	4	3	RIL-073	4.5	3	RIL-123	3.5	2.5
RIL-026	3	2	RIL-075	4	3	RIL-124	4	3
RIL-027	4	2.5	RIL-076	4.5	3	RIL-125	4.5	3
RIL-028	3	2.5	RIL-077	5	3	RIL-126	3	2.5
RIL-029	4	2	RIL-078	5	3	RIL-127	.	.
RIL-030	2	2.5	RIL-079	4.5	3.5	RIL-128	.	.
RIL-031	3.5	3	RIL-080	3	2.5	RIL-129	5	2.5
RIL-032	3	2.5	RIL-082	3.5	2	RIL-130	5	2
RIL-033	3	3	RIL-083	3.5	3	RIL-131	5	.
RIL-034	4.5	3	RIL-084	4.5	.	RIL-132	4.5	3
RIL-035	4.5	2.5	RIL-085	5	.	RIL-133	4.5	3
RIL-036	4.5	3	RIL-086	4	2	RIL-134	5	.
RIL-037	5	3	RIL-088	5	3.5	RIL-135	3	3
RIL-038	5	3	RIL-089	.	.	RIL-136	3.5	3
RIL-039	5	3	RIL-090	4.5	3	RIL-138	.	.
RIL-040	2	2	RIL-091	5	4	RIL-139	3.5	2.5
RIL-041	4	3	RIL-092	3	2.5	RIL-141	4.5	3
RIL-042	3.5	3	RIL-093	4	3	RIL-142	3	3
RIL-043	4.5	2.5	RIL-094	4	2	RIL-144	4.5	3
RIL-044	5	.	RIL-095	3	2.5	RIL-145	4.5	2.5
RIL-045	4	2.5	RIL-096	4.5	3.5	RIL-146	4	3.5
RIL-046	3.5	2.5	RIL-097	3.5	2.5			
RIL-047	3	3	RIL-098	3.5	2.5			

Anthracnose disease reaction according to Prom et al., (2009) where score 1.0-2.0 are resistant, 2.5-3.0 moderate resistant, and 3.0-5.0 susceptible. Dots (.) represents missing values.

Table A3. Anthracnose (*Colletotrichum sublineola*) disease reaction of sorghum (*Sorghum bicolor* (L.) Moench) recombinant inbred lines (RILs) derived from the cross of resistant line QL3 and susceptible line PI609251 evaluated in Puerto Rico (PR) and Texas (TX) during 2016.

RIL ID	PR	TX	RIL ID	PR	TX	RIL ID	PR	TX
RIL-001	.	2	RIL-050	2.5	.	RIL-102	4.5	3
RIL-002	3.5	3	RIL-051	4.5	3	RIL-103	4	2
RIL-004	5	.	RIL-052	5	3	RIL-104	.	.
RIL-005	5	3	RIL-054	3.5	2	RIL-105	5	3
RIL-007	4	3	RIL-055	5	.	RIL-106	5	.
RIL-008	4.5	2.5	RIL-056	4.5	3	RIL-107	4	3
RIL-009	4	2	RIL-057	4.5	3	RIL-108	.	.
RIL-010	5	2.5	RIL-058	2	2	RIL-109	.	3
RIL-011	4	3	RIL-059	5	3			
RIL-013	5	3.5	RIL-061	3	3			
RIL-014	4	3	RIL-062	5	3			
RIL-015	3.5	3	RIL-063	4	2.5			
RIL-016	5	.	RIL-065	.	2			
RIL-017	.	2	RIL-066	4.5	.			
RIL-018	5	2	RIL-067	5	3			
RIL-020	4	3.5	RIL-068	5	3.5			
RIL-021	5	2	RIL-069	.	.			
RIL-022	4.5	.	RIL-070	3.5	3			
RIL-023	3	2.5	RIL-071	5	2.5			
RIL-025	5	.	RIL-072	5	2			
RIL-026	4.5	2	RIL-073	5	3			
RIL-027	3	2	RIL-074	4	2.5			
RIL-028	5	2	RIL-075	4.5	3			
RIL-029	4	3	RIL-078	4	3			
RIL-030	5	3	RIL-079	5	3			
RIL-031	4	2	RIL-080	5	.			
RIL-032	5	3	RIL-081	4.5	3.5			
RIL-033	3	3	RIL-082	5	.			
RIL-034	3	.	RIL-083	.	.			
RIL-035	4	3.5	RIL-084	5	3			
RIL-036	3	3	RIL-085	.	2			
RIL-037	5	2.5	RIL-086	4.5	3			
RIL-038	5	2	RIL-087	.	3			
RIL-039	4	2	RIL-088	.	3			
RIL-040	3	3	RIL-089	4	.			
RIL-041	4	2	RIL-091	3	3			
RIL-042	4	2	RIL-092	5	.			
RIL-043	4.5	3.5	RIL-093	4	2			
RIL-044	.	.	RIL-094	.	2.5			
RIL-045	4	2.5	RIL-095	4				
RIL-046	2	2	RIL-097	3.5				
RIL-047	4	.	RIL-098	4				
RIL-048	3	2	RIL-099	4				
RIL-049	4.5	3	RIL-101	5				

