GENETIC ANALYSIS OF LEAF MORPHOLOGY AND YIELD-RELATED TRAITS IN SORGHUM POPULATIONS

By

WIRIYANAT PLOYARAM

(Under the Direction of Andrew H. Paterson)

ABSTRACT

Sorghum is human food, animal feed, and biofuel. This research aims to determine the relationship between genes and leaf morphology traits in sorghum populations, including leaf length, leaf width, leaf angle, and midrib diameter. Quantitative trait locus (QTL) mapping and Genome-wide association study (GWAS) analysis in a diversity panel were used to link leaf traits (phenotypes) to DNA marker (genotypic) data. Two RIL populations with a common parent (*S. bicolor* BTx623) were studied by QTL mapping using CIM. The candidate QTLs influencing leaf morphology and yield-related traits were compared to GWAS results in a sorghum diversity panel (SAP) and to other genes reported to influence leaf morphology traits in sorghum and other plants, especially turf grasses. Well-supported QTLs will aid marker-assisted sorghum breeding.

INDEX WORDS: QTL mapping, GWAS, leaf morphology, yield-related traits, GBS, SSR

GENETIC ANALYSIS OF LEAF MORPHOLOGY AND YIELD-RELATE TRAITS IN SORGHUM POPULATIONS

by

WIRIYANAT PLOYARAM

B.S., Iowa State University, USA, 2020

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2022

GENETIC ANALYSIS OF LEAF MORPHOLOGY AND YIELD-RELATE TRAITS IN SORGHUM POPULATIONS

by

WIRIYANAT PLOYARAM

Major Professor:	Andrew H. Paterson
Committee:	Peng-Wah Chee
	Jason G. Wallace

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2022

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to my major professor, Dr. Andrew Paterson, for his oversight and direction during my research. I feel the abilities and strategies I have acquired while working under his direction will motivate me to advance in my career. I would like to thank my committee members, Dr. Peng-Wah Chee and Dr. Jason Wallace, for their support, suggestions, and valuable time guiding me through this journey. I also would like to thank everyone at the Plant Genome Mapping Laboratory (PGML) for their support and assistance with this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Purpose of the study

This study aims to identify genomic regions affecting traits related to leaf morphology of sorghum in two populations sharing a common parent, *S. bicolor* BTx623, an elite inbred that was the source of the sorghum reference genome. This particular parent was crossed with either *S. bicolor* IS3620, representing race 'guinea' which is highly divergent from BTx623; or *S. propinquum*, a sister species within the genus Sorghum to create IS (*S.* BTx623 x *S.* IS3620) and PQ (*S. bicolor* BTx623 x *S. propinquum*) recombinant inbred line (RIL) populations respectively. Both populations, IS and PQRIL, that span much of the genetic diversity available in 'eusorghums'. Genotyping data obtained from genotyping-by-sequencing (GBS) as well as SSRs were utilized to perform quantitative trait loci (QTL) mapping in both sorghum populations. Similarly, genome-wide association study (GWAS) uses genotypes from multiple sorghum diversity panels obtained from multiple published GBS data. GWAS generally produces results with a high resolution; however, it has a high false-positive rate compared to QTL mapping. Conversely, QTL mapping results in a lower false positive rate, yet it has low resolution. Hence, the use of GWAS analysis complements QTL mapping, leading to high confidence in more precise regions within which to identify candidate genes which may contribute to the traits of interest. The approach of using both Next-Generation Sequencing

(NGS) technology and computational analyses in the study will allow the elucidation of the relationship between genes and leaf morphology from the two different populations. The QTLs or genes identified in this study will help breeders to optimize the leaf morphology in breeding programs. The identification of the genes will enhance genetic resources which will help identify the leaf morphology varieties that are more capable to resist climate change and pathogenic attacks for crop improvement.

Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] is one of the world's staple food resources and commonly grown in developing as well as developed countries. It is the world's fifth most important cereal crop after rice, wheat, maize, and barley. The world production of sorghum was reported to be 59.34 million metric tons in 2018 (FAO). Developing countries in Africa and Asia have the highest sorghum production globally, primarily using the crop for food purposes. In contrast, in developed countries, sorghum is used for animal feed (FAO). As a type of food resource, sorghum contained several macro and micronutrients which not only provide energy but also possess health benefits for humans (Anglani 1998). From a macronutrient perspective, one cup of sorghum grains provides approximately 632 calories, consisting 88% of carbohydrates, 9% of fat, and 3% of protein. Regarding micronutrients, this cereal crop has high contents of potassium, phosphorus, vitamin Bs such as thiamin, riboflavin, vitamin B6, biotin and niacin, however, it contains low calcium (Anglani 1998). In animal agriculture, sorghum plays an important role as a key ingredient of animal feed due to its low-cost production as well as high contents of energy and nutrients.

In addition to providing food security, sorghum has become a promising source of alternative energy in the form of biofuel produced by plant biomass. Biobutanol (butyl alcohol) is derived from cellulosic sugars in agricultural wastes such as corn stover, barley and wheat straw, lesquerella presscake and sweet sorghum bagasse. Biobutanol is considered a cleaner burning alternative gasoline. The benefits of biobutanol include having a higher energy content, a lower Reid vapor pressure (a common measure of the volatility of gasoline and other petroleum products), a higher energy security, and fewer emissions compared to ethanol. For instance, corn grain butanol meets the renewable fuel 20% greenhouse gas emission reduction threshold as required by the Renewable Fuel Standard. Sorghum, specifically sweet sorghum, shows promise as a biobutanol resources due to its drought tolerance, minimal water uptake and adaptability to wide-ranging growing conditions (USDA). Because of its C_4 photosynthesis pathway, sorghum has the ability to genetically tolerate hot and dry environments compared to most plants employing the C_3 photosynthesis mechanism. The C_4 photosynthesis pathway allows plants to accumulate carbon dioxide more efficiently with reduced water usage under high temperatures and light conditions (Mathur *et al.* 2017). Therefore, sorghum shows great potential as a resource to address many needs, such as food for humans, animal feed and biofuel.

Domestication of sorghum

The earliest records show that sorghum was discovered around approximately 5,000 – 6,000 years ago in Northeast-Central Africa (de Wet and Huckabay 1967; Winchell *et al.* 2017; Burgarella *et al.* 2021). Afterwards, the crop was introduced to different continents and countries including India, China, the United States of America, and Australia respectively (Burgarella *et al.* 2021). *Sorghum bicolor* was reported to be the earliest *Sorghum* species cultivated in the Indus Valley, India back in 2000 – 1700 BC. The Yellow River Valley is considered to be the area where the earliest sorghum was cultivated in China (Venkateswaran *et al.* 2019). The Chinese 'Kaoliang' line of sorghum originated from *Sorghum bicolor,* which was introduced from India (Doggett, 1998). During the 19th century, sorghum was introduced to the USA and Australia respectively. Although the exact number of species distributed across the globe based on morphological and molecular evidence, USDA considers the subgenus 'eusorghum' as the 'true sorghum' which consists of three species, *S. bicolor, S. propinquum, S. halepense* and a hybrid (*S. bicolor x S. halepense*) species called *Sorghum x almum* parodi (Dillon *et al.* 2007).

Domestication in sorghum has contributed to genotypic and phenotypic changes. Artificial selection throughout the domestication process has allowed the plant to be adapted under agricultural environments for human usage. Since humans prefer certain traits to be conferred on the crops, the domesticated and wild plant become different throughout time. The collection of traits caused by domesticated changes is referred to as 'domestication syndrome' where particular phenotypic traits become common among domesticated species (Paterson 2002a; Lai *et al.* 2018). For example, domesticated plants have changed in several characters such as size, shape and yield of seeds, seed dispersal and plant architecture. The phenotypic changes in domesticated species are associated with genetic changes.

Genetic diversity and germplasm base

Sorghum is known as a well-adaptive plant having great genetic diversity, potential to adapt locally under human and natural selection, as well as ability to efficiently grow in diverse environment. The genus *Sorghum* consists of five subgenera according to morphological characters: *Eu-sorghum, Chaetosorghum, Heterosorghum, Parasorghum,* and *Stipososorghum* (Ananda et al., 2020; Garber & Snyder, 1951; Harlan & Wet, 1972). The three major species in sorghum, *Sorghum bicolor, Sorghum halepense* (Johnson grass), and *Sorghum propinquum*, belong to the subgenera *Eu-sorghum*. Moreover, three subspecies*: subsp. bicolor, subsp. verticiliflorum*, and *subsp. drummondii* lie within the *S. bicolor* species. In fact, the *subsp.*

bicolor has five races: bicolor, caudatum, durra, guinea, and kafir (Lazarides *et al.* 1991; Ananda *et al.* 2020; Xin *et al.* 2021a). Based on hardiness and adaptation, the durra race is hardy and adapted to dry zones, whereas the guinea and bicolor races which are also dapted to wet zones. However, the kafir and caudatum races adapt to intermediate zone for high yielding. Nevertheless, sorghum is categorized based on its usage such as grain, forage, and sweet sorghum where each type of sorghum also possesses different characteristics. For example, sweet sorghum has relatively thicker stems served as a primary sink tissue for sugar production (Kanbar *et al.* 2019). Therefore, sorghum contains significant diversity within its species for further use in sorghum breeding programs to produce a wide range of elite and diverse sorghum lines among grain, forage, and sweet sorghum genotypes.

As for the sorghum germplasm bases, four major centers hold different collections of sorghum germplasm throughout the world. Two organizations in India, the international Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India maintains 37, 949 accessions assembled from 92 countries, and the National Bureau of Plant Genetic Resources (NBPGR) holds 20,221 accessions. The United States Department of Agriculture (USDA) National Plant Germplasm System (NPGS) preserves 45,000 accessions at the Plant Genetic Resources Conservation Unit in Griffin, GA. Lastly, the Institute of Crop Science, Chinese Academy of Agricultural Sciences (ICS-CAAS) in China maintains 18,263 accessions. Cultivated accessions account for 98.3%, whereas wild weedy relatives are 1.7% of the collection respectively (Xin *et al.* 2021b).

The variations of the mapping populations

Since *Sorghum bicolor* has many wild relatives such as *S. propinquum* and *S. halepense,* including hybrid species, there are many possibilities for intra- and interspecific crosses. The *S.*

bicolor accession IS3620 has diverged from *S. bicolor* BTx623 and represents the ‘guinea’ sorghum race. Based on the neighbor-joining method, a study reported that guinea subgroup formed a separate cluster and might represent an independent domestication compared to other sorghum races (Morris *et al.* 2013a). Next, *S. propinquum* is a perennial Southeast Asian native species. This species has rhizomatous characteristics that are absent from the species that gave rise to the cultigen, *S. bicolor*. Both *S. bicolor* and *S. propinquum* are diploid possessing the same number of chromosomes (2n = 20), where 2n is the somatic chromosome number having 2 complete sets (2x) of chromosome.

In ancient times, hybridization between *S. bicolor x S. propinquum* occurred naturally, forming *S. halepense* (Paterson *et al.* 2020). Unlike the ancestors whose set of chromosomes is diploid (2n = 2x = 20), *S. halepense* is tetraploid (2n = 4x = 40). In other words, a haploid cell of diploid sorghum is n = x =10 while a haploid cell of tetraploid *S. halepense* is n = 2x = 20. In interspecific crosses between *S. halepense* and *S. bicolor*, both triploid and tetraploid hybrids have been observed. Triploid progenies are prone to have sterile males as well as low female fertility, therefore they are less likely to be reproducible. Conversely, tetraploid progenies tend to form via an unreduced (2n = 2x = 20) sorghum gamete and a reduced (n = 2x = 20) *S. halepense* gamete (Hodnett *et al.* 2019) hence, their reproductive ability is not negatively affected.

Sorghum halepense, commonly known as Johnsongrass, is well-known to be an invasive and highly competitive plant. It is very destructive to other species due to its ability to invade and outcompete native plants in their habitats. Moreover, Johnsongrass has become a major threat to crop production since it shelters several agricultural pests and viruses (Klein and Smith 2021). Johnsongrass is considered an invasive weed species due to several factors namely, having an effective propagation via rapid flowering and disarticulation of mature inflorescences, as well as possessing underground rhizomes up to 70% of the plant's dry weight to store nutrients and to rapidly produce new vegetative growth. From a RIL population derived from *S. bicolor* and *S. propinquum*, a study of rhizomatousness and vegetative branching was conducted. Based on the discovery of five regions conferring rhizomatousness corresponding with the branching QTLs, study results have shown that the above ground vegetative branching and below ground rhizome growth are related to each other (Kong *et al.* 2015). Additionally, Johnsongrass is known to be herbicide-resistant and currently, there are no herbicides that can be used to eliminate the invasive plant without damaging sorghum.

Genetic mapping in sorghum

Genetic mapping is a powerful tool to understand the relationship between genes and the inheritance of traits from parents to offspring. Genetic mapping aims to identify the localization of genes influencing phenotypes based on correlation with DNA variation. Linkage analysis relies on polymorphic variants (markers) due to meiotic recombination during crosses between parents. Any marker that shows correlated segregation (linkage) with the trait, the marker is localized nearby in the genome (Altshuler *et al.* 2008). Genetic marker refers to a sequence of DNA used for identifying the presence and the location of other genes on a genetic map. The marker is derived from the difference in phenotypic expression controlled by genes. This difference can be used to compare individuals for studying recombination processes or identifying a target gene that is closely located nearby the marker. The molecular marker is a gene or DNA sequence that is linked with a certain location within the genome. This type of marker allows investigation of the inheritance of that genetic information of the gene.

Genetic mapping started in the early 1990s via the use of molecular markers such as RFLP (Restriction fragment length polymorphism) markers, AFLP (Amplified fragment length polymorphism), RAPD (Random amplified polymorphic DNA) markers, SSRs (simple sequence repeats), and DArT (Diversity Array technology) markers (Bhattramakki et al. 2000; Boivin et al., 1999; Chittenden et al., 1994; Ejeta & Knoll, 2007; Hulbert et al., 1990; Menz et al., 2004; Peng et al., 1999; Pereira et al., 1994; Singh and Lohithaswa 2006; Tao et al., 2000). These DNA markers are essential tools for genetic linkage map construction in many research areas such as marker-assisted selection, quantitative trait loci (QTL) mapping, and map-based cloning (Kirungu *et al.* 2018). However, the RAPD and AFLP techniques employ dominant markers which fail to differentiate heterozygous individuals from homozygous dominant ones. SSR markers which have reproducibility and use co-dominant markers became more common in sorghum gene mapping, genome evolution, molecular genetics and marker-assisted breeding (Tao *et al.* 1998; Wu and Huang 2007; Guan *et al.* 2011; Kong *et al.* 2013).

With rapid growth of advanced sequencing technology allows cheaper and more accessible sequencing platforms for genotyping a number of markers across almost any genome of interest. SNPs (single nucleotide polymorphism) markers referring to a single change between base pairs of a gene attracted many scientists in molecular genetics due to their abundance in the genomes and amenability for high-throughput platforms (Mammadov *et al.* 2012). The development of advanced high-throughput sequencing technology produces thousands of SNP markers for constructing high-density genetic maps. The higher resolution of the genetic map allows greater precision of QTL mapping. Genotyping-by-sequencing (GBS) provides a robust sequencing platform that generates high-quality SNP markers (Elshire *et al.* 2011; Nelson *et al.* 2011). These SNP markers play important roles in high-density map construction which have been studied and proved to be powerful and accurate for QTL mapping in agriculturally important traits (Yu *et al.* 2011; Wang *et al.* 2011a).

QTL mapping

Quantitative traits refer to traits measured numerically and controlled by intermediate and small quantitative trait loci (QTL), in contrast to qualitative traits that are controlled by one or a few genes. QTL mapping aims to identify a statistically significant relationship between phenotypic traits and DNA markers that segregate via chromosome recombination during meiosis. Genetic markers that are close together or tightly linked are passed on together from parent to progeny more frequently than genes or markers that are unlinked. Therefore, QTL and markers that are inherited together in the progeny allow the linkage of particular phenotypes with specific chromosomal regions. On the other hand, unlinked markers located far apart from the QTL are randomly inherited, indicating that there are no significant differences between means of the genotype groups. The method is commonly used for discovering genetic regions that influence phenotypic traits of interest in many crops including maize (Boer *et al.* 2007; Yang *et al.* 2020), barley (*Hordeum vulgare* L.) (Yin *et al.* 2005), rice (Wan *et al.* 2008; Marathi *et al.* 2012) and sorghum (Takanashi *et al.* 2021).

The first milestone in QTL mapping is the construction of a mapping population where the trait of interest is segregating. The parents selected for the mapping population differ for one or more traits of interest. A wide range of population structures such as backcross (BC), F_2, recombinant inbred lines (RIL), and doubled haploid (DH) is commonly used for QTL mapping. A suitable population typically requires at least 200 individuals for BC_1/F_2 populations in studies of different taxa and phenotypes (Paterson 2002b). There are three common methods for detecting QTLs, namely single-marker analysis, simple interval mapping and composite interval mapping. Single-marker analysis, the simplest method, utilizes single markers to detect their association with QTLs. This method relies on ANOVA, linear regression, and/or t-tests as the main statistical tools. Since the method does not always require a genetic map to detect the association between markers and phenotypes, the drawback for this method is that QTLs distant from a marker will be less likely to be detected. The simple interval mapping (SIM) method applies linkage maps and intervals between adjacent pairs of linked markers along chromosomes. This method is statistically more powerful since it incorporates linked markers for analysis and compensates for recombination between markers and the QTL. The composite interval mapping (CIM) method combines the advantages from the previous two methods. The method not only applies interval mapping with linear regression but also uses flanking marker information for the analysis. In this study, CIM method will be applied to identify single QTLs to further detect their interactions based on a statistical model (Collard *et al.* 2005).

Association mapping in sorghum

Association mapping, also known as genome-wide association study (GWAS), is based on the linkage disequilibrium (LD) within the populations of study. Linkage disequilibrium refers to the extent to which one SNP allele is inherited or associated with another SNP allele within a population. The rate of linkage disequilibrium decay is affected by the number of chromosomes in the population and the number of generations the population has existed (Bush and Moore 2012; Tam *et al.* 2019). GWAS allows the examination of an association of genetic influence on phenotypic traits. From the analysis conducted using computational software, a regression of the statistical test will be generated to identify genetic associations with phenotypic characteristics. In general, SNPs, single changes between base pairs in the DNA sequence within a genome, are commonly used as markers due to their small impact on biological systems within a genomic region and their abundance form of genetic variation in a genome. Typically, SNPs carry two forms of alleles, hence there are two possibilities of occurring base-pair for a SNP location within a population. The less common allele for the SNP location is known as ‘minor allele frequency’ (Bush and Moore 2012; Cano-Gamez and Trynka 2020).

GWAS is a powerful tool that is used for determining the genotypic variations and phenotypic diversity associations by identifying genetic regions that impact the trait of interest based on the natural variation from a population. The association mapping leverages genetic regions and phenotypic trait associations in many crops including maize (Mazaheri *et al.* 2019; Rashid *et al.* 2020), barley (*Hordeum vulgare* L.) (Cockram *et al.* 2010), rice (Yano *et al.* 2019; Bheemanahalli *et al.* 2021), wheat (Zanke *et al.* 2015; Bhatta *et al.* 2018), and sorghum (Morris *et al.* 2013a; Boyles *et al.* 2016). The markers in tight linkage disequilibrium with significant marker responsible for the phenotypic trait present significant association with the trait. While QTL analysis requires biparental crosses for constructing a mapping population, GWAS approach requires a large, diverse, and unrelated collection of samples. For example, the diversity panel population in the study derived from several published sorghum GBS datasets. Morris et al (2013) obtained 971 accessions from diverse sorghum germplasm from worldwide collections combining three diversity panels: the US sorghum association (SAP), the sorghum mini core collections (MC), and the Generation Challenge Program reference set (RS). The Sorghum Bioenergy Association Panel (BAP) was also genotypes with a total of 390 accessions obtaining 232,303 SNPs (Brenton *et al.* 2016). Lasky et al., 2015 assembled 1,943 sorghum georeferenced landrace lines with 404,627 SNP markers.

Biparental crosses have restricted allelic diversity and limited genomic resolution. On the other hand, a diverse population for GWAS analysis holds a large number of recombination events in the genetic history of the population and provides higher resolution (Platt *et al.* 2010; Brachi *et al.* 2011; Boyles *et al.* 2016). Utilizing mapping techniques, the variation of phenotypic traits in plants is directly linked to the underlying causal loci. To investigate phenotypic and genotypic differences, bi-parental QTL mapping populations (linkage mapping) or association mapping (GWAS analysis) of unrelated individuals are utilized. Thus, both mapping strategies aim to identify molecular markers associated with QTL (Alqudah *et al.* 2020).

Comparative analysis

Despite constructing a QTL map with a large population to negate the effects of statistical artifacts from sampling, the limited polymorphic loci between the two parents will influence the mapping accuracy (Holland 2007). In contrast, GWAS does not require constructing specific mapping populations, instead using high recombination of genes in natural populations, compared to linkage mapping. GWAS has been employed in the detection of quantitative loci by directly identifying associations between DNA markers and phenotypes in populations based on LD. However, the population structure in association analysis can produce a stronger LD between non-linked loci due to genetic drift and natural selection. Therefore, performing both association mapping and linkage mapping mitigates the false positives from associated loci due to high LD as well as establishing fine mapping of QTL intervals (He *et al.* 2017).