Anthracnose disease reaction according to Prom et al., (2009) where score 1.0-2.0 are resistant, 2.5-3.0 moderate resistant, and 3.0-5.0 susceptible. Dots (.) represents missing values.

Table A4. Anthracnose (*Colletotrichum sublineola*) disease reaction of sorghum (*Sorghum bicolor* (L.) Moench) recombinant inbred lines (RILs) derived from the cross of resistant line SC112-14 and susceptible line PI609251 evaluated in Puerto Rico (PR), Texas (TX), Florida (FL) and Georgia (GA) during 2016.

RIL ID	PR	FL	GA	TX	RIL ID	PR	FL	GA	TX	RIL ID	PR	FL	GA	TX
RIL-002	4.5	4	4.5	3	RIL-056	2	2	2.5	2	RIL-106	4	4.5	4.5	3
RIL-003	4	4	4.5	3.5	RIL-057	5	5	5	3	RIL-108	3	2	2	.
RIL-006	4.5	4.5	5	3.5	RIL-058	5	4	2.5	2	RIL-109	2	2	3	3
RIL-008	4	5	5	3.5	RIL-059	5	4.5	4	3	RIL-114	5	5	5	3
RIL-009	4.5	5	5	3	RIL-060	2	2	3	2	RIL-117	5	4	5	4
RIL-010	4.5	5	5	4	RIL-061	3	2	2	2	RIL-118	3.5	4.5	5	4
RIL-011	5	3.5	4	3	RIL-067	5	5	4.5	2.5	RIL-119	4	4.5	3	2.5
RIL-013	5	2.5	3	3	RIL-068	3.5	2	2	2	RIL-122	.	.	.	.
RIL-014	4	3.5	2.5	3	RIL-069	.	.	.	.	RIL-123	5	4.5	4.5	4.5
RIL-016B	2	2	2	2	RIL-070	.	.	.	3	RIL-125	3	2	4.5	2
RIL-017	5	3.5	3.5	3	RIL-071	5	3	3.5	3	RIL-126	5	4	4	3
RIL-020	5	5	5	3.5	RIL-075	.	.	.	.	RIL-127	5	5	5	3
RIL-021	2	2	3	2	RIL-076	5	4.5	3	.	RIL-129	.	.	.	.
RIL-022	2	2	2	2	RIL-078	2	2	2	2	RIL-130	.	.	.	.
RIL-023	2	3.5	3	2	RIL-079	.	.	.	3	RIL-131	2	2	2.5	2
RIL-024	.	.	.	.	RIL-080	2	2.5	2	2	RIL-132	2	3	2.5	2
RIL-025	.	.	.	.	RIL-081	.	.	.	.	RIL-133	5	3.5	3	3
RIL-026	2.5	2	2	2.5	RIL-082	2	2	2	2	RIL-134	.	.	.	.
RIL-027	.	.	.	3	RIL-083	.	.	.	.	RIL-135	2	2	2	2
RIL-028	.	.	.	2.5	RIL-085	4	2.5	2	2	RIL-136	5	3	3.5	3
RIL-032	2	2	2	2	RIL-086	4	5	4	3	RIL-137	4.5	2.5	4	3
RIL-033	2	2	2.5	2	RIL-088	4	2	3	2.5	RIL-140	3.5	2.5	2	2
RIL-034	5	4.5	4	3	RIL-089	.	.	.	2	RIL-141	4.5	3	2.5	3
RIL-035	4.5	2	2	2	RIL-090	3.5	4	3.5	3					
RIL-038	5	3.5	4	3	RIL-091	2	2	2	2					
RIL-039	2	2	2	2	RIL-092	4.5	3.5	4.5	4					
RIL-040	.	.	.	.	RIL-093	.	.	.	.					
RIL-041	3.5	2	2	.	RIL-094	.	.	.	.					
RIL-043	4.5	5	4.5	4	RIL-097	2	3	2	2					
RIL-050	4	2	5	2.5	RIL-098	.	.	.	.					
RIL-051	2	2	2	2	RIL-101	3	2	2	2					
RIL-052	4.5	2	2	2	RIL-102	.	2	2	2.5					
RIL-053	3	3	4	2	RIL-103	2	2	2	2					
RIL-055	2	2	2	2	RIL-104	.	.	.	.					