Several studies have reported the application of both QTL and GWAS mapping in plants. One study utilized GWAS analysis between significant SNPs and previously identified QTL-controlled sorghum architecture. The results including 3 overlaps (flowering time on chromosome 10 and stem circumference on chromosome 4 & 7) as well as 6 novel regions (flowering time on chromosome 8, internode number on chromosomes 6 & 8, panicle exertion on chromosome 6, panicle length on chromosome 3, and seed number on chromosome 6) were determined (Zhao *et al.* 2016). In *Brassica napus*, the identification of target regions controlling branch number by linkage mapping and association mapping was reported. The QTL region responsible for branch number through linkage mapping in rapeseed was also identified. Subsequently, a GWAS analysis was performed to verify the QTL region and to narrow down the QTL within a 1.51 Mb interval on chromosome C03. The GWAS analysis reduced the target region size and will aid in further identifying the candidate gene (He et al, 2017). Another study performed both QTL mapping and GWAS to study comparative genetics of seed size traits in sorghum and rice. Two regions located on the chromosome Sb04 and Sb10 were detected as gene candidates for sorghum seed size. Each region also significantly corresponded to seed size in rice (Zhang *et al.* 2015).

Discovering the genetic components of leaf morphology in sorghum

Leaves have a major role regarding plant anatomical and physiological functions. Leaves are the primary sites for photosynthesis, a process by which plants convert light energy to chemical energy using chlorophyll found within a leaf. The leaf absorbs energy form sunlight to oxidize water and produce sugar from carbon dioxide while releasing oxygen to the air. Accordingly, leaf architecture impacts plant growth and yield directly (Mantilla-Perez and Salas Fernandez 2017).

In sorghum, the effects of leaf angle not only determine the plant's efficiency of light interception but also affect the capacity of planting density (Mantilla-Perez and Salas Fernandez 2017). Wide leaf angle increases leaf shading leading to negative effects on photosynthesis under high plant density. Conversely, narrow leaf angle allows plants to grasp more sunlight leading to higher yield production (Kenchanmane Raju *et al.* 2020). In addition to leaf angle, another study also investigated the correlation between the green fodder yield of sorghum and plant architecture, including leaf length and leaf breadth in sorghum. The report observed that leaf length had positive correlation with leaf breadth (Prakash *et al.* 2010). Thus, leaf architecture greatly impacts plant growth and yield, positively affecting its morphology and physiology.

Leaf morphology in related grasses

Sorghum shares its evolutionary history with multiple cereal crops and grasses. Maize is a relatively close genetic relative to sorghum since both of them are from the same subfamily, called Panicoideae. Other members of the grass family (Poaceae) such as finger millet, bamboo, rice, wheat, and oat also share common ancestry with sorghum (Kellogg 1998). Although sorghum is an important crop, the topic of its leaf morphology has not been studied much compared to other model crops such as rice and maize. A number of studies have genetically dissected and analyzed genes and QTLs that have contributed to leaf architecture traits in both rice and maize. This section provides comparative literature reviews focusing on the studies of leaf architecture within the two crops, rice and maize respectively.

Rice (*Oryza sativa*) is considered as a staple food which more than half of the world's population consume regularly. The *NARROW LEAF 1* (*NAL1*) gene, located on chromosome 4, was found to impact the leaf morphology in rice. The gene affects various characteristics such as narrow leaf, the width of the flag leaf, total spikelet number per panicle, photosynthetic rate, and chlorophyll content. The *NAL1* gene exhibits genotypic variations that are associated with plant morphology depending on the rice accessions. A study reported that the *NAL1* allele in 'Koshihikari' (a temperate japonica cultivar) decreased the thickness of the flag leaf and increased the ratio of leaf area to dry mass. However, another study showed that the *NAL1* allele in 'Daringan' (a tropical japonica landrace) resulted in increased flag leaf width, more vascular bundles, greater root biomass, more spikelets, and increased grain yield per square meter (Taguchi-Shiobara *et al.* 2015). Ham et al. (2019) performed QTL analysis related to the flag leaf angle associated with photosynthetic efficiency and chlorophyll contents in a doubled haploid (DH) population of 120 lines, derived from a cross between ‘Cheongcheong’ and ‘Nagdong’ accessions. Four QTLs with LOD scores greater than 3 were detected for the investigated traits, including *qFA4* and *qFA11* from flag-leaf angle, *qPE3* from photosynthetic efficiency and *qCC11* from chlorophyll content. These QTLs were located on chromosomes 3, 4, and 11 respectively. In another study, the QTL *qFL1* was shown to influence a large proportion of the variation in flag leaf size (leaf length, width, and area) in populations derived from tow elite parental lines (Zhenshan 97 and 93-11). A QTL *qFL1* was detected using a large segregation population. The QTL was narrowed down to a 31 kb region on chromosome 1 based on advanced backcrossed populations (BC_2F_2 and BC_3F_2) derived from BRIL of Zhenshan 97 and 93-11. The study revealed that the QTL regions *qFL1* and *qFW1* were tightly linked for FL and FW. However, *qFL1/1sLL1* and *qsLW1* were independent for sLL and sLW. In fact, *qFL1* had a pleiotropic effect on flag leaf size and yield-related traits (Wang *et al.* 2011b).

Maize (*Zea mays*) is another important crop not only for human consumption but also for animal feed throughout the globe. A number of genes affecting leaf morphology have been identified in maize. For example, *NS1* (*Narrow sheath 1*) and *NS2* (*Narrow sheath 2*) perform redundant functions in maize leaf development causing extremely narrow sheath leaves. *DIL1* (*Dwarf and Irregular Leaf 1*) influences plant height, leaf width and length. The leaf width in maize is not controlled by the aforementioned qualitative genes but also by QTLs. Wang et al., (2018) conducted QTL mapping for leaf width in maize using a large RIL population and bin mapping technique. The study revealed *qLW4* to map to a 55-kb interval on chromosome 4. The QTL, *qLW4*, was described as a putative major effect QTL with dominant effect on leaf width, with no additional effect on leaf length. Another study by Zhang & Huang (2021) focused on maize leaf angle. A total of eight QTLs were detected on chromosomes 1, 2, 3, 4, and 8. Each QTL contributed to phenotypic variance, with additive effects ranging from 4.3 to 14.2 % of the leaf angle variance. A heterogeneous inbred family (HIF) and whole genome sequencing techniques were applied to confirm the validity of QTL as well as the potential candidate genes of the QTL regions. Ku et al. (2012) utilized QTL mapping and meta-QTL analysis to detect 21 QTLs and 17 meta-QTLs for leaf architecture in maize. The leaf orientation value (LOV) trait had the highest number of QTLs, six, on chromosomes 1, 2, 3, 7, 8, and 9. These QTLs contributed cumulatively 64.63% of the phenotypic variance. The five QTLs associated with leaf angle (LA) were located on chromosomes 1, 2, 3, 7, and 8 and contributed cumulatively 60.30% of the phenotypic variance. The five QTLs associated with leaf length (LL) were located on chromosomes 3, 5, and 7 and contributed cumulatively 53.16% of the phenotypic variance. The five QTLs associated with leaf width (LW) were located on chromosomes 1, 2, 7, and 8 and contributed cumulatively 34.13% of the phenotypic variance. In addition to QTLs, mQTLs were identified to study their association with the variation of multiple leaf morphology traits. The mQTL3-3 and mQTL7-2, located on chromosomes 3 and 7 respectively, controlled all four traits of interest (LA, LOV, LL, and LW). The mQTLs that controlled three out of four traits of interest include mQTL5-1 on chromosome 5 that influenced LA, LOV and LL as well as an mQTL on chromosome 9 that contributed to LA, LOV and LW phenotypic variances. Five mQTLs for LA and LOV were detected on chromosomes 1, 2, 3, and 8. Additionally, an mQTL on chromosome 3 controlled LA and LL and two mQTLs on chromosomes 1 and 4 controlled LA and LW. Two mQTLs for LA were detected on chromosomes 2 and 7, one mQTL controlling LOV was detected on chromosome 9, and one mQTL for LL was detected on chromosome 7.

Summary

Sorghum is an important resource for many needs including human food, animal feed, and bioenergy. The leaf plays an important role in plant growth and development, producing nutrition (sugar) through biochemical pathways using light, water, and carbon dioxide. Sorghum is able to produce higher yields than many other crops in non-favorable conditions such as heat and water stress. Identification of genomic regions controlling major leaf morphology traits in sorghum would provide information of potential value for improving aspects of productivity such as biomass production. In the study, we dissected the genetic control of morphological traits such as leaf length, leaf angle, leaf width, and diameter of midribs as well as yield-related traits such as plant height, flowering days, dry stalk weight, and dry leaf weight. Moreover, advancement in sequencing technology allows cheaper and more accessible choices for genotyping. In the study, genotyping-by-sequencing (GBS) is the method used for mapping the QTLs and the diversity panel. Identification of major and significant QTLs associated with different leaf morphology related traits may assist in molecular breeding and improvement of sorghum production and adaptation.

References

Alqudah, A. M., A. Sallam, P. Stephen Baenziger, and A. Börner, 2020 GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. J. Adv. Res. 22: 119.

Altshuler, D., M. J. Daly, and E. S. Lander, 2008 Genetic Mapping in Human Disease. Science 322: 881.

Ananda, G. K. S., H. Myrans, S. L. Norton, R. Gleadow, A. Furtado et al., 2020 Wild Sorghum as a Promising Resource for Crop Improvement. Front. Plant Sci. 11: 1–14.

Anglani, C., 1998 Sorghum for human food - A review. Plant Foods Hum. Nutr. 52: 85–95.

Bhatta, M., A. Morgounov, V. Belamkar, and P. S. Baenziger, 2018 Genome-Wide Association Study Reveals Novel Genomic Regions for Grain Yield and Yield-Related Traits in Drought-Stressed Synthetic Hexaploid Wheat. Int. J. Mol. Sci. 19:.

Bheemanahalli, R., M. Knight, C. Quinones, C. J. Doherty, and S. V. K. Jagadish, 2021 Genome-wide association study and gene network analyses reveal potential candidate genes for high night temperature tolerance in rice. Sci. Reports 2021 111 11: 1–17.

Boer, M. P., D. Wright, L. Feng, D. W. Podlich, L. Luo et al., 2007 A Mixed-Model Quantitative Trait Loci (QTL) Analysis for Multiple-Environment Trial Data Using Environmental Covariables for QTL-by-Environment Interactions, With an Example in Maize. Genetics 177: 1801.

Boivin, K., M. Deu, J. F. Rami, G. Trouche, and P. Hamon, 1999 Towards a saturated sorghum map using RFLP and AFLP markers. Theor. Appl. Genet. 1999 982 98: 320–328.

Boyles, R. E., E. A. Cooper, M. T. Myers, Z. Brenton, B. L. Rauh et al., 2016 Genome-Wide

Association Studies of Grain Yield Components in Diverse Sorghum Germplasm. Plant Genome 9: plantgenome2015.09.0091.

Brachi, B., G. P. Morris, and J. O. Borevitz, 2011 Genome-wide association studies in plants: The missing heritability is in the field. Genome Biol. 12:.

Brenton, Z. W., E. A. Cooper, M. T. Myers, R. E. Boyles, N. Shakoor et al., 2016 A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. Genetics 204: 21–33.

Burgarella, C., A. Berger, S. Glémin, J. David, N. Terrier et al., 2021 The Road to Sorghum Domestication: Evidence From Nucleotide Diversity and Gene Expression Patterns. Front. Plant Sci. 12:.

Bush, W. S., and J. H. Moore, 2012 Chapter 11: Genome-Wide Association Studies. PLoS Comput. Biol. 8:.

Cano-Gamez, E., and G. Trynka, 2020 From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. Front. Genet. 11: 424.

Chittenden, L. M., K. F. Schertz, Y. R. Lin, R. A. Wing, and A. H. Paterson, 1994 A detailed RFLP map of Sorghum bicolor x S. propinquum, suitable for high-density mapping, suggests ancestral duplication of Sorghum chromosomes or chromosomal segments. Theor. Appl. Genet. 87: 925–933.

Cockram, J., J. White, D. L. Zuluaga, D. Smith, J. Comadran et al., 2010 Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. Proc. Natl. Acad. Sci. U. S. A. 107: 21611–21616.

Collard, B. C. Y., M. Z. Z. Jahufer, J. B. Brouwer, and E. C. K. Pang, 2005 An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142: 169–196.

Dillon, S. L., F. M. Shapter, R. J. Henry, G. Cordeiro, L. Izquierdo et al., 2007 Domestication to crop improvement: Genetic resources for Sorghum and Saccharum (Andropogoneae). Ann. Bot. 100: 975–989.

Doggett, H., 1998 Sorghum, 2nd edition. Longman, UK

Ejeta, G., and J. E. Knoll, 2007 Marker-assisted selection in sorghum. Genomics-Assisted Crop Improv. 2: 187–205.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto et al., 2011 A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS One 6: e19379.

Guan, Y. an, H. lian Wang, L. Qin, H. wen Zhang, Y. bing Yang et al., 2011 QTL mapping of bio-energy related traits in Sorghum. Euphytica 182: 431–440.

Harlan, J. R., and J. M. J. de Wet, 1972 A Simplified Classification of Cultivated Sorghum1. Crop Sci. 12: 172–176.

He, Y., D. Wu, D. Wei, Y. Fu, Y. Cui et al., 2017 GWAS, QTL mapping and gene expression analyses in Brassica napus reveal genetic control of branching morphogenesis. Sci. Rep. 7: 1–9.

Hodnett, G. L., S. Ohadi, N. A. Pugh, M. V. Bagavathiannan, and W. L. Rooney, 2019 Sorghum bicolor x S. halepense interspecific hybridization is influenced by the frequency of 2n gametes in S. bicolor. Sci. Rep. 9:.

Holland, J. B., 2007 Genetic architecture of complex traits in plants. Curr. Opin. Plant Biol. 10: 156–161.

Hulbert, S. H., T. E. Richter, J. D. Axtell, and J. L. Bennetzen, 1990 Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. Proc. Natl. Acad. Sci. U. S. A. 87: 4251–4255.

Kanbar, A., E. Shakeri, D. Alhajturki, T. Horn, Y. Emam et al., 2019 Morphological and molecular characterization of sweet, grain and forage sorghum (Sorghum bicolor L.) genotypes grown under temperate climatic conditions. https://doi.org/10.1080/11263504.2019.1569568 154: 49–58.

Kellogg, E. A., 1998 Relationships of cereal crops and other grasses. Proc. Natl. Acad. Sci. U. S. A. 95: 2005–2010.

Kenchanmane Raju, S. K., M. Adkins, A. Enersen, D. Santana de Carvalho, A. J. Studer et al., 2020 Leaf Angle eXtractor: A high-throughput image processing framework for leaf angle measurements in maize and sorghum. Appl. Plant Sci. 8: 1–9.

Kirungu, J. N., Y. Deng, X. Cai, R. O. Magwanga, Z. Zhou et al., 2018 Simple Sequence Repeat (SSR) Genetic Linkage Map of D Genome Diploid Cotton Derived from an Interspecific Cross between Gossypium davidsonii and Gossypium klotzschianum. Int. J. Mol. Sci. 19:.

Klein, P., and C. M. Smith, 2021 Invasive Johnsongrass, a threat to native grasslands and agriculture. Biologia (Bratisl). 76: 413–420.

Kong, W., H. Jin, C. D. Franks, C. Kim, R. Bandopadhyay et al., 2013 Genetic Analysis of Recombinant Inbred Lines for Sorghum bicolor × Sorghum propinquum. G3 Genes|Genomes|Genetics 3: 101–108.

Kong, W., C. Kim, V. H. Goff, D. Zhang, and A. H. Paterson, 2015 Genetic analysis of rhizomatousness and its relationship with vegetative branching of recombinant inbred lines of Sorghum bicolor × S. propinquum . Am. J. Bot. 102: 718–724.

Ku, L. X., J. Zhang, S. L. Guo, H. Y. Liu, R. F. Zhao et al., 2012 Integrated multiple population analysis of leaf architecture traits in maize (Zea mays L.). J. Exp. Bot. 63: 261–274.

Lai, X., L. Yan, Y. Lu, and J. C. Schnable, 2018 Largely unlinked gene sets targeted by selection for domestication syndrome phenotypes in maize and sorghum. Plant J. 93: 843–855.

Lasky, J. R., H. D. Upadhyaya, P. Ramu, S. Deshpande, C. T. Hash et al., 2015 Genome-environment associations in sorghum landraces predict adaptive traits. Sci. Adv. 1:.

Lazarides, M., J. B. Hacker, and M. H. Andrew, 1991 Taxonomy, cytology and ecology of indigenous Australian sorghums (Sorghum Moench: Andropogoneae: Poaceae). Aust. Syst. Bot. 4: 591–635.

Mammadov, J., R. Aggarwal, R. Buyyarapu, and S. Kumpatla, 2012 SNP markers and their impact on plant breeding. Int. J. Plant Genomics 2012:.

Mantilla-Perez, M. B., and M. G. Salas Fernandez, 2017 Differential manipulation of leaf angle throughout the canopy: Current status and prospects. J. Exp. Bot. 68: 5699–5717.

Marathi, B., S. Guleria, T. Mohapatra, R. Parsad, N. Mariappan et al., 2012 QTL analysis of novel genomic regions associated with yield and yield related traits in new plant type based recombinant inbred lines of rice (Oryza sativa L.). BMC Plant Biol. 12: 1–19.

Mathur, S., A. V. Umakanth, V. A. Tonapi, R. Sharma, and M. K. Sharma, 2017 Sweet sorghum as biofuel feedstock: Recent advances and available resources. Biotechnol. Biofuels 10: 1–19.

Mazaheri, M., M. Heckwolf, B. Vaillancourt, J. L. Gage, B. Burdo et al., 2019 Genome-wide association analysis of stalk biomass and anatomical traits in maize. BMC Plant Biol. 19: 1–17.

Menz, M. A., R. R. Klein, N. C. Unruh, W. L. Rooney, P. E. Klein et al., 2004 Genetic Diversity of Public Inbreds of Sorghum Determined by Mapped AFLP and SSR Markers. Crop Sci. 44: 1236–1244.

Morris, G. P., P. Ramu, S. P. Deshpande, C. T. Hash, T. Shah et al., 2013 Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc. Natl. Acad. Sci. U. S. A. 110: 453–458.

Nelson, J. C., S. Wang, Y. Wu, X. Li, G. Antony et al., 2011 Single-nucleotide polymorphism discovery by high-throughput sequencing in sorghum. BMC Genomics 12: 1–15.

Paterson, A. H., 2002a What has QTL mapping taught us about plant domestication? New Phytol. 154: 591–608.

Paterson, A. H., 2002b What has QTL mapping taught us about plant domestication? New Phytol. 154: 591–608.

Paterson, A. H., W. Q. Kong, R. M. Johnston, P. Nabukalu, G. Wu et al., 2020 The Evolution of an Invasive Plant, Sorghum halepense L. (‘Johnsongrass’). Front. Genet. 11: 1–10.

Peng, Y., K. F. Schertz, S. Cartinhour, and G. E. Hart, 1999 Comparative genome mapping of Sorghum bicolor (L.) Moench using an RFLP map constructed in a population of recombinant inbred lines. Plant Breed. 118: 225–235.

Pereira, M. G., M. Lee, P. Bramel-Cox, W. Woodman, J. Doebley et al., 1994 Construction of an RFLP map in sorghum and comparative mapping in maize. Genome 37: 236–243.

Platt, A., B. J. Vilhjálmsson, and M. Nordborg, 2010 Conditions under which genome-wide association studies will be positively misleading. Genetics 186: 1045–1052.

Prakash, R., K. Ganesamurthy, A. Nirmalakumari, and P. Nagarajan, 2010 Correlation and path analysis in sorghum (Sorghum bicolor L. Moench). Electron. J. Plant Breed. 1: 315–318.

Rashid, Z., M. Sofi, S. I. Harlapur, R. M. Kachapur, Z. A. Dar et al., 2020 Genome-wide association studies in tropical maize germplasm reveal novel and known genomic regions for resistance to Northern corn leaf blight. Sci. Reports 2020 101 10: 1–16.

Taguchi-Shiobara, F., T. Ota, K. Ebana, T. Ookawa, M. Yamasaki et al., 2015 Natural variation in the flag leaf morphology of rice due to a mutation of the NARROW LEAF 1 gene in Oryza sativa L. Genetics 201: 795–808.

Takanashi, H., M. Shichijo, L. Sakamoto, H. Kajiya-Kanegae, H. Iwata et al., 2021 Genetic dissection of QTLs associated with spikelet-related traits and grain size in sorghum. Sci. Reports 2021 111 11: 1–16.

Tam, V., N. Patel, M. Turcotte, Y. Bossé, G. Paré et al., 2019 Benefits and limitations of genome-wide association studies. Nat. Rev. Genet. 2019 208 20: 467–484.