Anthracnose disease reaction according to Prom et al., (2009) where score 1.0-2.0 are resistant, 2.5-3.0 moderate resistant, and 3.0-5.0 susceptible. Dots (.) represents missing values.

Table A5. Number of single nucleotide polymorphisms (SNPs) per chromosome removed from linkage map based on the PERL script developed by Nave et al., (2016). SNPs were removed if within a particular interval of the linkage map they introduced a double crossing over (DCO) while another high-quality SNP was present in the interval (window size [WS]). The linkage maps were constructed from three sets of sorghum [*Sorghum bicolor* (L.) Moench] derived from anthracnose resistant lines IS18760, QL3, and SC112-14 crossed with the common susceptible line PI 609251.

	IS18760		**QL3**		**SC112-14**	
CHRM[1]	DCO[2]	WS[3]	DCO	WS	DCO	WS
1	59	51	53	10	73	53
2	66	111	76	15	32	35
3	85	117	61	13	29	30
4	2	62	76	10	35	62
5	43	73	89	17	28	26
6	53	29	55	13	42	25
7	18	12	30	2	25	12
8	37	9	64	23	29	12
9	52	40	63	11	32	23
10	35	38	71	26	27	31

[1] Chromosomes

[2] Number of markers that were removed because they introduced a double crossing over

[3] Number of markers removed because there was a high-quality SNP within the window size (WS)