Tao, Y. Z., R. G. Henzell, D. R. Jordan, D. G. Butler, A. M. Kelly et al., 2000 Identification of genomic regions associated with stay green in sorghum by testing RILs in multiple environments. Theor. Appl. Genet. 100: 1225–1232.

Tao, Y. Z., D. R. Jordan, R. G. Henzell, and C. L. McIntyre, 1998 Construction of a genetic map in a sorghum recombinant inbred line using probes from different sources and its comparison with other sorghum maps. Aust. J. Agric. Res. 49: 729–736.

Venkateswaran, K., M. Elangovan, and N. Sivaraj, 2019 Origin, Domestication and Diffusion of Sorghum bicolor. Breed. Sorghum Divers. End Uses 15–31.

Wan, X., J. Weng, H. Zhai, J. Wang, C. Lei et al., 2008 Quantitative Trait Loci (QTL) Analysis

For Rice Grain Width and Fine Mapping of an Identified QTL Allele gw-5 in a Recombination Hotspot Region on Chromosome 5. Genetics 179: 2239.

Wang, L., A. Wang, X. Huang, Q. Zhao, G. Dong et al., 2011a Mapping 49 quantitative trait loci at high resolution through sequencing-based genotyping of rice recombinant inbred lines. Theor. Appl. Genet. 122: 327–340.

Wang, P., G. Zhou, H. Yu, and S. Yu, 2011b Fine mapping a major QTL for flag leaf size and yield-related traits in rice. Theor. Appl. Genet. 123: 1319–1330.

Wang, B., Y. Zhu, J. Zhu, Z. Liu, H. Liu et al., 2018 Identification and fine-mapping of a major maize leaf width QTL in a Re-sequenced large recombinant inbred lines population. Front. Plant Sci. 9: 1–11.

de Wet, J. M. J., and J. P. Huckabay, 1967 The Origin of Sorghum bicolor. II. Distribution and Domestication. Evolution (N. Y). 21: 787.

Winchell, F., C. J. Stevens, C. Murphy, L. Champion, and D. Q. Fuller, 2017 Evidence for sorghum domestication in fourth millennium BC Eastern Sudan: Spikelet morphology from ceramic impressions of the Butan group. Curr. Anthropol. 58: 673–683.

Wu, Y. Q., and Y. Huang, 2007 An SSR genetic map of Sorghum bicolor (L.) Moench and its comparison to a published genetic map. Genome 50: 84–89.

Xin, Z., M. Wang, H. E. Cuevas, J. Chen, M. Harrison et al., 2021a Sorghum genetic, genomic, and breeding resources. Planta 254: 114.

Xin, Z., M. Wang, H. E. Cuevas, J. Chen, M. Harrison et al., 2021b Sorghum genetic, genomic, and breeding resources. Planta 254: 1–24.

Yang, J., Z. Liu, Q. Chen, Y. Qu, J. Tang et al., 2020 Mapping of QTL for Grain Yield Components Based on a DH Population in Maize. Sci. Reports 2020 101 10: 1–11.

Yano, K., Y. Morinaka, F. Wang, P. Huang, S. Takehara et al., 2019 GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture. Proc. Natl. Acad. Sci. U. S. A. 116: 2162–21267.

Yin, X., P. C. Struik, F. A. Van Eeuwijk, P. Stam, and J. Tang, 2005 QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. J. Exp. Bot. 56: 967–976.

Yu, H., W. Xie, J. Wang, Y. Xing, C. Xu et al., 2011 Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. PLoS One 6:.

Zanke, C. D., J. Ling, J. Plieske, S. Kollers, E. Ebmeyer et al., 2015 Analysis of main effect QTL for thousand grain weight in European winter wheat (Triticum aestivum L.) by genome-wide association mapping. Front. Plant Sci. 6: 644.

Zhang, N., and X. Huang, 2021 Mapping quantitative trait loci and predicting candidate genes for leaf angle in maize. PLoS One 16: 1–15.

Zhang, D., J. Li, R. O. Compton, J. Robertson, V. H. Goff et al., 2015 Comparative Genetics of seed size traits in divergent cereal lineages represented by sorghum (Panicoidae) and Rice (Oryzoidae). G3 Genes, Genomes, Genet. 5: 1117–1128.

Zhao, J., M. B. Mantilla Perez, J. Hu, and M. G. Salas Fernandez, 2016 Genome-Wide Association Study for Nine Plant Architecture Traits in Sorghum. Plant Genome 9:.

CHAPTER 2

GENETIC ANALYSIS OF LEAF MORPHOLOGY AND YIELD-RELATED TRAITS IN SORGHUM POPULATIONS

Ployaram W.[1], W. Kong, and Andrew H. Paterson.

To be submitted to *G3: Genes / Genomes / Genetics*

Abstract

Sorghum is an important C_4 crop produced for the production of grain, fodder, sugar, and bioenergy. The objective of this research is to determine the relationship between genes and leaf morphology traits including leaf length, leaf width, leaf angle, and midrib diameter in sorghum populations. To link traits of sorghum leaves to DNA marker data, QTL mapping was performed in two recombinant inbred line (RIL) populations using the composite interval mapping (CIM) method, complemented by genome-wide association (GWAS) analysis in a diversity panel. A total of 101 QTLs were identified in the ISRIL population, and 20 in the PQRIL population. QTL mapping revealed 12 colocalized intervals on chromosome 7 for plant height, dry biomass, dry stalk weight, leaf area, leaf angle, and leaf width from two environments. Overall, LOD score peaks ranging from 4.5 to 31.0 show that the genomic loci remained stable and unaffected by the environments. A substantial stable QTL for PH (*qPHT6.1*) with a LOD score of 21.0 in our study, accounting for 21.47% of the phenotypic variation, was supported by previously published data. GWAS analysis identified SNP markers associated with six out of eight traits. While QTL mapping minimizes false positive associations, GWAS improves precision in determining candidate QTL regions. Well-supported QTLs will be a useful resource for further sorghum breeding via marker-assisted selection.

Introduction

Sorghum (*Sorghum bicolor* (L.) Moench) is the fifth most important crop in the world, cultivated for multiple purposes including food, feed, forage, and fuel. Because of its C_4 photosynthesis pathway, sorghum is advantageous over many other crops under unfavorable conditions such as drought and heat. The C_4 photosynthesis pathway allows plants to accumulate carbon dioxide more efficiently with reduced water usage under high temperatures and light conditions (Mathur *et al.* 2017). Moreover, sorghum outperforms other C_4 plants such as maize and sugarcane, due to lower water and fertilizer requirements (Rooney *et al.* 2007). Hence, sorghum is a great candidate to solve challenges that need to be overcome in order to improve agriculture, including enhancing food production sustainably in order to feed a growing global population as well as a growing demand for water resources (Bouman 2007; Drewry *et al.* 2014).

Leaves are the primary sites for photosynthesis, a process by which plants convert light energy to chemical energy using chlorophyll found within a leaf. The leaf absorbs energy form sunlight to oxidize water and produce sugar from carbon dioxide while releasing oxygen to the air (Mantilla-Perez & Salas Fernandez, 2017). Leaf morphology is one of the important traits determining plant growth and yield production since leaves have a major role regarding plant anatomical and physiological functions (Tsukaya 2004; Fu *et al.* 2019). The effects of leaf angle in sorghum not only regulate the plant's efficiency of light absorption, but also influence the plant's capacity for dense planting (Mantilla-Perez & Salas Fernandez, 2017). Higher planting densities and grain yields are possible with a narrow leaf angle, which has been the subject of positive selection in agricultural systems. This is probably due to the benefits of better RUE and increased nitrogen content of canopies (Drewry *et al.* 2014; Warnasooriya and Brutnell 2014; Truong *et al.* 2015). Under conditions of high plant density, a broad leaf angle causes leaf shading, which reduces photosynthesis. Positive selection may have favored broad leaf angles in nature, as their capacity to block light and nutrients from neighboring plants reduced competition for scarce resources (Schmitt *et al.* 2003; Drewry *et al.* 2014). In contrast, narrow leaf angle permits plants to absorb more sunlight, resulting in greater yield production (Kenchanmane Raju *et al.* 2020). Canopy architecture is also affected by leaf width, which in turn influences the transmission of photosynthetically active radiation and light signals. There is a tradeoff between the leaf width and light capture. Leaves that are extremely narrow are beneficial for light transparency but greatly limit light capture. However, extremely wide leaves reduce light transmitted to the middle and lower leaves leading to a decrease in the overall photosynthetically active radiation, causing shading avoidance (Wang *et al.* 2018). Accordingly, leaf architecture impacts plant growth and yield. Understanding the genetic factors influencing leaf morphology may allow genetic improvement of yield and adaptation in sorghum. In fact, sorghum has a relatively small genome size of ~730 Mb which makes the crop an attractive model to study functional genomics of Saccharine and other C_4 grasses (Paterson *et al.* 2009).

This study aims to identify genomic regions affecting traits related to leaf morphology of sorghum in two populations, IS and PQRIL, that span much of the genetic diversity available in 'eusorghums'. Both populations share a common parent, *S. bicolor* BTx623, an elite inbred that was the source of the sorghum reference genome. This particular parent was crossed with either *S. bicolor* IS3620, representing race 'guinea' which is highly divergent from BTx623; or *S. propinquum*, a sister species within the genus Sorghum to create IS and PQRIL populations respectively. Genetic analysis in the study was conducted via QTL mapping, GWAS and comparative analysis.

QTL mapping is a powerful tool for identifying genetic regions harboring genes associated with a trait of interest using a bi-parental population. GWAS analysis is another, complementary approach to unravel the molecular genetic basis underlying natural phenotypic variation (Alqudah *et al.* 2020). Biparental crosses used in QTL mapping have restricted allelic diversity and limited genomic resolution. On the other hand, a diverse population for GWAS analysis holds a large number of recombination events from the genetic history of the population and provides higher resolution (Platt *et al.* 2010; Brachi *et al.* 2011; Boyles *et al.* 2016); however, it has a high false-positive rate compared to QTL mapping (He *et al.* 2017). This research will facilitate both powerful tools; QTL mapping and GWAS analysis, providing finer QTL regions via comparative analysis leading to more precise regions which in turn leads to being able to identify candidate genes which may contribute to the traits of interest.

In this study, we described 20 QTLs and 101 QTLs, found in the PQRIL population and ISRIL population respectively. While GWAS increases the accuracy in identifying candidate QTL areas, QTL mapping reduces false positive relationships. The identification of candidate genes will be conducted via co-localized SNP markers in GWAS and QTL mapping, in addition to comparisons to genes from previously published research that are known to influence leaf morphological characteristics in sorghum and related grass species.

Materials and Methods

Plant materials for QTL analysis

Two different mapping populations, ISRIL and PQRIL populations were used in the study. The first population derived from *Sorghum bicolor* BTx623 crossed with *Sorghum bicolor* IS3620C, comprising 399 F7-8 RILs derived by selfing a single F1 plant as described in Kong et al (2018). The second population derived from *Sorghum bicolor* BTx623 crossed with *Sorghum propinquum* as described in Kong et al (2015). The mapping population comprised 161 F5 recombinant inbred lines derived from a controlled cross between single plants of *S. bicolor* BTx623 and *S. propinquum* (unnamed accession) by single seed descent from single F2 plants based on previous study (Paterson *et al.* 1995).

Development of mapping population and phenotyping for QTL analysis

Hereinafter, ISRIL and PQRIL will be used for representing the populations derived from *S. bicolor* BTx623 to *S. bicolor* IS3620C and *S. bicolor* BTx623 and *S. propinquum* respectively. Based on Kong et al (2018), the mapping population derived from ISRIL was planted in Watkinsville, GA at the University of Georgia Plant Science Farm in the summer of 2011 and 2012. Single 3-meter row plots of each population were machine-planted in a completely randomized design. Similarly, the mapping population for PQRIL was also planted at the University of Georgia Plant Science Farm, Watkinsville, Georgia in 2009, 2010, and 2011 for F5, F6 and F7 generations respectively according to Kong et al (2015). Single 1.5-meter plots of each RIL were transplanted in 2009 and 2011 or directly seeded in 2010 using a completely randomized design.

Leaf morphological traits for both populations were measured, including leaf length (LL), leaf width (LW), leaf angle (LA), and diameter of midribs (MR) of fully expanded and mature leaves (usually the fourth leaf below the flag). Also, biomass yield related traits were evaluated including plant height (PH), dry leaf weight (DLW), dry stalk weight (DSW), dry biomass (DB), and leaf area (LAR) from the ISRIL population; and plant height (PH) as well as days to flower (DF) in the PQRIL population. Plant height was measured at physiological maturity of each plant. Dry leaf weight and leaf area reflect overall production of sorghum leaves, allowing the investigation of the impact of leaf morphology on leaf yield. Dry biomass was measured from the vegetative parts, including leaves, stalks and inflorescence. The trait allows the study of leaf morphology effect on the sorghum yield. Flowering time is an important factor in sorghum biomass output because it determines the transition between vegetative and reproductive growth.

The length of vegetative development in sorghum is directly proportional to its increased biomass production (Childs *et al.* 1997).

The Pearson method and the cor() function of the R software were used to create the trait correlation matrix (R Core Development Team, 2013). The significance of each correlation was evaluated using the cor.test() function in R. The Correlation() function from the PerformanceAnalytics package was used to create scatter plots and histograms (Boyles et al. 2016). The lme4 software and the variation were used to calculate the broad-sense heritability.

Genotyping for QTL analysis

Samples were frozen at -80 C and lyophilized for 48 hr. DNA extraction was performed based on Aljanabi et al (1999) using leaf samples from the RIL populations. Two different methods were used for genotyping in each population. The F7 - F8 recombinant inbred line (RIL) population of 399 individuals was analyzed from the mapping population derived from *S. bicolor* BTx623 to *S. bicolor* IS3620C (ISRIL) as described in Kong et al (2018). The ISRIL population was genotyped via genotyping-by-sequencing (GBS) technique with the combination of a Multiplex Shotgun Genotyping (MSG) and the Tassel GBS analysis pipeline. A total of 7,103 raw SNPs were obtained using an Illumina Miseq with single-end sequencing. The reference genome sequence of *Sorghum bicolor* from Paterson et al (2009) was aligned for SNP discovery. Heterozygosity at a locus is called if two alleles are each inferred to be present at a probability greater than that of sequencing error. Raw SNP data from the TASSEL GBS pipeline were further filtered if the minor allele frequency is less than 5% or the proportion of missing genotypes greater than 40%. SNPs are further merged if the Pearson's correlation between them is larger than 0.95. Meanwhile, the 161 RILs from F5, F6, and F7 of the PQRIL mapping population was analyzed for the research as described in Kong et al (2015). The PQRIL population was genotyped for simple sequence repeats (SSR) as described (Kong *et al.* 2013).

Linkage map construction and QTL analysis

A genetic map for 393 individuals derived from the ISRIL population was constructed in Kong et al (2018) with a total of 381 bins of 616 GBS-based SNP markers. The genetic map was constructed using R/qtl (Broman *et al.* 2003) by merging markers within 1 cM for bin assignment for each chromosome. A genetic map for the PQRIL population of 161 individuals from three generations (F5, F6, and F7) derived from *S. bicolor* BTx623 to *S. propinquum* was generated from 141 SSR markers using MAPMAKER (Lander *et al.* 1987), covering a genetic distance of 773.1 cM, with an average interval between consecutive markers of 5.48 cM based on Kong et al (2015).

Composite interval mapping (CIM) was performed for QTL analysis using Windows QTL Cartographer version 2.5 (Wang et al., 2012). The standard model, forward and backward regression models were used to detect QTL regions (Model 6). The window size was 10 cM with walk speed of 1 cM. The LOD threshold values for each trait were determined by performing 1,000 permutations at a genome-wide significance level of 0.05 ($P < 0.05$). The putative QTLs were selected and declared significant based on the LOD score. Those QTLs with likelihood peaks that were at least 20 cM apart and did not share any overlapping genomic regions within 90% (1-LOD) likelihood intervals were considered distinct. The contribution rate (R^2) was estimated as the percentage of variance explained by each QTL in proportion to the total phenotypic variance. The QTLs were named following a nomenclature system that was previously described (McCouch et al., 1997; Kong et al., 2018). The name starts with a ‘q’ following by an abbreviation for each phenotypic trait of interest, then the chromosome number and a decimal number to differentiate multiple QTL on the same chromosome.

Genotyping and phenotyping for GWAS

The SNP dataset for the GWAS study was obtained from the Dryad Data Repository (doi:10.50601/dryad.63h8fd4) based on Hu et al (2019). This dataset includes a total of 459,304 SNPs of 10,323 sorghum genotypes from multiple sorghum diversity panels that was analyzed based on published sorghum studies (Morris *et al.* 2013, 2013b; Lasky *et al.* 2015; Brenton *et al.* 2016; Yu *et al.* 2016; Bouchet *et al.* 2017). The sorghum SNP dataset was constructed by combining multiple published GBS data from diverse landraces, breeding lines, and biparental mapping families, including the sorghum association panel, the bioenergy association panel, and NAM population.

The phenotypic dataset for this study was obtained separately from the genotypes. A total of 354 accessions from the sorghum diversity panel was planted in 2009 and 2010, near Watkinsville, Georgia. The leaf morphology and yield-related traits were evaluated in each year. Leaf morphological traits that were measured include leaf length (LL), leaf width (LW), leaf angle (LA), and diameters of midribs (MR). Biomass yield related traits were also evaluated from the diversity panel, including plant height (PH), dry leaf weight (DLW), dry stalk weight (DSW), and flowering days (DF). Only the genotypic data from the reference sorghum SNP dataset that was matched with the obtained phenotypes was analyzed for the study. Best linear unbiased estimations (BLUEs) for all phenotypic traits were calculated with polyqtlR package (Bourke *et al.* 2021) for combining the dataset from the two environments (2009 and 2010). A total of 405 accessions from the US sorghum association panel (SAP) (Casa et al., 2008) was used for performing GWAS analysis. The trait correlation matrix was constructed using the

Pearson method and the R software cor() function (R Core Development Team, 2013). Using the cor.test() function in R, the significance of each correlation was determined. The chart.Correlation() function within the PerformanceAnalytics package was used to create scatter plots and histograms (Boyles *et al.* 2016). The broad-sense heritability was determined using the lme4 program and the variation. Using the following model, the phenotypic variation captured by the population structure was analyzed:

$$y = PC1 + PC2 + PC3 + PC4 + PC5$$

where y is the phenotypic data and *PC1* to *PC5* are the first five principal components (PCs) (Hu et al., 2019).

Association analysis

The study utilized models available within TASSEL V.5 software (Bradbury *et al.* 2007), which are General Linear Model (GLM) and Mixed Linear Model (MLM). GLM involves only population structure (Q) to perform association analysis by testing association between segregating sites and phenotypes with fixed effects linear model. MLM involves both population structure (Q) and kinship matrix (K), including both fixed and random effects. Preliminary analyses were performed to determine which model was more suitable for association mapping analysis. Manhattan plots of MLM method showed low signals for association between SNP markers and phenotypic traits compared to GLM in which contrast to the results from QTL mapping. Therefore, the association analysis using a general linear model (GLM) with 1,000 permutations was performed in TASSEL where the first five PCs were used to account for population structure. Single nucleotide polymorphism (SNPs) with a minor allele frequency (MAF) greater than 0.05 were filtered. A total of 194,349 hypothesis tests were conducted for each trait. An empirical p-value ($P < 0.01$) from the permutations was used to determine the threshold of GWAS analysis. The significance of an association between SNPs and traits was determined when the P-value was less than the threshold value of 0.01. Manhattan and QQ plots were created with the qqman R package (Turner 2018).

Results

Phenotypic analysis

Descriptive statistics of ISRIL, PQRIL and the diversity panel are shown in Table 2.1 for leaf morphology related traits and Table 2.2 for yield related traits. ANOVA tests for each population demonstrated statistically significant ($P < 0.05$) effects of genotype, environment, and their interaction for the traits of interest. (Table 2.3-2.5). Of the three populations, ISRIL has the highest mean values for most traits, including leaf length, leaf angle, diameter of midribs, dry leaf weight, dry stalk weight, dry biomass and leaf area. In addition, the diversity panel shows the highest mean values in leaf width and plant height, whereas PQRIL contains the highest mean values for days to flowering.

Correlation coefficients between traits were different among the three populations (Table 2.6-2.8). Significant correlations between the leaf morphological and yield-related traits were observed. In ISRIL, dry leaf weight, leaf area and leaf length were significantly correlated with all other studied traits. Dry stalk weight and dry biomass were correlated with all other traits except leaf angle. Similarly, leaf width and diameter of midribs were correlated with all other traits except plant height (Table 2.6). No significant association was observed among plant height, leaf width, and diameter of midribs as well as among leaf angle, dry biomass, and dry stalk weight in ISRIL. In PQRIL, all traits were significantly correlated except leaf angle (Table 2.7). In the diversity panel, leaf length, width, and angle were significantly correlated to all other traits except dry stalk weight (Table 2.8). Correlation analysis of traits across different years shows significant moderate to weak correlation for leaf morphology and yield-related traits in all populations (Figure 2.5-2.7). In addition, the highest correlation coefficients were observed in the ISRIL between leaf area and leaf length (r = 0.88, $P < 0.001$) and between leaf area and leaf width (r = 0.88, $P < 0.001$). Additionally, strong correlations were exhibited between dry biomass and dry leaf weight (r = 0.89, $P < 0.001$) as well as dry biomass and dry stalk weight (r = 0.95, $P < 0.001$). Consequently, leaf area was significantly and positively correlated with both leaf length and leaf area, and dry biomass was also significantly and strongly correlated with dry leaf and stalk weight.