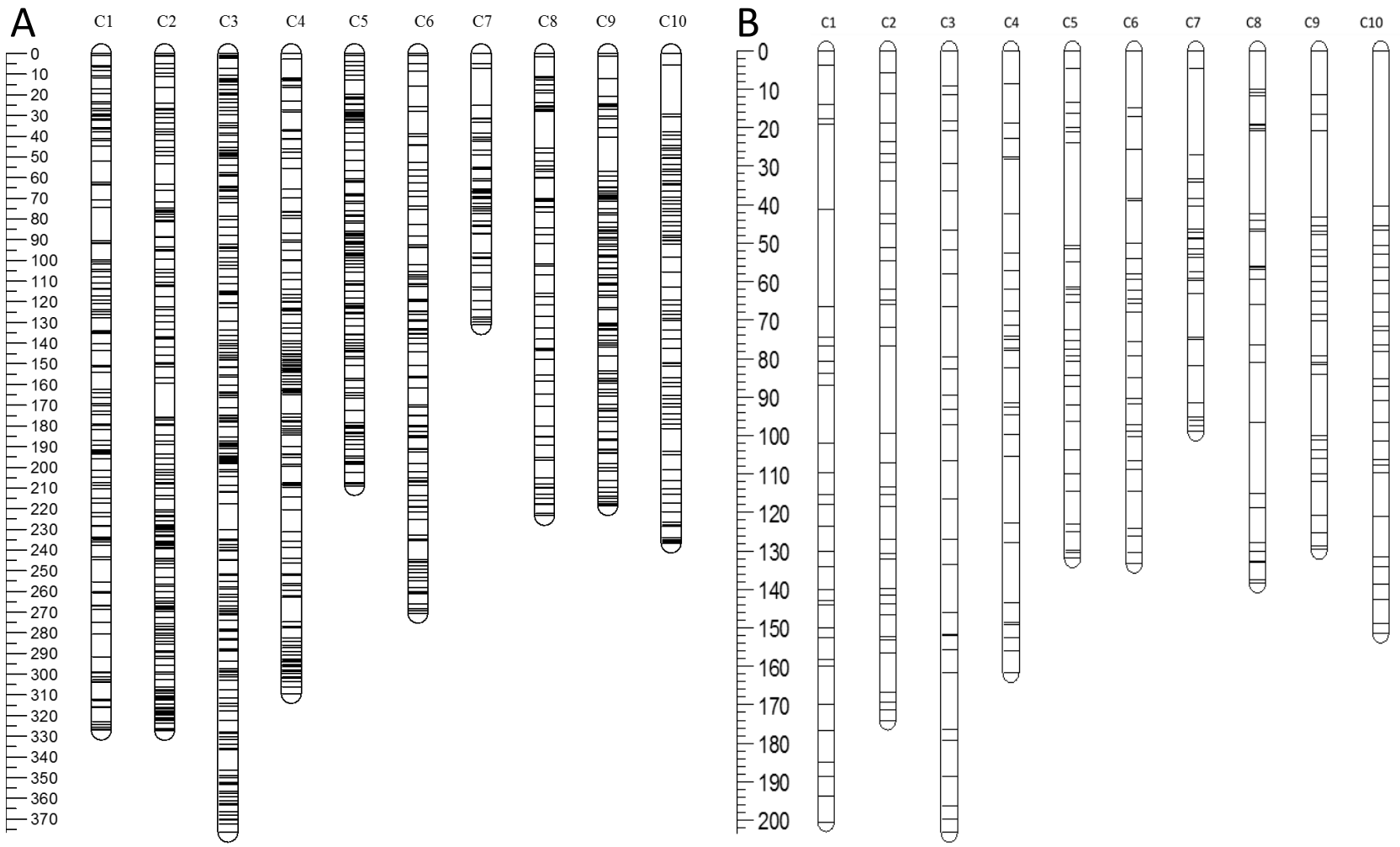

Figure A1. Recombination linkage map based on genotype-by sequence analysis of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred line (RILs) derived from the cross of anthracnose (*Colletotrichum sublineola*) resistant line IS18760 and the susceptible line PI609251. (A) high density linkage map based on 1,355 single nucleotide polymorphism (SNPs) and (B) a subset of 303 SNPs selected base on recombination information.

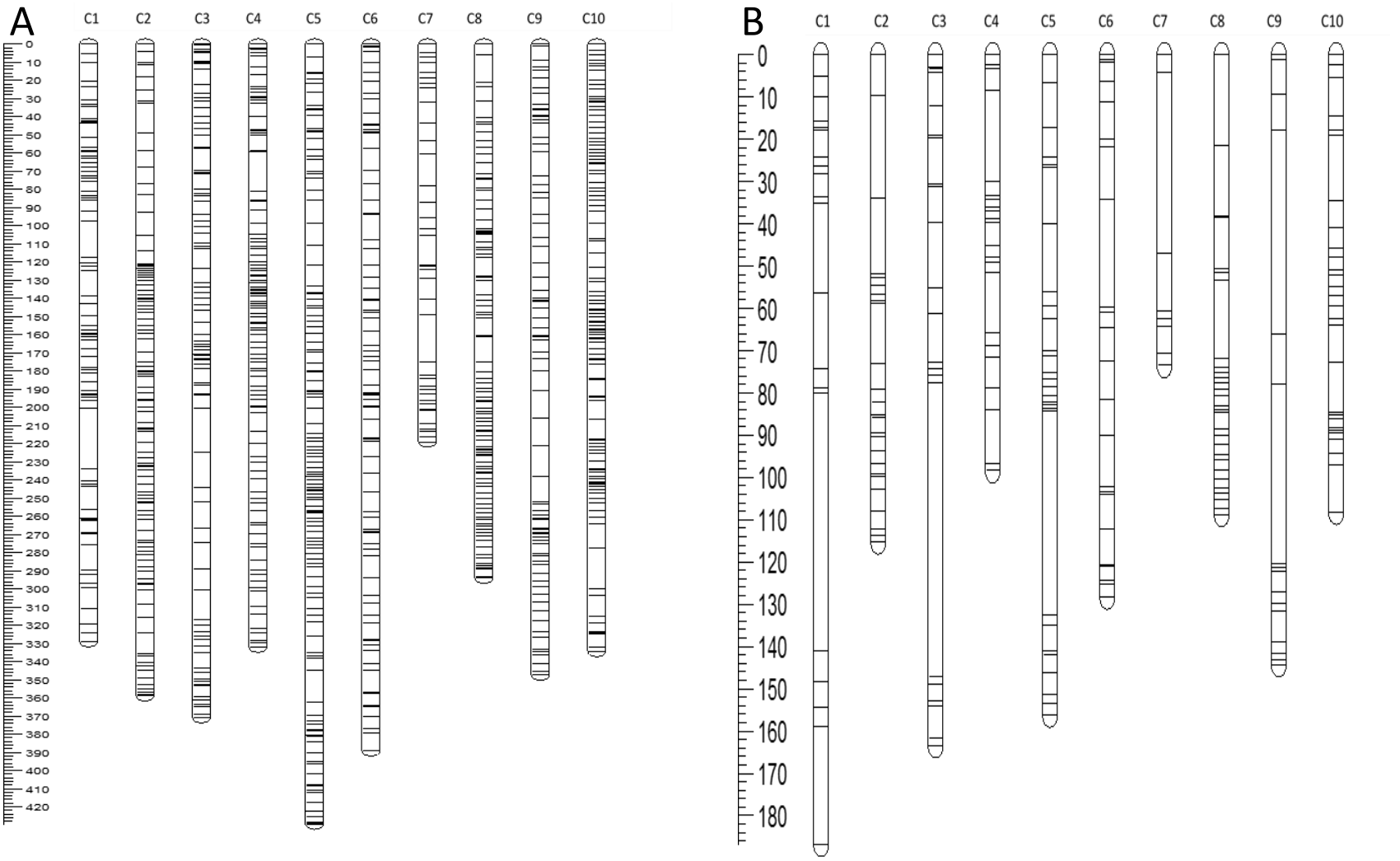

Figure A2. Recombination linkage map based on genotype-by sequence analysis of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred line (RILs) derived from the cross of anthracnose (*Colletotrichum sublineola*) resistant line QL3 and the susceptible line PI609251. (A) high density linkage map based on 1,008 single nucleotide polymorphism (SNPs) and (B) a subset of 230SNPs selected base on recombination information.