QTLs for leaf morphology traits

A total of 19 QTLs were identified for the four leaf morphology traits in the ISRIL population based on LS-means calculated across years (Table 2.9). Five QTLs (*qLL1.1*, *qLL1.2*, *qLL3.1*, *qLL8.1*, and *qLL9.1*) were identified to influence leaf length (LL). Four QTLs (*qLW3.1*, *qLW6.1*, *qLW7.1*, and *qLW8.1*) were found for leaf width (LW). Seven QTLs (*qLA1.1*, *qLA3.1*, *qLA3.2*, *qLA3.3*, *qLA6.1*, *qLA7.1*, and *qLA9.1*) were detected to affect leaf angle (LA). All three QTLs (*qMR3.1*, *qMR6.1*, and *qMR8.1*) detected for diameter of midribs (MR) are novel. The phenotypic variance explained by a single QTL varied from 2.34% to 30.87% within the population.

In the PQRIL population, three QTLs (*qLL4.1*, *qLL8.2*, and *qLL9.2*) were detected to impact leaf length. Two QTLs (*qLW7.2* and *qLW8.2*) were identified for leaf width. One QTL (*qLA1.2*) was found to influence leaf angle.

From individual environment analysis, a total of nine QTLs affecting leaf length (LL) were detected on chromosomes 1(3), 3(2), 4, 8(2), and 9 from the two RIL populations (seven from ISRIL and two from PQRIL). Three QTLs were detected on chromosome 1 in the ISRIL population. Two QTLs were identified on chromosomes 3 and 8 (one in each population). One QTL was detected on chromosomes 4 (PQRIL) and 9 (ISRIL). The phenotypic variance explained by these QTLs ranges from 4.32 -7.90 % (Table 2.11). The additive effect of the detected QTLs in either ISRIL or PQRIL population showed that the both the paternal line (*S. bicolor* BTx623) and the maternal line (*S. bicolor* IS3620C or *S. propinquum*), contributed alleles for particular traits. Seven of nine QTLs exhibited positive additive effects, suggesting the contribution of alleles from the paternal line (BTx623) to the increase in leaf length. A total of 16 QTLs affecting leaf angle (LA) were identified on chromosomes 1 (5), 3 (2), 6 (2), 7 (5), 8, and 9 (11 from ISRIL and five from PQRIL). The phenotypic variance explained by these QTLs ranges from 2.59 – 26.8 % (Table 2.12). Both populations showed 11 QTLs with negative additive effects, indicating contribution of alleles from the respective maternal lines for higher leaf angle. A total of 11 QTLs were detected on chromosomes 1, 2 (2), 3 (2), 6 (2), 7 (2), and 8 influencing leaf width (LW), explaining 3.42 - 12.42% of phenotypic variance (Table 2.13). Nine QTLs were from ISRIL, and two from PQRIL. In both populations, the trait was mostly influenced by the parental line (BTx623) due to the presence of 10 QTLs with positive additive effects. Seven QTLs were detected on chromosomes 1 (2), 3 (2), 6 (2), and 8 responsible for diameter of midribs (MR) in the ISRIL population, explaining 3.18 – 11.03% of phenotypic variance (Table 2.14). All of the identified QTLs exhibited positive additive effects, indicating that the favorable alleles for increased diameter of midribs were contributed from the parental line (BTx623).

QTLs for yield-related traits

A total of 17 QTLs were identified for the six yield-related traits based on the LS-means calculated across years in the ISRIL population (Table 2.10). Five QTLs (*qPH3.1, qPH6.1, qPH7.1, qPH8.1,* and *qPH9.1*) were detected to influence plant height (PH). We detected novel

QTLs for yield-related phenotypes, including three (*qDSW1.1, qDSW6.1*, and *qDSW7.1*) that were identified for dry stalk weight (DSW), two (*qDLW1.1* and *qDLW4.1*) that were detected for dry leaf weight (DLW), and two (*qDB1.1* and *qDB7.1*) that were found to affect dry biomass (DB). These three traits shared a QTL interval (~22.3-44.5 Mb) on chromosome 1. Five QTLs (*qLAR1.1, qLAR1.2, qLAR3.1, qLAR7.1*, and *qLAR8.1*) were identified for leaf area (LAR). The phenotypic variance explained by a single QTL varied from 2.69% to 21.47%. In the PQRIL population, one QTL (*qPH9.1*) was detected for plant height. Also, one QTL (*qDF4.1*) was identified to affect days to flowering in sorghum.

From individual environment analysis in the ISRIL population, a total of eight QTLs affecting PH were detected on chromosomes 2, 3, 6 (2), 7 (2), 8, and 10 (Table 2.15). The phenotypic variance explained by these QTLs ranges from 3.28% to 18.71%. Two QTLs for days of flowering were detected in PQRIL on chromosomes 9 and 10, explaining 7.91% to 11.6% of phenotypic variance (Table 2.16). Both QTLs showed positive additive effects, suggesting the contribution of alleles for days to flowering from the paternal line. A total of six QTLs affecting DSW were detected on chromosomes 1, 3, 6 (2), 7 (2), explaining 3.12% to 6.83% of phenotypic variance (Table 2.17). Half of the detected QTLs showed a positive sign of the additive effects in both PH and DSW, suggesting the contribution of alleles from the respective parental lines conferring increases in height and stalk weight of sorghum. The additive effect of the detected QTLs in ISRIL showed that both the paternal line (BTx623) and the maternal line (IS3620C), contributed alleles for increasing plant height and dry stalk weight. Three QTLs were detected on chromosomes 1, 3, and 8 contributing to DLW in ISRIL. The phenotypic variance explained by QTLs ranges from 3.47% to 6.43 % (Table 2.18). Two-thirds of the detected QTLs showed a positive sign of the additive effects. A total of five QTLs were detected on chromosomes 1, 3 (2), and 7 (2) responsible for DB in the ISRIL population. The phenotypic variance explained by QTLs ranges from 3.35% to 4.07% (Table 2.19). Two of five QTLs showed positive additive effects, suggesting that the favorable alleles for increased sorghum biomass were contributed from the paternal line (BTx623). However, for the remaining three QTLs, the maternal line (*S. propinquum*) contributed alleles for increased dry biomass. A total of eight QTLs on chromosome 1 (2), 3 (2), 7, and 8 (3) were detected for LAR. The phenotypic variance explained by QTLs ranges from 3.59% to 7.09% (Table 2.20). All of the detected QTLs exhibited positive additive effects, indicating that the favorable alleles for enhanced sorghum biomass were inherited from the paternal line.

Overlapping QTLs

While 101 QTLs were detected in the ISRIL population and 20 in the PQRIL population, including using overall LS means and single year data, overlaps among these QTLs have the consequence that they can be accounted for by as few as 16 (ISRIL) and 2 (PQRIL) distinct locations. QTL regions for multiple traits overlapped at some loci (Table 2.21), including chromosomes 1, 3, 7, and 9 from the ISRIL population. QTL regions for dry stalk weight, dry leaf weight, and dry biomass overlapped on chromosome 1; and for plant height, dry stalk weight, dry biomass, and leaf area on chromosome 7. In addition, QTL regions for leaf length, width and angle overlapped on chromosomes 3 and 9.

For individual environments in the ISRIL population, QTL regions for more than one trait overlapped at several loci (Table 2.22). Eight QTLs overlapped on chromosome 1 for multiple traits and years, including leaf area, dry leaf weight, dry biomass, leaf length and diameter of midribs. Ten QTLs were located in similar intervals on chromosome 3 for many traits and environments, including dry biomass, dry stalk weight, dry leaf weight, leaf area, leaf length, leaf width, leaf angle, and diameter of midribs. Three additional QTLs for leaf area and leaf length from the 2012 environment were detected on chromosome 3. Four QTLs shared similar intervals on chromosome 6 for plant height and dry stalk weight, with three additional QTLs influencing leaf width and diameter of midribs elsewhere on the chromosome. Three different groups of QTLs overlapped at different intervals on chromosome 7: one group included six QTLs for plant height, dry biomass, dry stalk weight, and leaf width; another comprised four QTLs influencing plant height, leaf area, and leaf angle; and finally, two QTLs for leaf area and leaf width. Four overlapping QTLs for dry leaf weight, leaf area, leaf length, and leaf width were detected on chromosome 8. Two overlapping QTLs were found to influence leaf length and leaf angle on chromosome 9. For each environment in the PQRIL population (Table 2.23), QTL regions were identified to affect leaf morphology traits. While four QTLs on chromosome leaf width shared similar intervals with leaf angle, two QTLs for leaf length and leaf angle were overlapped.

QTL correspondence with other studies

The Sorghum QTL Atlas (Mace *et al.* 2019) aided comparisons of QTL intervals across different studies based on their physical positions on the sorghum reference genome (Paterson et al., 2009). Nine of 25 QTLs for leaf length, width, and angle in sorghum detected in this study were found to correspond with other QTLs reported to influence these traits from published studies. Two QTLs for LL (*qLL4.1* and *qLL9.1*) were also found to overlap with those from prior studies (Feltus *et al.* 2006; McCormick *et al.* 2016). Four QTLs for LW (*qLW3.1, qLW6.1* and *qLW7.1*, *qLW8.1*) were identified in other studies sharing similar physical distances (Feltus et al., 2006; McCormick et al., 2016; Shehzad & Okuno, 2015; Zhi et al., 2022). Three QTLs for LA (*qLA1.1, qLA6.1,* and *qLA7.1*) overlapped with genetic regions on chromosome 1 (Truong *et al.*

2015), chromosome 6 (Hart *et al.* 2001), and chromosome 7 (Truong *et al.* 2015; McCormick *et al.* 2016; Zhao *et al.* 2016).

A total of 50 QTLs were found to influence yield-related phenotypes in the ISRIL population (more QTLs were detected for each trait within single environments). Some of those were found to correspond to genes reported in other studies such as PH and DF via QTL mapping and GWAS analysis that were extensively used in studying traits in sorghum (Lin *et al.* 1995; Hart *et al.* 2001; Kebede *et al.* 2001; Brown *et al.* 2006; Srinivas *et al.* 2009; Morris *et al.* 2013a; Kong *et al.* 2015). All QTLs detected for PH in the ISRIL population have been reported from previous studies (Hart *et al.* 2001; Feltus *et al.* 2006; Brown *et al.* 2006; McCormick *et al.* 2016; Kong *et al.* 2018b). Three QTLs (*qPH6.1*, *qPH7.1*, and *qPH9.1*) for PH also found in two GWAS studies (Morris *et al.* 2013a; Kong *et al.* 2015), are likely to correspond to three dwarfing genes in sorghum (*Dw2*, *Dw3*, and *Dw1*) (Quinby and Karper 1953). Among these genes, *Dw1* was mapped to chromosome 9 (~52.0-55.8 Mb), *Dw2* was on chromosome 6 (~ 42.2-45.5 Mb), and *Dw3* was on chromosome 7 (~58.4-59.5 Mb). A QTL (*qPH9.1*) for PH that was observed in the PQRIL population coincided with a QTL found in the ISRIL population based on physical distance, validating our QTL's accuracy (Li *et al.* 2015). One QTL controlling DF in the PQRIL showed some overlap with QTL from another study (E. S. Mace et al., 2013).

GWAS for the traits of interest

To further investigate the genetic basis of leaf morphology and yield-related traits in a sorghum diversity panel for which genotypic data was published and we measured phenotypes (See methods), we conducted GWAS using GLM with 1,000 permutations on PH, DF, DSW, DLW, LL, LW, LA, and MR (Figure 2.8-2.15). The GWAS identified significant SNP markers for yield-related traits with an association peak on chromosome 9 (~57 Mb) for PH; three for DF on chromosomes 3(~4 Mb), 4(~62 Mb), 7 (~58 Mb), and 10 (~54 Mb); but none for DSW and DLW. For leaf morphology, the GWAS identified five association peaks on chromosome 1 (~56 Mb), 3 (~4 Mb), and 10 (~18 Mb) for LL; four for LW on chromosomes 1 (~63 Mb), 2 (~66 Mb), 6 (57 Mb), and 8 (5 Mb); two for LA on chromosomes 1 (~7 Mb), and 7 (~65 Mb); and four for MR on 6 (~28 Mb), 7 (~15 Mb), and 8 (~17 Mb).

Correspondence between QTL and GWAS evidence narrows locations of causal genes for leaf morphology and yield-related traits

Among the sixteen (ISRIL) and two (PQRIL) distinct locations harboring QTLs, two gain support from significant GWAS evidence that helps to more precisely map causal loci. For PH QTLs suggested by flanking markers within a 3.8 Mb (52.0 -55.8 Mb) on chromosome 9, comparing to peak SNPs (S9_56472283 to S9_58933338) from GWAS. The SNPs on chromosome 9 spanned a distance of 2.46 Mb, nearby the QTL region and perhaps suggesting multiple QTLs. For leaf morphology traits, an LA QTL on chromosome 1 within a 9.1 Mb (2.9 - 12.0 Mb) included SNP marker from GWAS (S1_6543997) on chromosome 1, implicating a very small region in genetic control of the trait.

Discussion

In this study, significant phenotypic and genotypic variability for leaf morphology and yield-related characteristics were demonstrated by two RIL populations (ISRIL and PQRIL) generated from two separate species (IS3620C and *S. propinquum*) with a common parent (BTx623), collectively representing much of the genetic variability available within ‘eusorghums’. The use of evidence from both biparental QTL mapping and GWAS in the study allows identification of genetic locations controlling the traits of interest. While using a large population to build a QTL map would reduce the impact of sampling-based statistical artifacts, the small number of polymorphic loci between the parents will still have an effect on the precision of the map (Holland 2007). Unlike linkage mapping, which necessitates the construction of dedicated mapping populations, genome-wide association studies (GWAS) may take advantage of the high levels of recombination that occur naturally within populations. The population structure used in association analysis, however, can lead to increased LD between loci that are not really connected, because of genetic drift and natural selection. As a result, it is important to undertake both association mapping and linkage mapping to reduce the number of false positives caused by highly-associated loci with high LD and to establish accurate QTL interval mapping (He *et al.* 2017).

Leaf architecture has a significant effect on plant growth and yield, as well as its morphology and physiology. Correlations among the leaf morphology and yield-related traits from the study showed a large variety, ranging from strong to weak associations. For example, the highest correlation coefficients were observed in the ISRIL population between leaf area and leaf length ($r = 0.88$, $P < 0.001$) and between leaf area and leaf width ($r = 0.88$, $P < 0.001$). Additionally, strong correlations between the dry biomass and dry leaf weight exhibited ($r = 0.89$, $P < 0.001$) as well as the dry biomass and dry stalk weight ($r = 0.95$, $P < 0.001$). Consequently, not only was leaf area significantly and positively correlated with leaf length and leaf width, but dry biomass was also significantly and strongly correlated with dry leaf and stalk weight. A positive and significant correlation was observed between diameter of midribs and other traits such as leaf area, leaf length, and leaf width, suggesting that the size of midribs influences leaf morphological characteristics by increasing the yield and size of leaves.

QTL mapping showed 12 colocalized intervals for plant height, dry biomass, dry stalk weight, leaf area, leaf angle, and leaf width to locate near each other on chromosome 7 from 2011 and 2012 environments (Figure 2.16). Each trait in the QTL mapping exhibited very similar LOD score peaks ranging from 4.5 to 31.0 in both years, indicating that the genetic regions were not affected by the environments and consistent between the years. Leaf angle exhibited the highest LOD score, followed by plant height, dry stalk weight, leaf width, dry biomass, and leaf area respectively. Colocalized QTLs for yield-related traits and leaf morphology would be helpful for simultaneously improving yield as well as achieving optimal leaf angle and width. A significant stable QTL for PH (*qPHT6.1*) in our analysis with LOD score of 21.0, which accounted for 21.47 % of the phenotypic variance was supported by previously published data (Morris et al. 2013; Hilley et al. 2017). The QTL colocalized with the gene *Dw2* that was found as a protein kinase encoded by Sobic.006G067700 that shares similarities with LOC Os12g29580 (rice), GRMZM2G412524 (maize), GRMZM2G128319 (maize), and At3G52890 (maize) (Arabidopsis) (Hilley et al., 2017). Four loci (*Dw1*, *Dw2*, *Dw3*, and *Dw4*) have been identified in earlier investigations as affecting plant height (Quinby and Karper 1953). Brown et al. (2008) mapped *Dw1* on chromosome 9, which were found in our QTL mapping and GWAS analysis. *Dw1* was mapped to a region on chromosome 9 between 56.8-57.1 Mb. By using map-based cloning, the gene corresponding to *Dw1* was identified as Sobic.009G229800, which controls internode cell proliferation and encodes a putative membrane protein with no known function. GWAS identified the SNPs for PH on 9 (*Dw1*) were found to locate within the QTL intervals (*qPH9.1*). Combining GWAS and QTL provide functional markers for sorghum breeding.

In aspects of both plant morphology and genetic organization, sorghum is closely related to maize (Multani et al., 2003). Several traits from the study, including leaf width, plant height, and leaf angle, were compared to priori candidate genes identified in other studies. Zhi et al (2022) reported two significant haplotypes which were discovered in Sobic.008G070600 (homolog of *leafbladeless1* gene (*lbl*, GRMZM2G020187), co-located with the QTL from our study, *qLW8.1*, that was mapped similarly to their QTL, *qLW_dtf8.1*, at 52 Mb. The genetic region was explained with haplotype I consisting primarily of caudatum types and haplotype II consisting primarily of Asian durra types. In fact, Sobic.008G070600 is the ortholog of maize *leafbladeless1* (*lbl1*, GRMZM2G020187), described as a potential gene that specifies adaxial and abaxial organ polarity. The *lbl1* recessive mutations, which were mostly expressed in the shoot apical meristem, vasculature, and adaxially in leaf primordia, altered the abaxialization and breadth of the leaves (Nogueira et al. 2007; Zhi et al., 2022). Plant height in sorghum was previously linked to *Dw3*, which we found to be located in a specific genomic region. *Dw3*, which encodes a P-glycoprotein auxin transporter, has been described molecularly (Sb07g023730) (Li *et al.* 2015; Kong *et al.* 2018b). As reported by Hart et al. (2001), stem shortening, and leaf skewness were associated with *dwarf3* (*dw3*) alleles. Four independent dwarfing mutations (*dw1*, *dw2*, *dw3*, and *dw4*) are extensively used for reducing plant height in sorghum breeding programs (Multani et al., 2003). The ortholog of *dw3* in maize is *brachytic* (*br2*), which has been used by maize breeders. Similar dwarfing mechanisms have been identified to contain the recessive *brachytic2* (*br2*) mutation, which are distinguished by their compact lower stalk internodes. Due to the absence of a P-glycoprotein that regulates polar auxin transport in the maize stalk, the *br2* recessive mutant has a shorter plant height. Furthermore, the pleiotropic influence of *dw3* in sorghum on both plant height and leaf angle—an essential agronomic feature for crop architecture and yield production—was used to explain a positive association between both parameters. It was observed that auxin regulates the development and expansion of the preligule band in maize with *liguleless1, liguleless2,* and *liguleless* narrow null alleles (Moon et al., 2013). The regulator can also change the angle of leaf inclination in sorghum by up to 34 degrees because *dw3* encodes a P-glycoprotein that modifies polar auxin transport (Truong et al., 2015). Reduced polar auxin transport from the shoot apical meristem is most likely responsible for the effect of the null *dw3* gene on leaf angle (Multani *et al.* 2003; Knöller *et al.* 2010).

In summary, this study identified significant genetic variation for leaf morphology and yield-related traits in sorghum. Several of the QTLs identified in this study overlapped with previously identified QTL areas for the majority of trait studies, hence validating our findings. Additionally, we uncovered some putatively unique QTLs that cannot be verified by a comparative approach. We also identified colocalized QTLs controlling several traits, which may be the result of a single gene affecting multiple traits or the presence of multiple genes controlling distinct traits. These findings will further contribute to the resources and data utilized in the genetic enhancement of sorghum breeding programs.

References

Aljanabi, S. M., L. Forget, and A. Dookun, 1999 An Improved and Rapid Protocol for the Isolation of Polysaccharide- and Polyphenol-Free Sugarcane DNA. Plant Mol. Biol. Report. 1999 173 17: 281–281.

Alqudah, A. M., A. Sallam, P. Stephen Baenziger, and A. Börner, 2020 GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. J. Adv. Res. 22: 119.

Bouchet, S., M. O. Olatoye, S. R. Marla, R. Perumal, T. Tesso *et al.*, 2017 Increased Power To Dissect Adaptive Traits in Global Sorghum Diversity Using a Nested Association Mapping Population. Genetics 206: 573–585.

Bouman, B. A. M., 2007 A conceptual framework for the improvement of crop water productivity at different spatial scales. Agric. Syst. 93: 43–60.