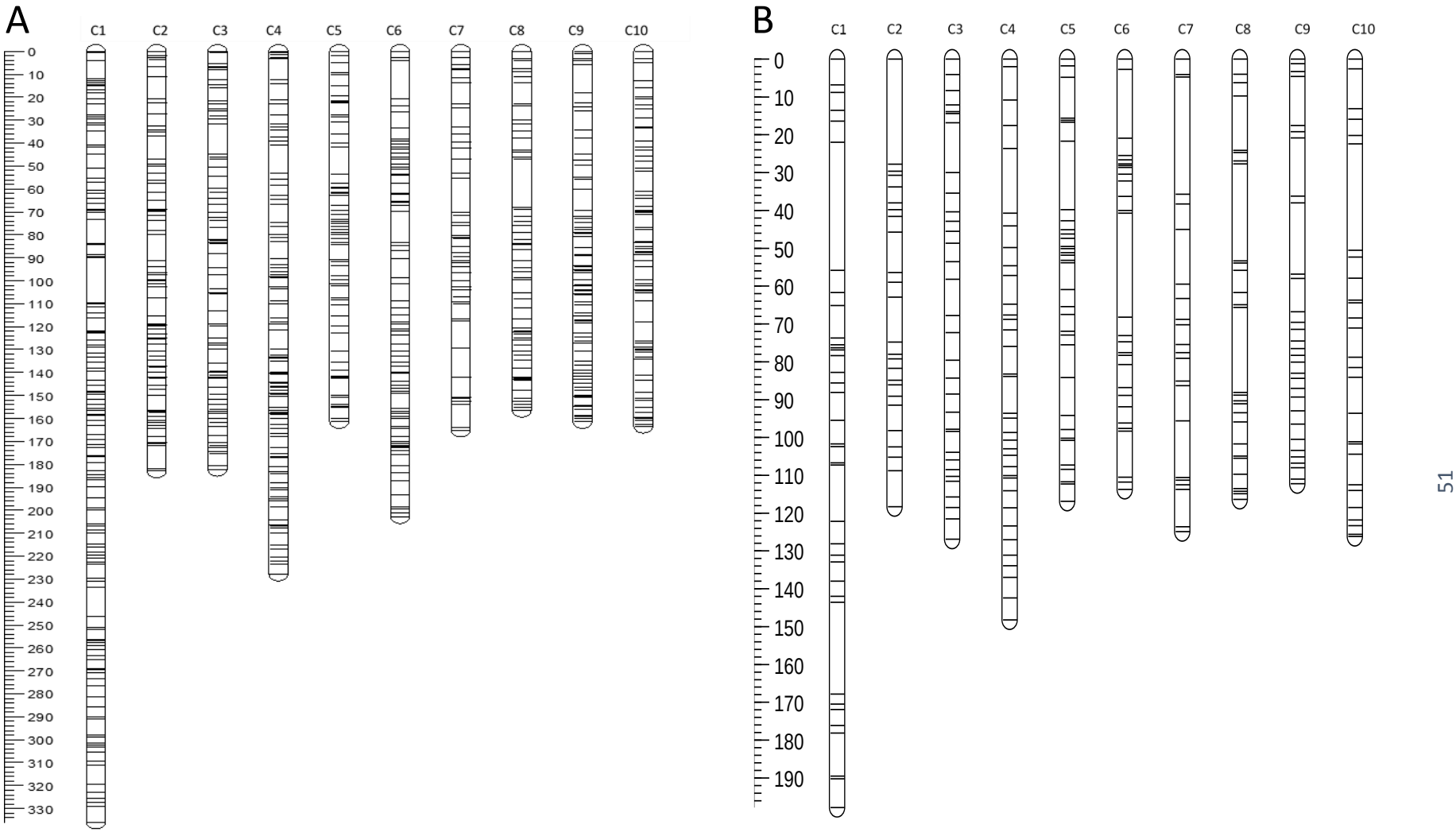

Figure A3. Recombination linkage map based on genotype-by sequence analysis of sorghum [*Sorghum bicolor* (L.) Moench] recombinant inbred line (RILs) derived from the cross of anthracnose (*Colletotrichum sublineola*) resistant line SC112-14 and the susceptible line PI609251. (A) high density linkage map based on 952 single nucleotide polymorphism (SNPs) and (B) a subset of 306 SNPs selected base on recombination information.

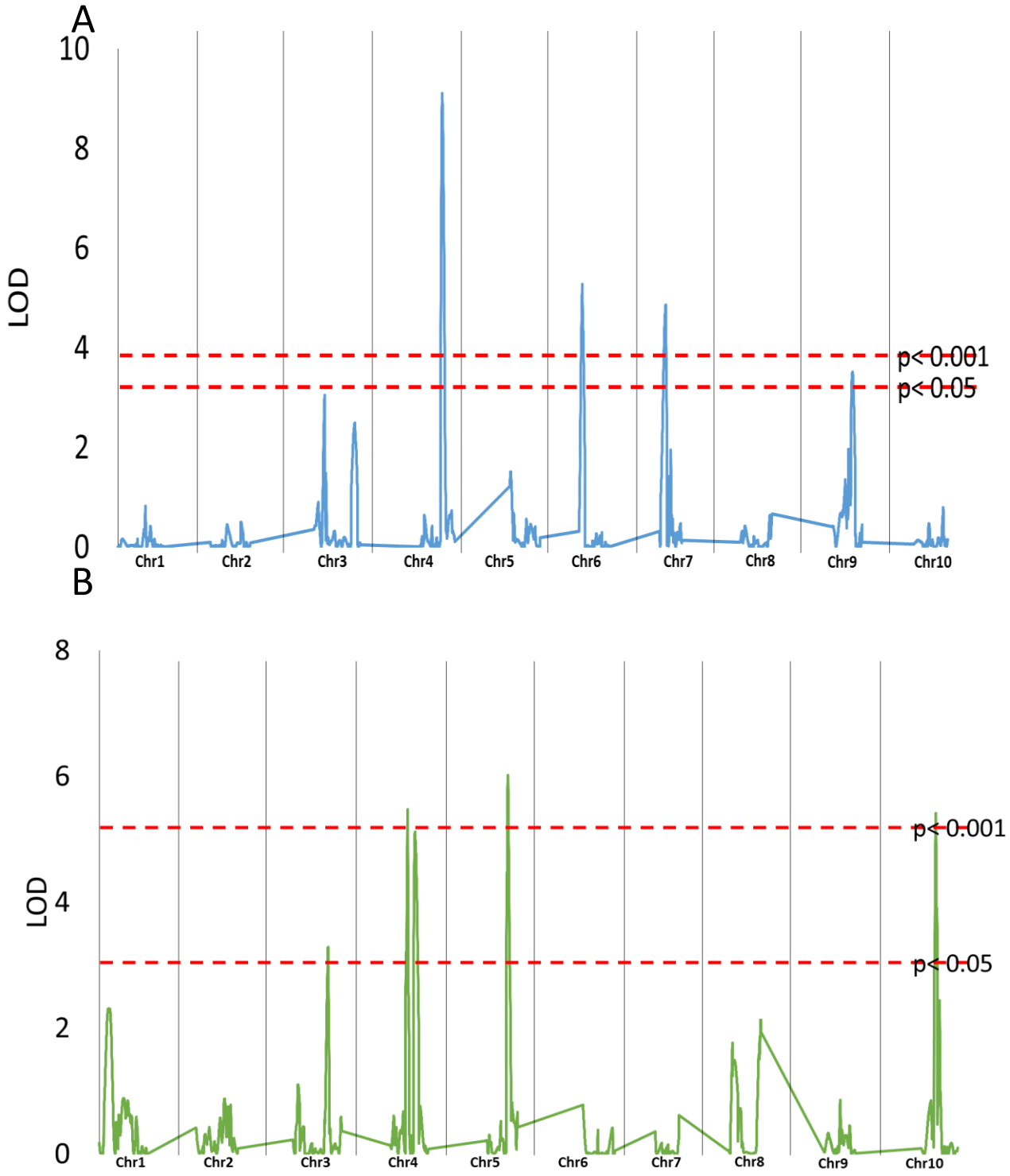

Figure A4. Composite interval mapping for anthracnose (*Colletotrichum sublineola*) resistant response in sorghum [*Sorghum bicolor* (L.) Moench] using a recombinant inbred line (RILs) population derived from the resistant line IS18760 crossed with the susceptible line PI609251. The RILs were evaluated for anthracnose resistant response in (A) Puerto Rico and (B) Texas in 2016.

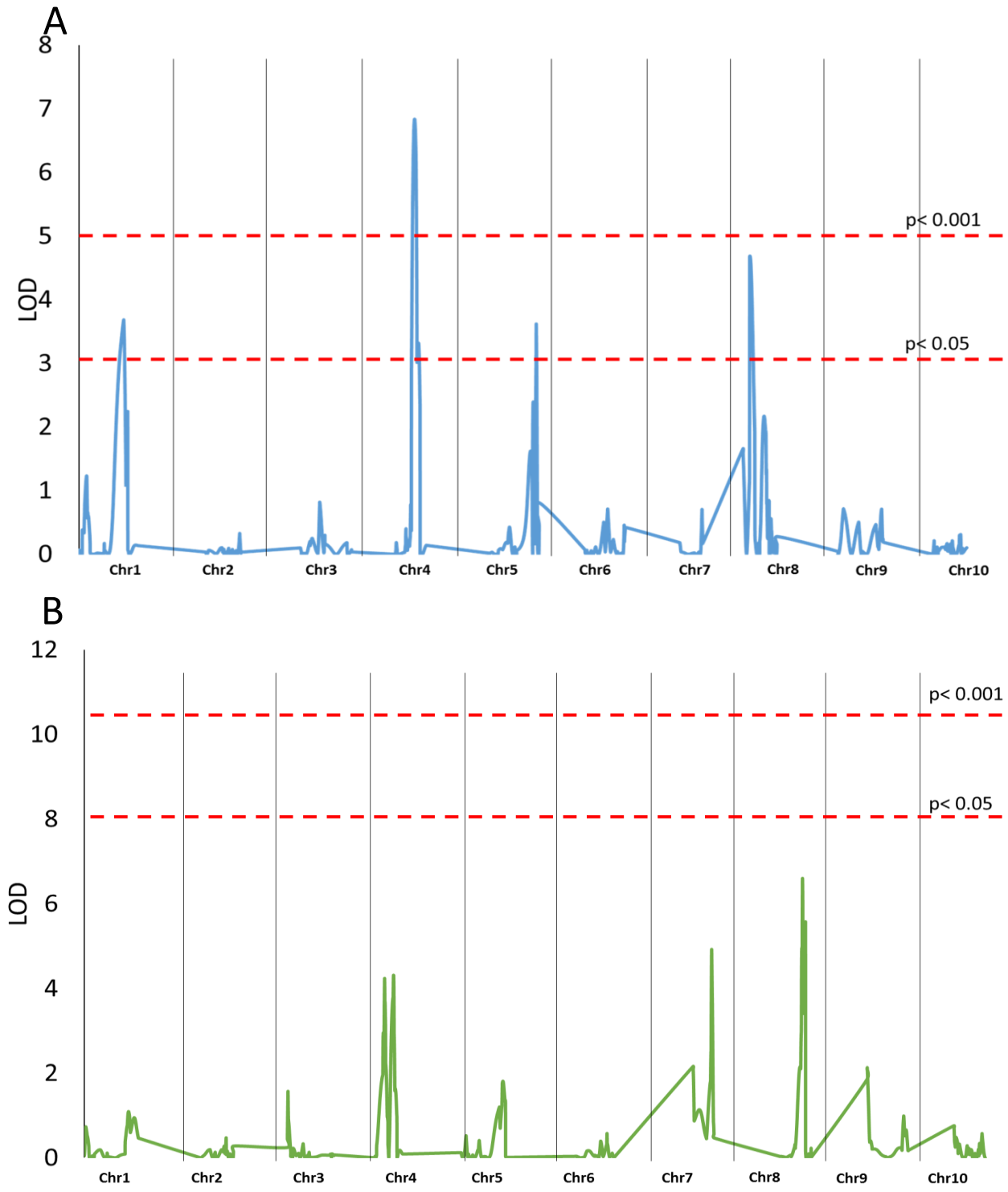

Figure A5. Composite interval mapping for anthracnose (*Colletotrichum sublineola*) resistant response in sorghum [*Sorghum bicolor* (L.) Moench] using a recombinant inbred line (RILs) population derived from the resistant line QL3 crossed with the susceptible line PI609251. The RILs were evaluated for anthracnose resistant response in (A) Puerto Rico and (B) Texas in 2016.