Bourke, P. M., R. E. Voorrips, C. A. Hackett, G. van Geest, J. H. Willemsen *et al.*, 2021 Detecting quantitative trait loci and exploring chromosomal pairing in autopolyploids using polyqtlR. Bioinformatics 37: 3822–3829.

Boyles, R. E., E. A. Cooper, M. T. Myers, Z. Brenton, B. L. Rauh *et al.*, 2016 Genome-Wide Association Studies of Grain Yield Components in Diverse Sorghum Germplasm. Plant Genome 9: plantgenome2015.09.0091.

Brachi, B., G. P. Morris, and J. O. Borevitz, 2011 Genome-wide association studies in plants: The missing heritability is in the field. Genome Biol. 12:.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss *et al.*, 2007 TASSEL: software for association mapping of complex traits in diverse samples. Bioinforma. Appl. 23: 2633–2635.

Brenton, Z. W., E. A. Cooper, M. T. Myers, R. E. Boyles, N. Shakoor *et al.*, 2016 A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. Genetics 204: 21–33.

Broman, K. W., H. Wu, Ś. Sen, and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889–890.

Brown, P. J., P. E. Klein, E. Bortiri, C. B. Acharya, W. L. Rooney *et al.*, 2006 Inheritance of inflorescence architecture in sorghum. Theor. Appl. Genet. 113: 931–942.

Brown, P. J., W. L. Rooney, C. Franks, and S. Kresovich, 2008 Efficient Mapping of Plant Height Quantitative Trait Loci in a Sorghum Association Population With Introgressed Dwarfing Genes. Genetics 180: 629.

Casa, A. M., G. Pressoir, P. J. Brown, S. E. Mitchell, W. L. Rooney *et al.*, 2008 Community resources and strategies for association mapping in Sorghum. Crop Sci. 48: 30–40.

Childs, K. L., F. R. Miller, M.-M. Cordonnier-Pratt, L. H. Pratt, P. W. Morgan *et al.*, 1997 The Sorghum Photoperiod Sensitivity Gene, Ma, Encodes a Phytochrome B'. Plant Physiol 113: 61–62.

Drewry, D. T., P. Kumar, and S. P. Long, 2014 Simultaneous improvement in productivity, water use, and albedo through crop structural modification. Glob. Chang. Biol. 20: 1955–1967.

Feltus, F. A., G. E. Hart, K. F. Schertz, A. M. Casa, S. Kresovich *et al.*, 2006 Alignment of genetic maps and QTLs between inter- and intra-specific sorghum populations. Theor. Appl. Genet. 112: 1295–1305.

Fu, Y., G. Xu, H. Chen, X. Wang, Q. Chen *et al.*, 2019 QTL mapping for leaf morphology traits in a large maize-teosinte population. Mol. Breed. 39: 1–13.

Hart, G. E., K. F. Schertz, Y. Peng, and N. H. Syed, 2001 Genetic mapping of Sorghum bicolor (L.) Moench QTLs that control variation in tillering and other morphological characters. Theor. Appl. Genet. 103: 1232–1242.

He, Y., D. Wu, D. Wei, Y. Fu, Y. Cui *et al.*, 2017 GWAS, QTL mapping and gene expression analyses in Brassica napus reveal genetic control of branching morphogenesis. Sci. Rep. 7: 1–9.

Hilley, J. L., B. D. Weers, S. K. Truong, R. F. McCormick, A. J. Mattison *et al.*, 2017 Sorghum Dw2 Encodes a Protein Kinase Regulator of Stem Internode Length. Sci. Reports 2017 71 7: 1–13.

Holland, J. B., 2007 Genetic architecture of complex traits in plants. Curr. Opin. Plant Biol. 10: 156–161.

Hu, Z., M. O. Olatoye, S. Marla, and G. P. Morris, 2019 An Integrated Genotyping-by-Sequencing Polymorphism Map for Over 10,000 Sorghum Genotypes. Plant Genome 12: 180044.

Kebede, H., P. K. Subudhi, D. T. Rosenow, and H. T. Nguyen, 2001 Quantitative trait loci influencing drought tolerance in grain sorghum (Sorghum bicolor L. Moench). Theor. Appl. Genet. 2001 1032 103: 266–276.

Kenchanmane Raju, S. K., M. Adkins, A. Enersen, D. Santana de Carvalho, A. J. Studer *et al.*, 2020 Leaf Angle eXtractor: A high-throughput image processing framework for leaf angle measurements in maize and sorghum. Appl. Plant Sci. 8: 1–9.

Knöller, A. S., J. J. Blakeslee, E. L. Richards, W. A. Peer, and A. S. Murphy, 2010 Brachytic2/ZmABCB1 functions in IAA export from intercalary meristems. J. Exp. Bot. 61: 3689–3696.

Kong, W., H. Jin, C. D. Franks, C. Kim, R. Bandopadhyay *et al.*, 2013 Genetic Analysis of Recombinant Inbred Lines for Sorghum bicolor × Sorghum propinquum. G3 Genes|Genomes|Genetics 3: 101–108.

Kong, W., C. Kim, V. H. Goff, D. Zhang, and A. H. Paterson, 2015 Genetic analysis of rhizomatousness and its relationship with vegetative branching of recombinant inbred lines of Sorghum bicolor × S. propinquum . Am. J. Bot. 102: 718–724.

Kong, W. Q., C. Kim, D. Zhang, H. Guo, X. Tan *et al.*, 2018a Genotyping by sequencing of 393 sorghum bicolor BTx623 × IS3620C recombinant inbred lines improves sensitivity and resolution of QTL detection. G3 Genes, Genomes, Genet. 8: 2563–2572.

Kong, W. Q., C. Kim, D. Zhang, H. Guo, X. Tan *et al.*, 2018b Genotyping by sequencing of 393 sorghum bicolor BTx623 × IS3620C recombinant inbred lines improves sensitivity and resolution of QTL detection. G3 Genes, Genomes, Genet. 8: 2563–2572.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181.

Lasky, J. R., H. D. Upadhyaya, P. Ramu, S. Deshpande, C. T. Hash *et al.*, 2015 Genome-environment associations in sorghum landraces predict adaptive traits. Sci. Adv. 1:.

Li, X., X. Li, E. Fridman, T. T. Tesso, J. Yu *et al.*, 2015 Dissecting repulsion linkage in the dwarfing gene Dw3 region for sorghum plant height provides insights into heterosis. Proc. Natl. Acad. Sci. U. S. A. 112: 11823–11828.

Lin, Y. R., K. F. Schertz, and A. H. Paterson, 1995 Comparative Analysis of Qtls Affecting Plant Height and Maturity across the Poaceae, in Reference to an Interspecific Sorghum Population. Genetics 141: 391.

Mace, E., D. Innes, C. Hunt, X. Wang, Y. Tao *et al.*, 2019 The Sorghum QTL Atlas: a powerful tool for trait dissection, comparative genomics and crop improvement. Theor. Appl. Genet. 132: 751–766.

Mantilla-Perez, M. B., and M. G. Salas Fernandez, 2017 Differential manipulation of leaf angle throughout the canopy: Current status and prospects. J. Exp. Bot. 68: 5699–5717.

Mathur, S., A. V. Umakanth, V. A. Tonapi, R. Sharma, and M. K. Sharma, 2017 Sweet sorghum as biofuel feedstock: Recent advances and available resources. Biotechnol. Biofuels 10: 1–19.

McCormick, R. F., S. K. Truong, and J. E. Mullet, 2016 3D sorghum reconstructions from depth images identify QTL regulating shoot architecture. Plant Physiol. 172: 823–834.

McCouch, S., 1997 Report on QTL nomenclature. Rice Genet. News1. 14: 11-13.

Moon, J., H. Candela, and S. Hake, 2013 The Liguleless narrow mutation affects proximal-distal signaling and leaf growth. Development 140: 405–412.

Morris, G. P., P. Ramu, S. P. Deshpande, C. T. Hash, T. Shah *et al.*, 2013a Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc. Natl. Acad. Sci. U. S. A. 110: 453–458.

Morris, G. P., D. H. Rhodes, Z. Brenton, P. Ramu, V. M. Thayil *et al.*, 2013b Dissecting genome-wide association signals for loss-of-function phenotypes in sorghum flavonoid pigmentation traits. G3 Genes, Genomes, Genet. 3: 2085–2094.

Multani, D. S., S. P. Briggs, M. A. Chamberlin, J. J. Blakeslee, A. S. Murphy *et al.*, 2003 Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. Science 302: 81–84.

Nogueira, F. T. S., S. Madi, D. H. Chitwood, M. T. Juarez, and M. C. P. Timmermans, 2007

Two small regulatory RNAs establish opposing fates of a developmental axis. Genes Dev. 21: 750.

Paterson, A. H., J. E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood *et al.*, 2009 The Sorghum bicolor genome and the diversification of grasses. Nature 457: 551–556.

Paterson, A. H., K. F. Schertz, Y. R. Lin, S. C. Liu, and Y. L. Chang, 1995 The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, Sorghum halepense (L.) Pers. Proc. Natl. Acad. Sci. U. S. A. 92: 6127.

Platt, A., B. J. Vilhjálmsson, and M. Nordborg, 2010 Conditions under which genome-wide association studies will be positively misleading. Genetics 186: 1045–1052.

Quinby, J., and R. E. Karper, 1953 Inheritance of height in sorghum. undefined.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rooney, W. L., J. Blumenthal, B. Bean, and J. E. Mullet, 2007 Designing sorghum as a dedicated bioenergy feedstock. Biofuels, Bioprod. Biorefining 1: 147–157.

Schmitt, J., J. R. Stinchcombe, M. S. Heschel, and H. Huber, 2003 The Adaptive Evolution of Plasticity: Phytochrome-Mediated Shade Avoidance Responses 1. 43: 459–469.

Shehzad, T., and K. Okuno, 2015 QTL mapping for yield and yield-contributing traits in sorghum (Sorghum bicolor (L.) Moench) with genome-based SSR markers. Euphytica 203: 17–31.

Srinivas, G., K. Satish, R. Madhusudhana, R. Nagaraja Reddy, S. Murali Mohan *et al.*, 2009 Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellite markers in sorghum. Theor. Appl. Genet. 118: 1439–1454.

Truong, S. K., R. F. McCormick, W. L. Rooney, and J. E. Mullet, 2015 Harnessing genetic variation in leaf angle to increase productivity of sorghum bicolor. Genetics 201: 1229–1238.

Tsukaya, H., 2004 Leaf shape: genetic controls and environmental factors. Int. J. Dev. Biol. 49: 547–555.

Turner, S. D., 2018 qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J. Open Source Softw. 3: 731.

Wang, B., Y. Zhu, J. Zhu, Z. Liu, H. Liu *et al.*, 2018 Identification and fine-mapping of a major maize leaf width QTL in a Re-sequenced large recombinant inbred lines population. Front. Plant Sci. 9: 1–11.

Wang S., C. J. Basten, and Z.-B. Zeng (2012). Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm)

Warnasooriya, S. N., and T. P. Brutnell, 2014 Enhancing the productivity of grasses under high-density planting by engineering light responses: from model systems to feedstocks. J. Exp. Bot. 65: 2825–2834.

Yu, X., X. Li, T. Guo, C. Zhu, Y. Wu *et al.*, 2016 Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nat. Plants 2016 210 2: 1–7.

Zhao, J., M. B. Mantilla Perez, J. Hu, and M. G. Salas Fernandez, 2016 Genome-Wide Association Study for Nine Plant Architecture Traits in Sorghum. Plant Genome 9:.

Zhi, X., G. Hammer, A. Borrell, Y. Tao, A. Wu *et al.*, 2022 Genetic basis of sorghum leaf width and its potential as a surrogate for transpiration efficiency. 135: 3057–3071.

Table 2.1. Descriptive statistics for leaf morphology related traits in three sorghum populations.

a) *S. bicolor* BTx623 X *S. bicolor* IS3620C

Trait	Unit	2011				2012			
		Mean	SD	Min	Max	Mean	SD	Min	Max
LL	cm	55.91	9.15	28.1	83.8	54.05	11	14.9	82.85
LW	cm	6.45	1.03	2.55	9.55	6.04	1.2	2.15	9.2
LA	degree	55.55	17.81	16	108	55.97	17	17.5	122.5
MR	cm	4.39	1.16	1	7.45	4.36	1	1.8	7

b) *S. bicolor* BTx623 X *S. propinquum*

Trait	Unit	2009				2010				2011			
		Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
LL	cm	15.95	3.88	8	27	51.74	12	23.5	77.5	55.5	11	33	86
LW	cm	3.58	1.03	1.4	6.7	5.09	1.7	2.25	10.7	5.39	1.3	2.4	11
LA	degree	38	12.7	13	90	45.71	15	12.5	87.5	48.4	21	5.5	130

c) The Sorghum Association Panel (Casa et al., 2008)

Trait	Unit	2009				2010			
		Mean	SD	Min	Max	Mean	SD	Min	Max
LL	cm	54.49	10.5	16	81	59.67	11	31.9	110
LW	cm	6.30	1.26	3.1	10.8	7.38	3.1	2.4	46.1
LA	degree	51.53	11.4	10	78	50.9	15	5	86.5
MR	cm	2.12	0.54	0.8	4	5.7	2	1.5	24

LL: leaf length; LW: leaf width, LA; leaf angle; MR: diameter of midribs

Table 2.2. Descriptive statistics for yield related traits in three sorghum populations.

a) *S. bicolor* BTx623 X *S. bicolor* IS3620C

Trait	Unit	2011				2012			
		Mean	SD	Min	Max	Mean	SD	Min	Max
PH	cm	97.12	21	45	175	99.95	24	54.2	181
DLW	g	49.32	29.4	6	184	52.56	33	4.5	202
DSW	g	58.54	44.7	5.5	371	60.53	50	7	476
DB	g	107.86	71.5	12	533	113.1	79	13	678
LAR	cm^2	366.43	97.9	97	704	336.2	113	32.5	693

b) *S. bicolor* BTx623 X *S. bicolor* IS3620C

Trait	Unit	2009				2010				2011			
		Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
PH	cm	83.96	33.8	29	198	88.81	29	49.2	195	99.8	22	48	201
DF	day	71.81	16.6	39	125	64.68	9.7	46.2	97	64.1	9.8	40	87

c) The Sorghum Association Panel (Casa et al., 2008)

Trait	Unit	2009				2010			
		Mean	SD	Min	Max	Mean	SD	Min	Max
PH	cm	103.4	39	49	302	108.8	54	50.1	370
DF	day	67.4	12.3	43	122	59.95	11	42.4	131
DSW	g	38.19	52.9	3.9	502	61.01	64	6.5	548
DLW	g	29.02	15.2	5.1	93	41.71	27	6	239

PH: plant height; DLW: dry leaf weight; DSW: dry stalk weight; DB: dry biomass; LAR: leaf area; DF: day to flower (flowering days)

Table 2.3. Effect of genotype and environment in the *S. bicolor* BTx623 x *S. bicolor* IS3620C (ISRIL) population.

Trait	Effects	df	F-value	Significance
LL	Genotype (G)	392	2.71	***
	Year (Y)	1	12.75	***
	G x Y	392	2.71	***
LW	Genotype (G)	392	2.87	***
	Year (Y)	1	48.72	***
	G x Y	392	2.86	***
LA	Genotype (G)	391	4.62	***
	Year (Y)	1	0.03	NS
	G x Y	391	4.64	***
MR	Genotype (G)	392	2.42	***
	Year (Y)	1	0.2	NS
	G x Y	392	2.43	***
PH	Genotype (G)	392	4.49	***
	Year (Y)	1	6.35	*
	G x Y	392	4.49	***
DLW	Genotype (G)	392	2.06	***
	Year (Y)	1	4.53	*
	G x Y	392	2.06	***
DSW	Genotype (G)	392	2.03	***
	Year (Y)	1	0.81	NS
	G x Y	392	2.03	***
DB	Genotype (G)	392	2.00	***
	Year (Y)	1	2.05	NS
	G x Y	392	2.00	***
LAR	Genotype (G)	392	2.69	***
	Year (Y)	1	27.64	***
	G x Y	392	2.68	***

LL, leaf length; LW, leaf width; LA, leaf angle; MR, diameter of midribs; PH, plant height; DLW, dry leave weight; DSW, dry stalk weight; DB, dry biomass; LAR, leaf area
*p<0.05, **p<0.01 ***p<0.001, NS not significant

Table 2.4. Effect of genotype and environment in the *S. bicolor* BTx623 x *S. propinquum* (PQRIL) population.

Trait	Effects	df	F-value	Significance
LL	Genotype (G)	160	3.08	***
	Year (Y)	2	996.75	***
	G x Y	160	2.08	***
LW	Genotype (G)	160	2.86	***
	Year (Y)	2	113.42	***
	G x Y	160	2.69	***
LA	Genotype (G)	160	2.59	***
	Year (Y)	2	24.25	***
	G x Y	160	2.08	***
PH	Genotype (G)	160	1.49	**
	Year (Y)	2	12.98	***
	G x Y	160	1.48	**
DF	Genotype (G)	159	2.7	***
	Year (Y)	2	20.53	***
	G x Y	159	2.62	***

LA, leaf angle; LAR, leaf area; LL, leaf length; LW, leaf width; PH, plant height; DF, day to flowering

*p<0.05, **p<0.01 ***p<0.001

Table 2.5. Effect of genotype and environment in the sorghum diversity panel (SAP).

Trait	Effects	df	F-value	Significance
LL	Genotype (G)	406	3.36	***
	Year (Y)	1	87.99	***
	G x Y	406	3.31	***
LW	Genotype (G)	406	1.53	***
	Year (Y)	1	48.29	***
	G x Y	406	1.48	***
LA	Genotype (G)	406	1.63	***
	Year (Y)	1	0.98	NS
	G x Y	406	1.15	NS
MR	Genotype (G)	406	1.36	**
	Year (Y)	1	1185.81	***
	G x Y	406	1.15	NS
PH	Genotype (G)	406	3.69	***
	Year (Y)	1	5.4	*
	G x Y	406	3.69	***
DF	Genotype (G)	407	4.33	***
	Year (Y)	1	214.79	***
	G x Y	407	4.30	***
DLW	Genotype (G)	405	2.01	***
	Year (Y)	1	106.99	***
	G x Y	405	1.99	***
DSW	Genotype (G)	405	2.26	***
	Year (Y)	1	52.98	***
	G x Y	405	2.26	***

LL, leaf length; LW, leaf width; LA, leaf angle; MR, diameter of midribs; PH, plant height; DF, day to flowering; DLW, dry leave weight; DSW, dry stalk weight
*$p<0.05$, **$p<0.01$ ***$p<0.001$, NS not significant

Table 2.6. Correlation coefficients among leaf morphology and yield related traits in the ISRIL population calculated across years.

	PH	DB	DSW	DLW	LAR	LL	LW	LA	MR
PH	1								
DB	0.4542***	1							
DSW	0.5738***	0.9455***	1						
DLW	0.2702***	0.8891***	0.7662***	1					
LAR	0.1177*	0.3819***	0.2946***	0.4600***	1				
LL	0.2963***	0.4897***	0.4182***	0.5226***	0.8777***	1			
LW	-0.0674	0.1985***	0.1161**	0.3046***	0.8778***	0.5694***	1		
LA	0.2962***	-0.0441	0.0889	-0.1920***	-0.4286***	-0.3691***	-0.3970***	1	
MR	0.0915	0.3676***	0.2863***	0.4017***	0.6804***	0.5814***	0.6290***	-0.3010***	1

PH, plant height; DB, dry biomass; DSW, dry stalk weight; DLW, dry leave weight; LAR, leaf area; LL, leaf length; LW, leaf width; LA, leaf angle; MR, diameter of midribs
*p<0.05, **p<0.01, ***p<0.001

Table 2.7. Correlation coefficients among leaf morphology and yield related traits in the PQRIL population calculated across years.

	PH	DF	LL	LW	LA
PH	1				
DF	0.4714***	1			
LL	0.2622***	0.3167***	1		
LW	0.2103**	0.1640*	0.4552***	1	
LA	0.0505	-0.2909***	-0.3496**	-0.2670***	1

PH, plant height; DF, day to flowering; LL, leaf length; LW, leaf width; LA, leaf angle
*p<0.05, **p<0.01, ***p<0.001

Table 2.8. Correlation coefficients among leaf morphology and yield related traits in the diversity panel calculated across years.

	PH	DF	DSW	DLW	LL	LW	LA	MR
PH	1							
DF	-0.0114	1						
DSW	0.3503***	0.4148***	1					
DLW	0.0108	0.5270***	0.6414***	1				
LL	0.1444**	0.6380***	0.4377***	0.4972***	1			
LW	-0.1283*	0.5064***	0.2369***	0.4019***	0.5678***	1		
LA	-0.1376**	0.2206***	0.0318	0.1435**	0.1756***	0.1292*	1	
MR	-0.0862	0.2900***	0.1399**	0.2767***	0.4289***	0.6179***	0.1024*	1

PH, plant height; DF, day to flowering; DSW, dry stalk weight; DLW, dry leave weight; LL, leaf length; LW, leaf width; LA, leaf angle; MR, diameter of midribs
*p<0.05, **p<0.01 ***p<0.001

Table 2.9. QTLs affecting leaf morphology traits using overall LS means in two sorghum populations.