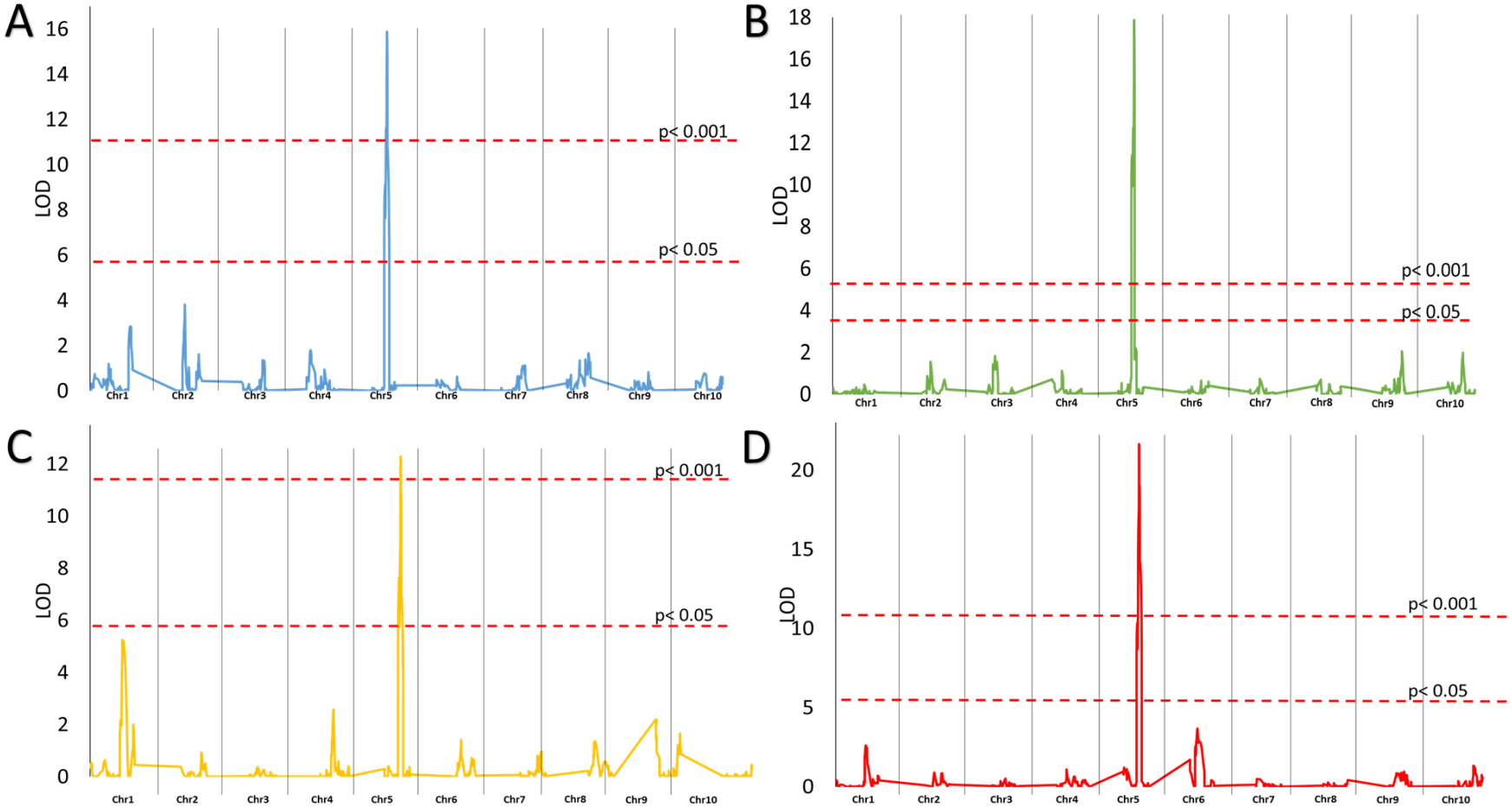

Figure A6. Composite interval mapping for anthracnose (*Colletotrichum sublineola*) resistant response in sorghum [*Sorghum bicolor* (L.) Moench] using a recombinant inbred line (RILs) population derived from the resistant line SC112-14 crossed with the susceptible line PI609251. The RILs were evaluated for anthracnose resistant response in (A) Puerto Rico, (B) Texas, (C) Florida, and (D) Georgia in 2016