Population	Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	Physical position (Mb)[b]	LOD	Additive	R^2 (%)[c]
	LL	qLL1.1	1	18.51	17.45-21.55	5.6-7.0	8.8	2.6466	8.15
	LL	qLL1.2	1	72.01	69.6-74.75	24.5-45.9	5.4	-2.3824	5.43
	LL	qLL3.1	3	110.41	107.45-115.2	60.6-62.5	6.7	2.1011	6.28
	LL	qLL8.1	8	29.31	17.05-38.8	2.7-6.1	4.3	1.8781	4.97
	LL	qLL9.1	9	121.01	111.6-122	56.1-59.5	3.8	-1.6691	3.93
	LW	qLW3.1	3	110.31	106.15-112	60.6-61.7	6.5	0.2321	6.02
	LW	qLW6.1	6	63.41	56.9-73.5	51.7-55.5	4.2	0.1853	3.77
	LW	qLW7.1	7	91.01	84.85-96.65	58.4-59.5	7.7	0.2667	8.02
	LW	qLW8.1	8	84.31	81.05-88.05	51.8-53.1	5.9	0.2211	5.40
ISRIL	LA	qLA1.1	1	75.41	70.95-81.45	24.5-49.8	3.3	2.8091	2.37
	LA	qLA3.1	3	113.81	105.4-119	60.2-64.2	3.5	-2.5576	2.67
	LA	qLA3.2	3	134.71	128.95-138.9	66.6-69.6	3.1	2.6476	2.34
	LA	qLA3.3	3	157.71	150.7-158.7	71.5-73.2	3.0	2.5102	2.54
	LA	qLA6.1	6	26.61	25.1-31.3	45.5-46.1	9.0	4.2293	7.25

	LA	qLA7.1	7	91.01	89.05-93.95	58.4-59.5	31.0	-8.6338	30.87
	LA	qLA9.1	9	120.01	112.8-121	56.1-59.5	5.2	3.5292	4.99
	MR	qMR3.1	3	115.81	114.55-117.85	61.9-63.3	12.9	0.3374	13.31
	MR	qMR6.1	6	58.91	54.65-65.85	51.1-53.7	5.2	0.2106	5.14
	MR	qMR8.1	8	22.31	6.4-33.65	2.4-6.1	3.3	0.2016	4.70
	LL	qLL4.1	4	71.01	62.8-77	58.0-64.9	5.0	2.9816	15.4176
	LL	qLL8.2	8	66.81	58.6-76.65	51.5-51.9	3.5	2.0608	8.4715
PQRIL	LL	qLL9.2	9	43.31	36.3-50.3	50.2-59.2	4.2	-2.2135	9.2984
	LW	qLW7.2	7	15.91	1.9-23.9	0.9-8.4	3.5	0.3182	10.6159
	LW	qLW8.2	8	55.91	50.95-61.7	5.3-49.1	4.3	0.306	9.9203
	LA	qLA1.2	1	22.01	17.35-30.45	2.9-12.0	3.5	3.7971	7.9755
	LA	qLA7.2	7	45.61	36.6-46	51.1-58.4	10.1	-5.662	22.0464

LL, leaf length; LW, leaf width; LA, leaf angle; MR, diameter of midribs
[a]1.5-LOD support interval of the QTL.
[b]Based on flanking DNA marker location in the published genome sequence (Paterson et al., 2009).
[c]Percentage of the variation explained by the QTL.

Table 2.10. QTLs affecting yield-related traits using overall LS means in two sorghum populations.

Population	Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	Physical position (Mb)	LOD	Additive	R^2 (%)[b]
	PH	qPH3.1	3	19.61	15.8-24.5	3.2-4.9	3.5	3.4796	3.14
	PH	qPH6.1	6	22.51	20.55-24.15	42.4-45.5	21.0	9.2562	21.47
	PH	qPH7.1	7	90.01	86.7-94.25	58.4-59.5	11.4	-6.6891	11.76
	PH	qPH8.1	8	69.81	60.8-75.45	47.8-51.8	3.9	-3.5909	3.30
	PH	qPH9.1	9	92.41	86.5-100.8	52.0-55.8	3.2	-3.2199	2.69
	DSW	qDSW1.1	1	68.81	66.95-73.5	22.3-44.5	4.4	-7.0527	4.38
	DSW	qDSW6.1	6	17.91	11.4-20.5	39.6-44.7	6.3	6.9597	6.84
ISRIL	DSW	qDSW7.1	7	89.01	86.25-92.9	58.4-59.5	8.4	-7.7224	8.72
	DB	qDB1.1	1	68.81	66.45-73.1	22.3-44.5	6.0	-14.7133	6.71
	DB	qDB7.1	7	89.01	83.75-94.65	57.7-59.5	4.0	-9.5298	4.65
	LAR	qLAR1.1	1	25.71	23.3-27.4	7.1-8.1	4.8	24.5871	4.49
	LAR	qLAR1.2	1	92.51	86.25-94.35	50.3-52.8	4.3	-18.8778	4.00
	LAR	qLAR3.1	3	114.81	107.9-118.25	60.6-63.3	6.5	23.2727	6.59
	LAR	qLAR7.1	7	95.01	86.7-102.6	58.4-60.7	4.5	20.7852	5.33
	LAR	qLAR8.1	8	84.31	79.3-89.15	51.8-53.1	5.5	21.421	5.54

	DLW	qDLW1.1	1	68.81	66.4-73.2	22.3-44.5	6.2	-6.6351	6.95
	DLW	qDLW4.1	4	157.91	156.1-163.1	66.9-67.6	3.0	-3.6689	3.10
PQRIL	PH	qPH9.1	9	33.91	31.45-39.9	50.2-54.5	2.9	-5.1294	8.89
	DF	qPH4.1	4	52.01	44.5-58.75	12.5-58.8	3.5	-3.4873	11.52

PH, plant height; DSW, dry stalk weight; DB, dry biomass; LAR, leaf area; DLW, dry leave weight; DF, day to flowering
[a]1.5-LOD support interval of the QTL.
[b]Based on flanking DNA marker location in the published genome sequence (Paterson et al., 2009).
[c]Percentage of the variation explained by the QTL.

Table 2.11. QTLs affecting leaf length in ISRIL and PQRIL populations using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
LL_2011	qLL2011-1	1	18.51	17.00-20.95	6.2	2.5	5.98	ISRIL
	qLL2011-3	3	113.81	109.15-117.30	4.9	2	4.75	ISRIL
LL_2012	qLL2012-1a	1	20.81	19.45-25.00	8.2	3.5	7.9	ISRIL
	qLL2012-1b	1	68.81	66.25-70.00	5.5	-3.2	5.45	ISRIL
	qLL2012-3	3	110.41	107.45-113.40	5.2	2.4	5.01	ISRIL
	qLL2012-8	8	82.61	74.70-87.35	4.4	2.3	4.32	ISRIL
	qLL2012-9	9	120.01	115.50-121.00	4.4	-2.5	5.14	ISRIL
LL_2010	qLL2010-4	4	72.01	62.75-86.55	2.8	4.4	10.46	PQRIL
	qLL2010-8	8	67.81	63.30-73.90	3.3	3.8	8.98	PQRIL

LL, leaf length
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.12. QTLs affecting leaf angle in ISRIL and PQRIL populations using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
LA_2011	qLA2011-1	1	170.11	165.15-176.35	6.1	-4.2	5.35	ISRIL
	qLA2011-3a	3	114.81	111.20-119.60	3.9	-3.4	3.57	ISRIL
	qLA2011-3b	3	157.71	140.90-158.70	3.2	3.2	3.17	ISRIL
	qLA2011-6	6	28.41	19.00-38.80	3.5	3.2	3.16	ISRIL
	qLA2011-7	7	92.01	88.70-95.05	23.4	-9.3	26.8	ISRIL
	qLA2011-9	9	121.01	114.15-122.00	7.1	5	7.68	ISRIL
LA_2012	qLA2012-1a	1	123.11	113.70-130.80	3.7	-3.3	3.39	ISRIL
	qLA2012-1b	1	167.51	158.10-176.75	3.2	-3	2.83	ISRIL
	qLA2012-1c	1	75.21	70.75-82.05	3	3.3	2.59	ISRIL
	qLA2012-6	6	26.61	24.60-31.2	7.7	4.7	7.07	ISRIL
	qLA2012-7	7	91.01	89.00-93.95	22	-8.8	25.21	ISRIL
LA_2009	qLA2009-7	7	43.41	36.80-46.00	8.3	-5.6	19.25	PQRIL
LA_2010	qLA2010-1	1	69.91	64.40-76.30	3.2	-5.3	9.17	PQRIL
	qLA2010-7	7	37.71	32.55-41.40	4.7	-6.1	15.52	PQRIL

LA-2011	qLA2011-7	7	35.71	32.50-40.90	6.4	-9.3	19.97	PQRIL
	qLA2011-8	8	71.11	68.10-76.70	3.6	-6.5	9.29	PQRIL

LA, leaf angle
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.13. QTLs affecting leaf width in ISRIL and PQRIL populations using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
LW_2011	qLW2011-1	1	185.11	179.90-188.10	3.3	0.2	3.43	ISRIL
	qLW2011-2	2	61.91	57.20-65.80	4.6	-0.2	4.96	ISRIL
	qLW2011-3	3	92.01	82.50-98.40	4.2	0.2	4.76	ISRIL
	qLW2011-6	6	63.41	62.00-71.30	3.2	0.2	3.25	ISRIL
	qLW2011-7	7	89.01	86.35-95.25	5.9	0.3	6.46	ISRIL
LW_2012	qLW2012-3	3	115.81	112.00-121.45	5.8	0.3	5.92	ISRIL
	qLW2012-6	6	70.01	65.50-74.65	4.7	0.3	4.92	ISRIL
	qLW2012-7	7	95.01	91.85-100.60	6.8	0.3	8.13	ISRIL
	qLW2012-8	8	84.31	80.70-86.45	9.2	0.4	9.85	ISRIL
LW_2009	qLW2009-2	2	13.71	6.75-20.75	4.6	0.4	12.42	PQRIL
LW_2011	qLW2011-7	7	45.61	33.60-46.00	3.8	0.4	10.44	PQRIL

LW, leaf width
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.14. QTLs affecting diameter of midribs in the ISRIL population using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
MR_2011	qMR2011-1	1	32.81	26.20-34.25	3.2	0.4	3.18	ISRIL
	qMR2011-3	3	115.81	109.70-119.30	6.9	0.3	6.92	ISRIL
	qMR2011-6	6	57.01	54.10-61.80	5	0.3	4.97	ISRIL
MR_2012	qMR2012-1	1	20.51	16.80-24.05	3.8	0.2	3.72	ISRIL
	qMR2012-3	3	116.61	115.40-118.25	11.1	0.3	11.03	ISRIL
	qMR2012-6	6	70.01	67.30-75.70	4	0.2	4.04	ISRIL
	qMR2012-8	8	22.31	13.30-35.60	3.9	0.2	5.54	ISRIL

MR, diameter of midribs
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.15. QTLs affecting plant height in ISRIL and PQRIL populations using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
PH_2011	qPH2011-6	6	21.51	20.50-23.85	19.5	9.3	18.71	ISRIL
	qPH2011-7	7	89.01	85.95-92.75	10.2	-7.3	11.85	ISRIL
	qPH2011-8	8	65.41	58.05-69.80	4.7	-5.5	6.3	ISRIL
PH_2012	qPH2012-2	2	117.11	108.40-125.70	3.2	-4.8	3.82	ISRIL
	qPH2012-3	3	19.61	15.70-24.35	4.4	5.2	4.38	ISRIL
	qPH2012-6	6	22.51	17.40-25.20	12.3	9.2	13.14	ISRIL
	qPH2012-7	7	91.01	86.75-96.35	7.5	-7.3	8.77	ISRIL
	qPH2012-10	10	65.01	63.50-70.95	3.6	4.5	3.28	ISRIL

PH, plant height
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.16. QTLs affecting the number of flowering days in the PQRIL population using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
DF_2009	qDF2009-9	9	35.91	28.40-45.10	3.1	-4.9	7.91	PQRIL
	qDF2009-10	10	89.71	84.70-93.95	3.9	-7.3	11.6	PQRIL

DF, day to flower (flowering days)
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.17. QTLs affecting dry stalk weight in the ISRIL population using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
DSW_2011	qDSW2011-6	6	22.51	9.85-26.20	4	9.7	4.38	ISRIL
	qDSW2011-7	7	89.01	83.30-94.15	6.4	-11.8	6.87	ISRIL
DSW_2012	qDSW2012-1	1	96.51	90.20-103.75	3.1	-9.5	3.12	ISRIL
	qDSW2012-3	3	117.61	114.45-121.90	3.5	9.9	3.92	ISRIL
	qDSW2012-6	6	15.91	5.70-25.75	4.4	11.3	4.99	ISRIL
	qDSW2012-7	7	89.01	84.6-93.45	6.5	-13.1	6.83	ISRIL

DSW, dry stalk weight
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.18. QTLs affecting dry leaf weight in the ISRIL and PQRIL populations using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
DLW_2011	qDLW2011-1	1	69.81	66.95-72.35	6	-9.7	6.43	ISRIL
	qDLW2011-8	8	96.71	89.10-99.70	3.2	5.5	3.47	ISRIL
DLW_2012	qDLW2012-3	3	117.61	113.80-122.55	3.6	6.6	3.94	ISRIL

DLW, dry leave weight
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.19. QTLs affecting dry biomass in the ISRIL population using composite interval mapping.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
DB_2011	qDB2011-1	1	68.81	62.15-73.15	3.5	-18.6	4.02	ISRIL
	qDB2011-3	3	116.61	112.95-119.95	3.2	18	3.35	ISRIL
	qDB2011-7	7	89.01	82.40-94.95	3.6	-14.5	4.07	ISRIL
DB_2012	qDB2012-3	3	117.61	113.80-125.15	3.1	15.1	3.55	ISRIL
	qDB2012-7	7	89.01	81.50-94.7	3.5	-15.5	3.79	ISRIL

DB, dry biomass
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.20. QTLs affecting leaf area in ISRIL using composite interval mapping method.

Trait	QTL	Chr.	Position (cM)	Interval (cM)[a]	LOD	Additive	R^2 (%)[b]	Source[c]
LAR_2011	qLAR2011-1	1	18.51	17.20-21.40	3.5	21.1	3.59	ISRIL
	qLAR2011-3	3	114.81	110.20-118.25	5.9	24.8	6.19	ISRIL
	qLAR2011-7	7	92.01	79.75-100.50	3.2	19.5	3.92	ISRIL
LAR_2012	qLAR2012-1	1	25.71	23.90-27.15	5.2	33.9	5.45	ISRIL
	qLAR2012-3	3	110.31	106.10-113.8	5.8	28.9	6.35	ISRIL
	qLAR2012-7	7	95.01	89.00-103.65	4.3	26.4	5.4	ISRIL
	qLAR2012-8a	8	82.61	78.20-86.00	6.9	30.7	7.09	ISRIL
	qLAR2012-8b	8	21.31	11.20-39.4	2.9	24.5	4.55	ISRIL
	qLAR2012-8c	8	62.41	50.50-69.80	3.1	23.8	3.93	ISRIL

LAR, leaf area
[a]1.5-LOD support interval of the QTL
[b]Percentage of the variation explained by the QTL
[c]ISRIL: *S. biocolor* BTx623 X *S. biocolor* IS3620C, PQRIL: *S. biocolor* BTx623 X *S. propinquum*

Table 2.21. Overlapping QTL regions based on LS means in the ISRIL population.

Chr.	PH	DB	DSW	DLW	LAR	LL	LW	LA	MR	Start[a] (cM)	End[a] (cM)
1	-	qDB1.1	qDSW1.1	qDLW1.1	-	-	-	-	-	66.4	73.5
3	-	-	-	-	-	qLL3.1	qLW3.1	qLA3.1	-	105.4	119.0
7	qPH7.1	qDB7.1	qDSW7.1	-	qLAR7.1	-	-	-	-	83.75	102.6
9	-	-	-	-	-	qLL9.1	-	qLA9.1	-	112.8	122.0

PH, plant height; DB, dry biomass; DLW, dry leave weight; DSW, dry stalk weight; LA, leaf angle; LAR, leaf area; LL, leaf length; LW, leaf width; MR, diameter of midribs

[a]1.5-LOD support interval of the QTL.

Table 2.22. Overlapping QTL regions in the ISRIL population for individual environments.

Chr.	PH	DB	DSW	DLW	LAR	LL	LW	LA	MR	Start[a] (cM)	End[a] (cM)
1	-	-	-	-	qLAR2011-1 qLAR2012-1	qLL2011-1 qLL2012-1a	-	-	qMR2012-1	16.8	27.15
1	-	qDB2011-1	-	-	-	qLL2012-1b	qDLW2011-1	-	-	62.15	73.15
3	-	qDB2011-3 qDB2012-3	qDSW2012-3	qDLW2012-3	qLAR2011-3	qLL2011-3	qLW2012-3	qLA2011-3a	qMR2011-3 qMR2012-3	109.15	125.15
3	-	-	-	-	qLAR2012-3	qLL2012-3	-	-	-	106.1	113.8
6	qPH2011-6 qPH2012-6	-	qDSW2011-6 qDSW2012-6	-	-	-	-	-	-	5.7	26.2
6	-	-	-	-	-	-	qLW2011-6 qLW2012-6	-	qMR2012-6	62	75.7
7	qPH2011-7	qDB2012-7 qDB2011-7	qDSW2011-7 qDSW2012-7	-	-	-	qLW2011-7	-	-	81.5	94.95
7	qPH2012-7	-	-	-	qLAR2011-7	-	-	qLA2011-7 qLA2012-7	-	79.75	100.5
7	-	-	-	-	qLAR2012-7	-	qLW2012-7	-	-	89	103.65
8	-	-	-	qDLW2011-8	qLAR2012-8a	qLL2012-8	qLW2012-8	-	-	74.2	99.7
9	-	-	-	-	-	qLL2012-9	-	qLA2011-9	-	114.15	122

DB, dry biomass; DLW, dry leave weight; DSW, dry stalk weight; LA, leaf angle; LAR, leaf area; LL, leaf length; LW, leaf width; MR, diameter of midribs; PH, plant height.
[a]1.5-LOD support interval of the QTL.

Table 2.23. Overlapping QTL regions in the PQRIL population for individual environments.

Chr.	PH	DF	LL	LW	LA	Start (cM)[a]	End (cM)[a]
7				qLW2011-7	qLA2009-7	32.55	46
					qLA2010-7		
					qLA2011-7		
8			qLL2010-8		qLA2011-8	63.3	76.6

DF, day to flowering; LA, leaf angle; LAR, leaf area; LL, leaf length; LW, leaf width; PH, plant height.
[a]1.5-LOD support interval of the QTL.

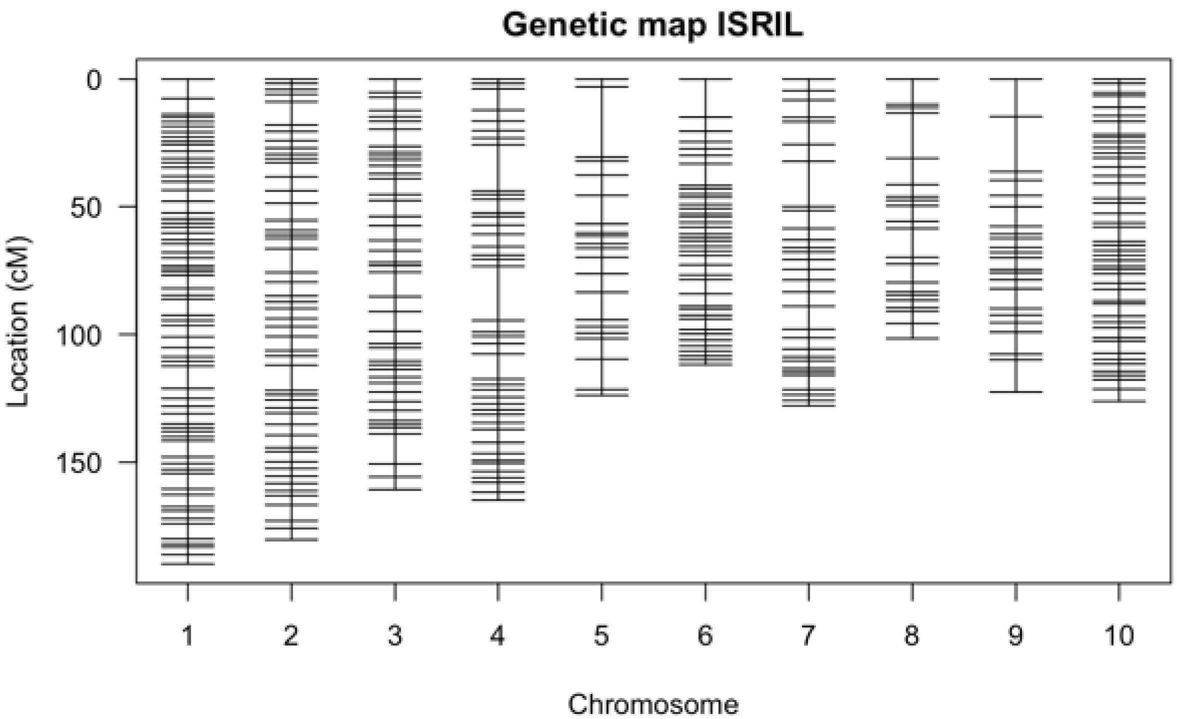

Figure 2.1. Genetic map of *S. bicolor* BTx623 X *S. bicolor* IS3620C (ISRIL) population constructed after merging markers within 1 cM into single bins based on Kong et al., 2018.

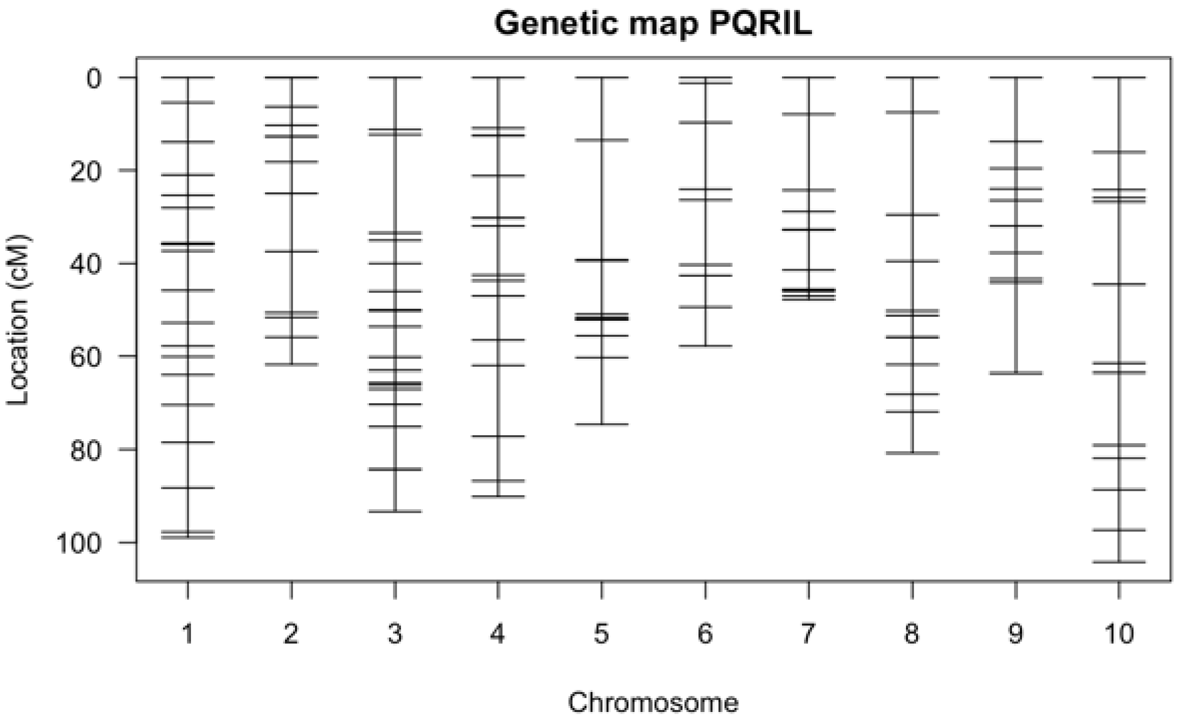

Figure 2.2. Genetic map of *S. bicolor* BTx623 X *S. propinquum* (PQRIL) population spanning a genetic distance of 773.1 cM, with an average interval between consecutive markers of 5.48 cM based on Kong et al., 2015.

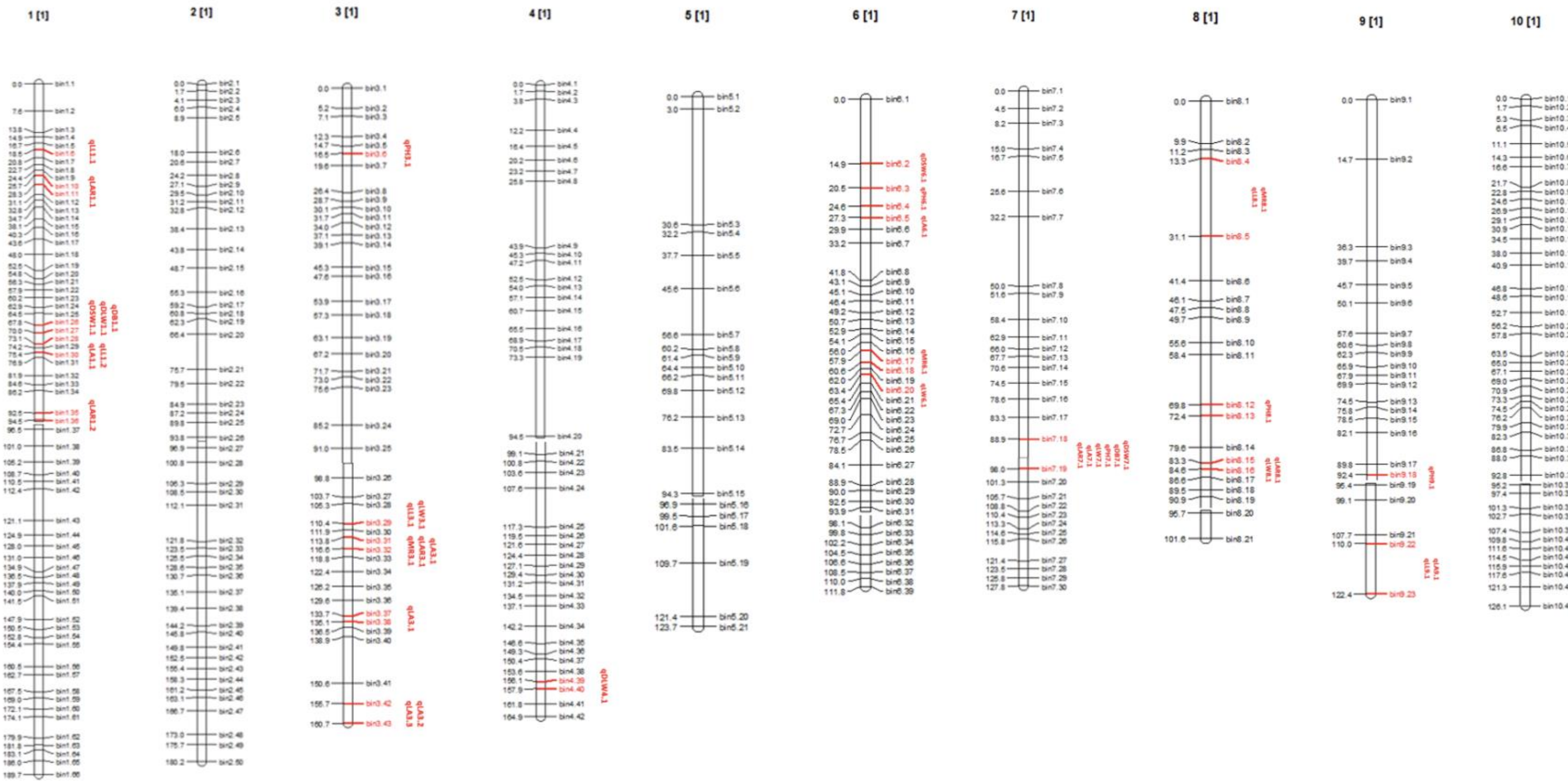

Figure 2.3. Projection of all QTLs detected in ISRIL using overall LS means.

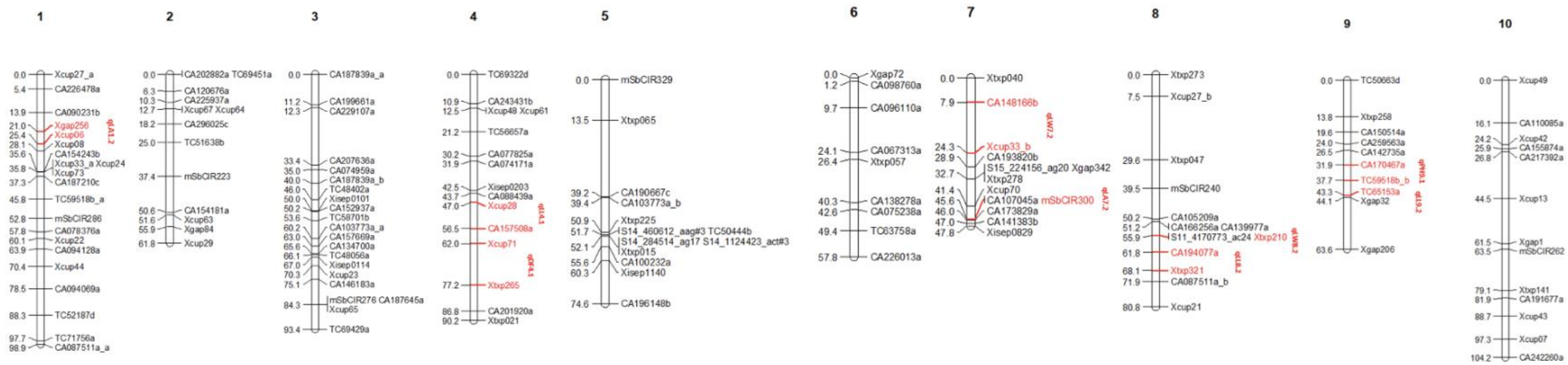

Figure 2.4. Projection of all QTLs detected in PQRIL using overall LS means.

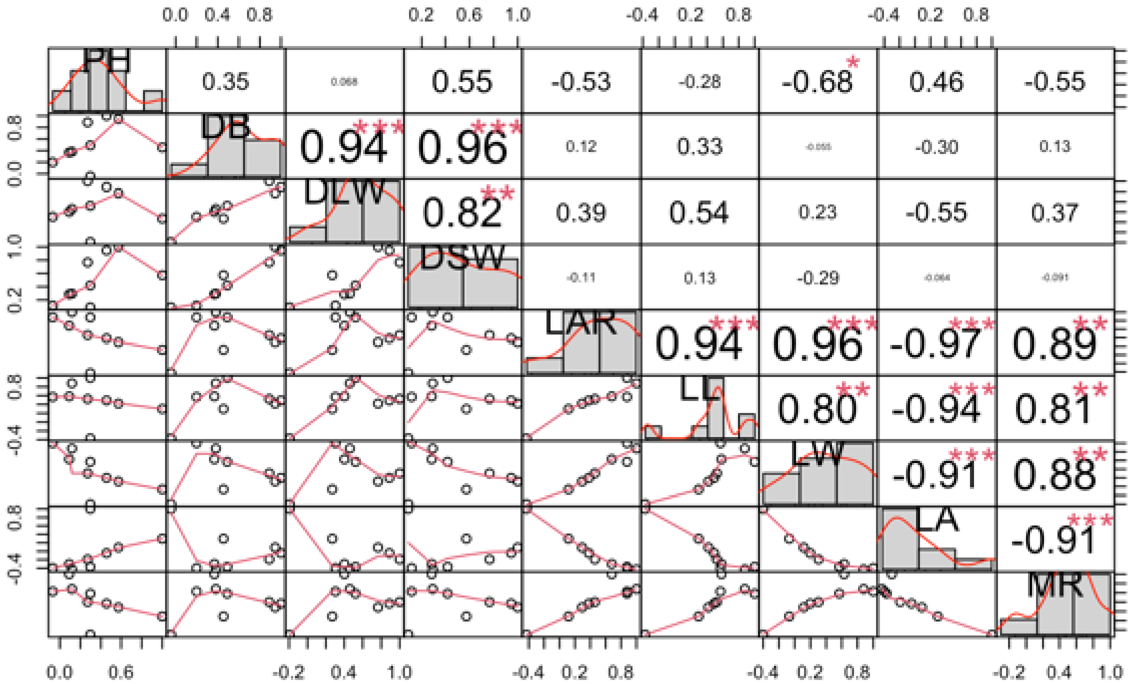

Figure 2.5. Variation and Pearson pairwise correlations among leaf morphology and yield-related traits from the ISRIL population. To the left and the diagonal are scatter plots and histograms representing correlation coefficients for each phenotypic trait (PH; plant height, DB; dry biomass, DLW; dry leaf weight, DSW; dry stalk weight, LAR; leaf area, LL; leaf length, LW; leaf width, LA; leaf angle, MR; midrib diameter). The red line through the scatter plot represents the line of best fit. Pearson correlation coefficients among the traits are shown above and to the right of the diagonal. The correlation significance levels are: *p = 0.05, **p = 0.01, and ***p = 0.001, and the size of the coefficient values are proportional to the strength of the correlation.

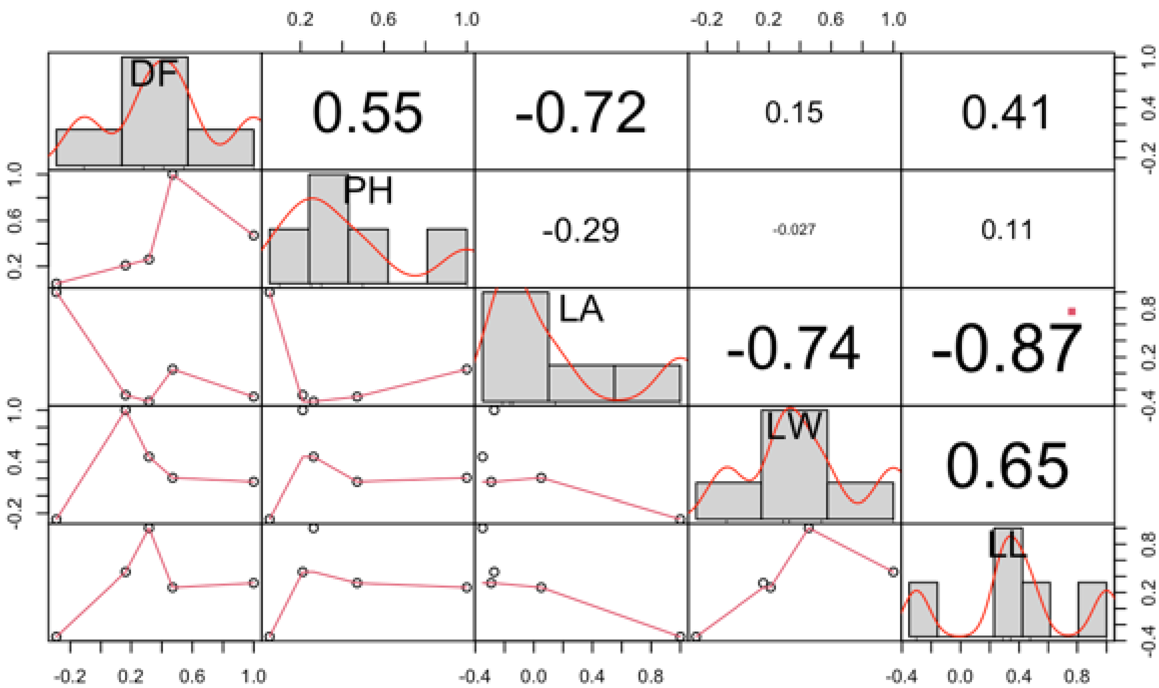

Figure 2.6. Variation and Pearson pairwise correlations among leaf morphology and yield-related traits from the PQRIL population. To the left and the diagonal are scatter plots and histograms representing correlation coefficients for each phenotypic trait (DF; days to flower, PH; plant height, LA; leaf angle, LW; leaf width, LL; leaf length). The red line through the scatter plot represents the line of best fit. Pearson correlation coefficients among the traits are shown above and to the right of the diagonal. The correlation significance levels are: *p = 0.05, **p = 0.01, and ***p = 0.001, and the size of the coefficient values are proportional to the strength of the correlation.

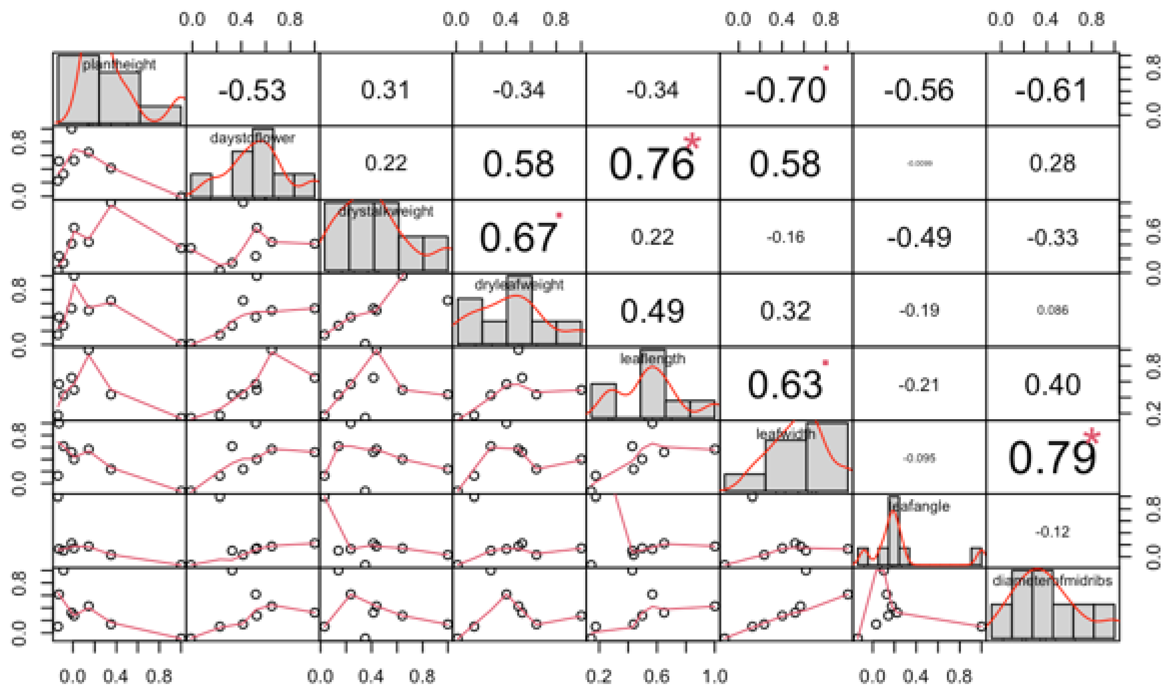

Figure 2.7. Variation and Pearson pairwise correlations among leaf morphology and yield-related traits from the sorghum diversity panel. To the left and the diagonal are scatter plots and histograms representing correlation coefficients for each phenotypic trait. The red line through the scatter plot represents the line of best fit. Pearson correlation coefficients among the traits are shown above and to the right of the diagonal. The correlation significance levels are: *p = 0.05, **p = 0.01, and ***p = 0.001, and the size of the coefficient values are proportional to the strength of the correlation.

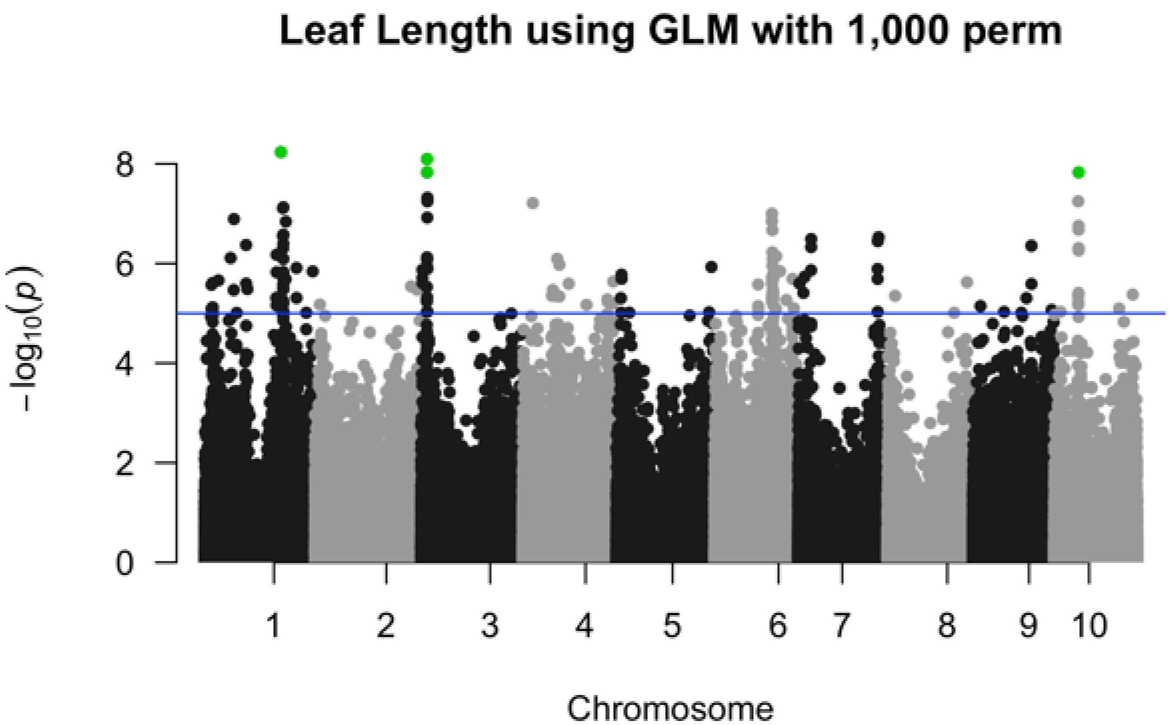

Figure 2.8. Manhattan plots displaying genome-wide association study result for leaf length from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with leaf length. The blue horizontal line denotes the suggestive line at -log10(1e-5). Those SNPs below the threshold from an empirical p-value from the permutations (*P* < 0.01) are highlighted as green on the plot.

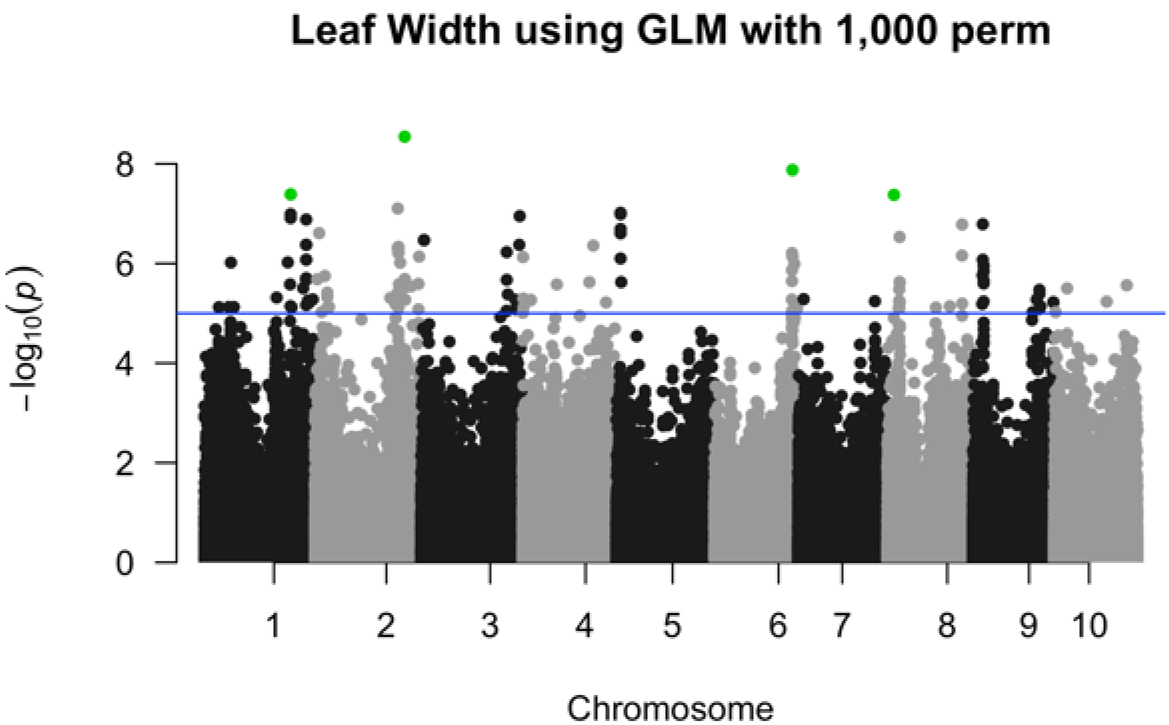

Figure 2.9. Manhattan plots displaying genome-wide association study result for leaf width from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with leaf width. The blue horizontal line denotes the suggestive line at -log10(1e-5). Those SNPs below the threshold from an empirical p-value from the permutations ($P < 0.01$) are highlighted as green on the plot.

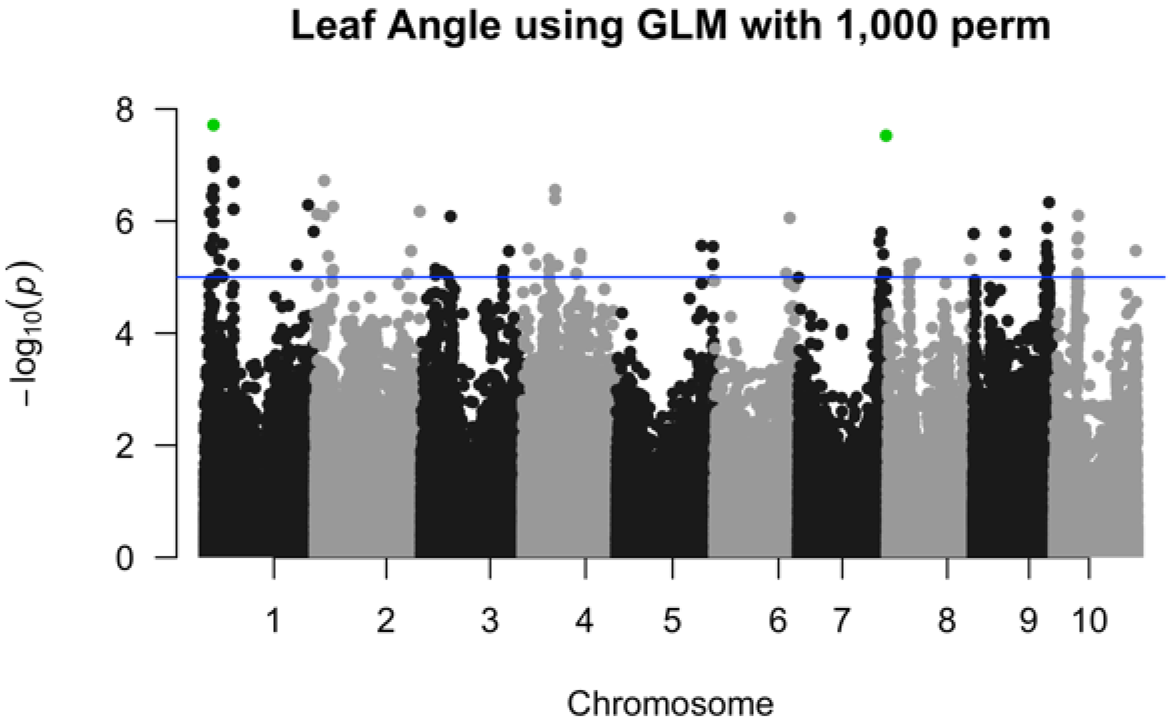

Figure 2.10. Manhattan plots displaying genome-wide association study result for leaf angle from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with leaf angle. The blue horizontal line denotes the suggestive line at -log10(1e-5). Those SNPs below the threshold from an empirical p-value from the permutations ($P < 0.01$) are highlighted as green on the plot.

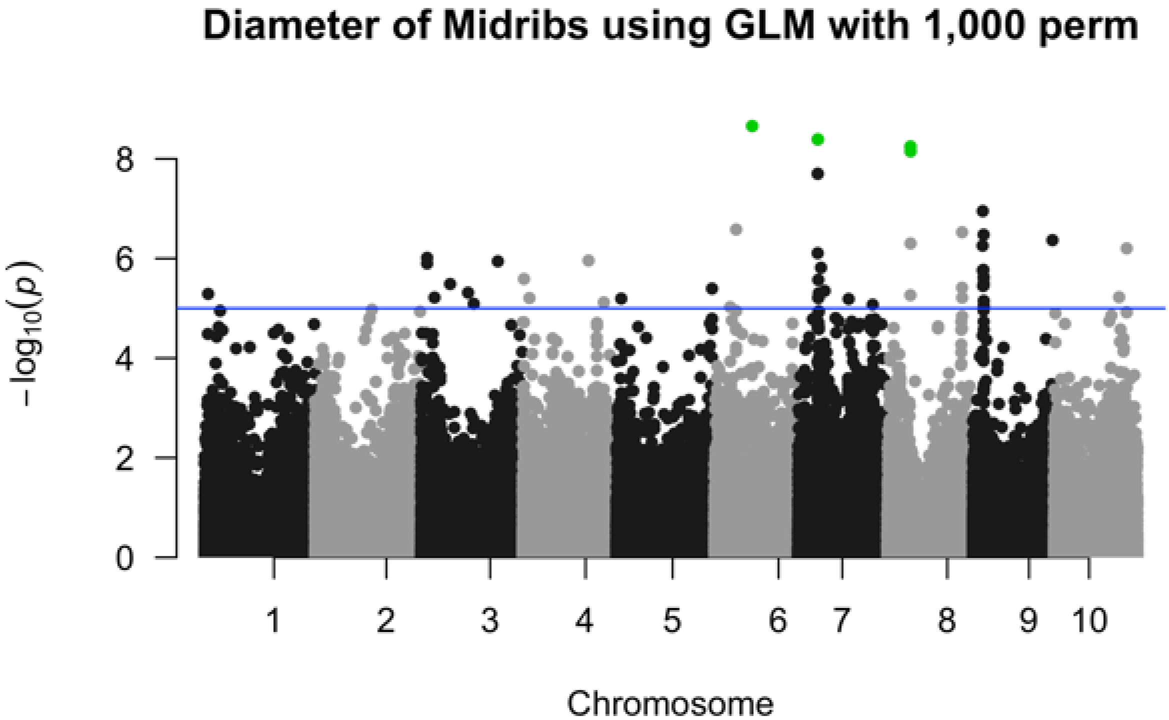

Figure 2.11. Manhattan plots displaying genome-wide association study result for diameter of midribs from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with diameter of midribs. The blue horizontal line denotes the suggestive line at -log10(1e-5). Those SNPs below the threshold from an empirical p-value from the permutations ($P < 0.01$) are highlighted as green on the plot.

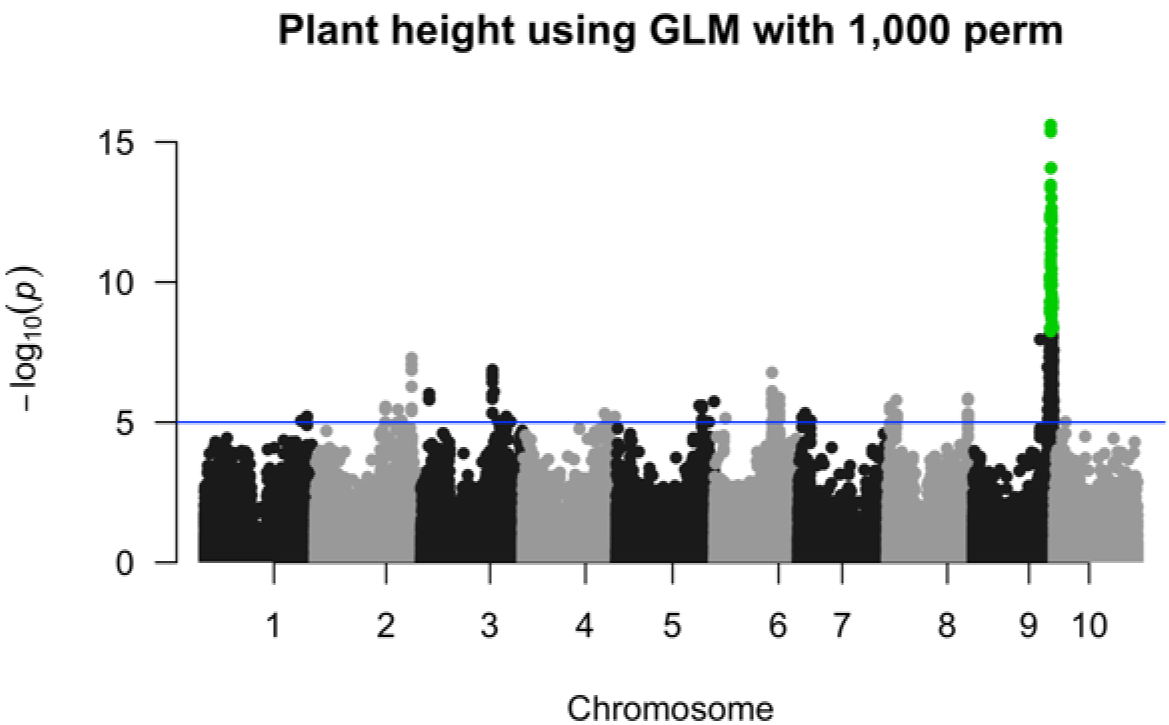

Figure 2.12. Manhattan plots displaying genome-wide association study result for plant height from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with plant height. The blue horizontal line denotes the suggestive line at -log10(1e-5). Those SNPs below the threshold from an empirical p-value from the permutations (*P* < 0.01) are highlighted as green on the plot.

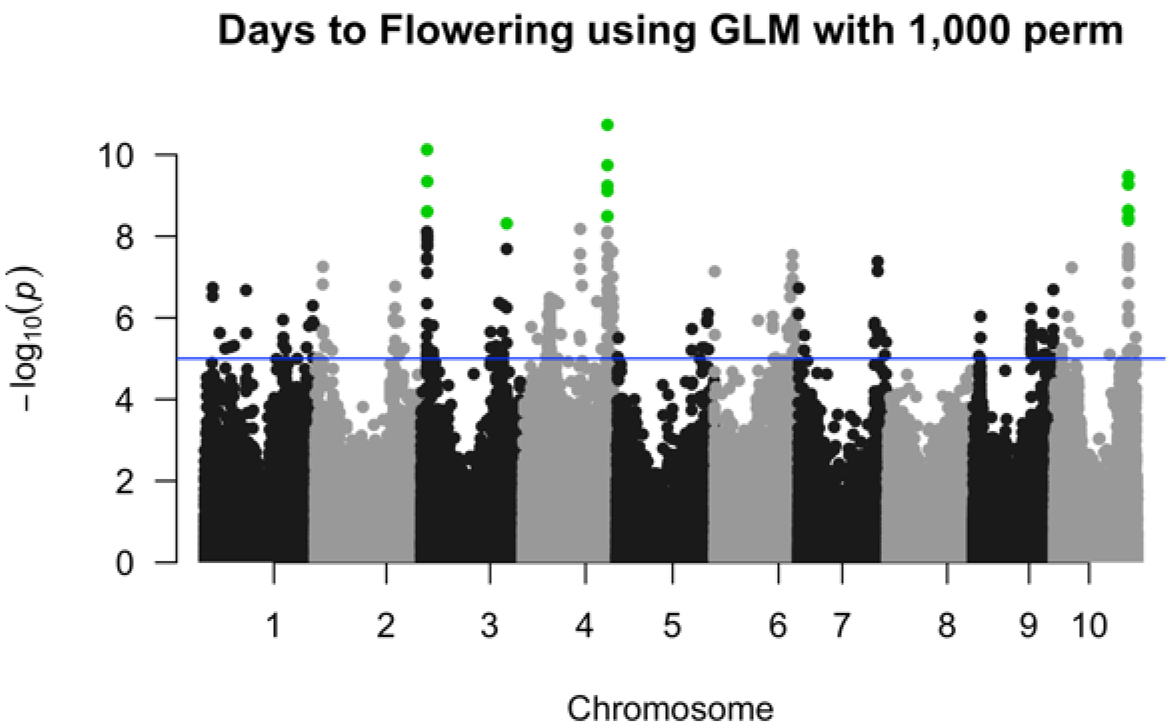

Figure 2.13. Manhattan plots displaying genome-wide association study result for days to flowering from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with days to flowering. The blue horizontal line denotes the suggestive line at -log10(1e-5). Those SNPs below the threshold from an empirical p-value from the permutations ($P < 0.01$) are highlighted as green on the plot.

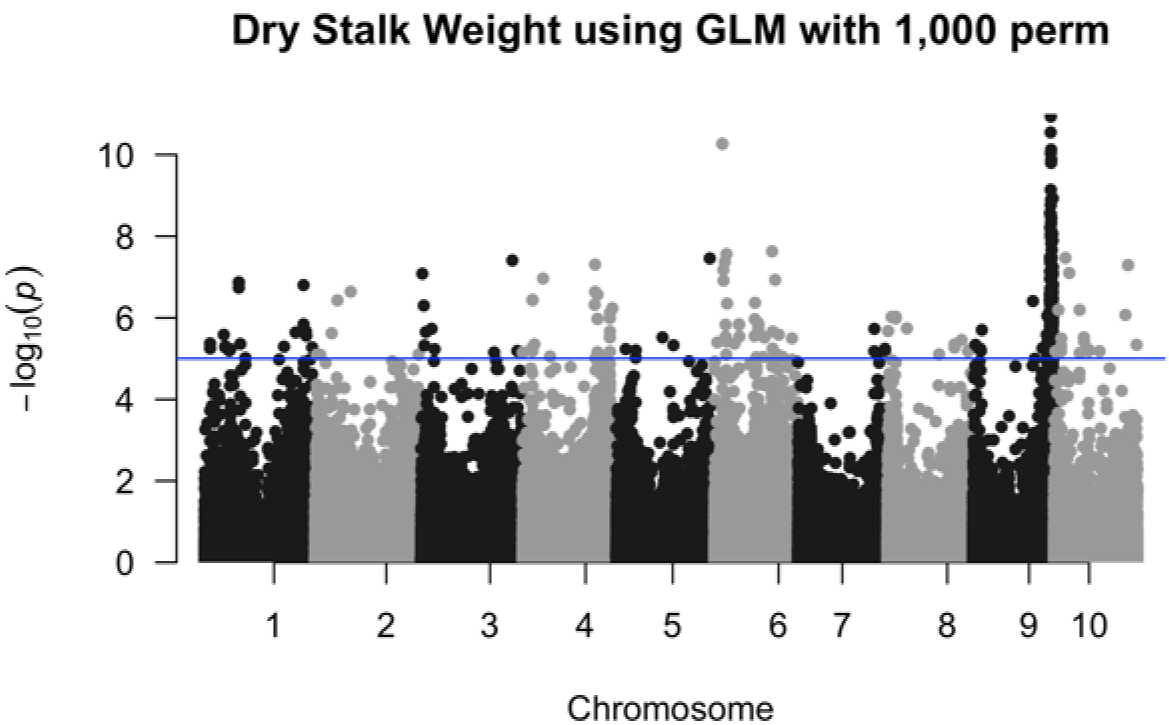

Figure 2.14. Manhattan plots displaying genome-wide association study result for dry stalk weight from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with dry stalk weight. The blue horizontal line denotes the suggestive line at -log10(1e-5). There are no SNPs below the threshold from an empirical p-value from the permutations ($P < 0.01$).

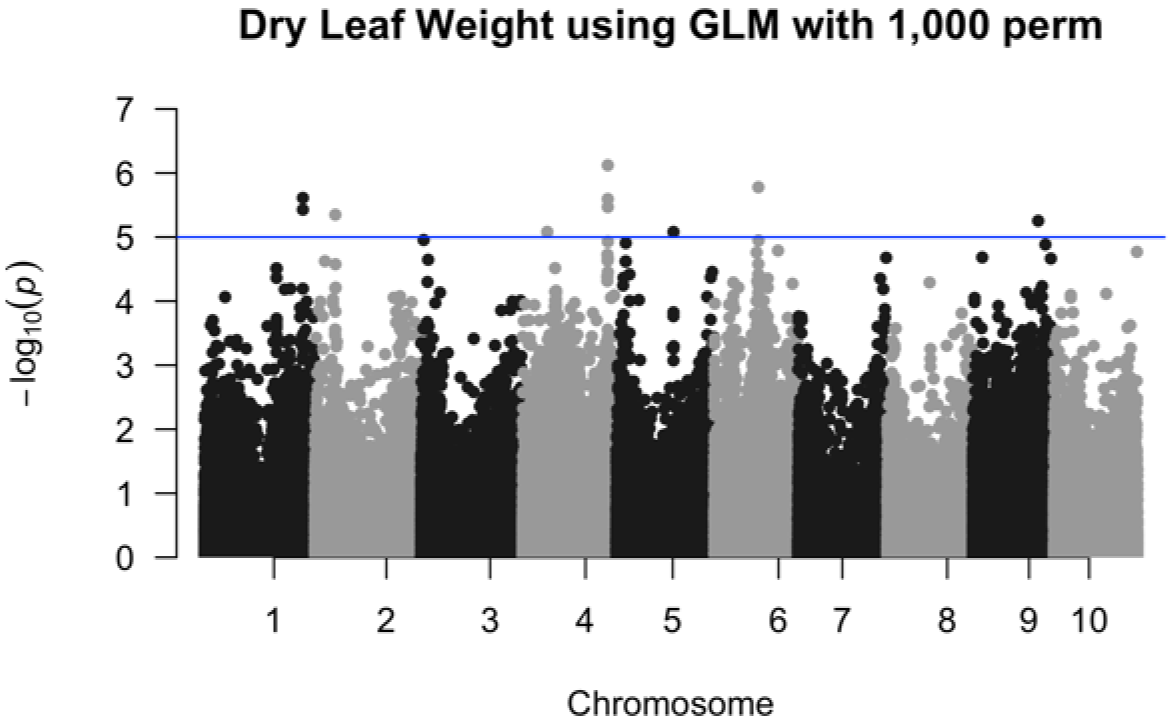

Figure 2.15. Manhattan plots displaying genome-wide association study result for dry leaf weight from the general linear model (GLM) in TASSEL. Physical single-nucleotide polymorphism (SNP) position on the genome is provided on the x-axis. The y-axis shows SNPs associated with dry leaf weight. The blue horizontal line denotes the suggestive line at -log10(1e-5). There are no SNPs below the threshold from an empirical p-value from the permutations (P < 0.01).

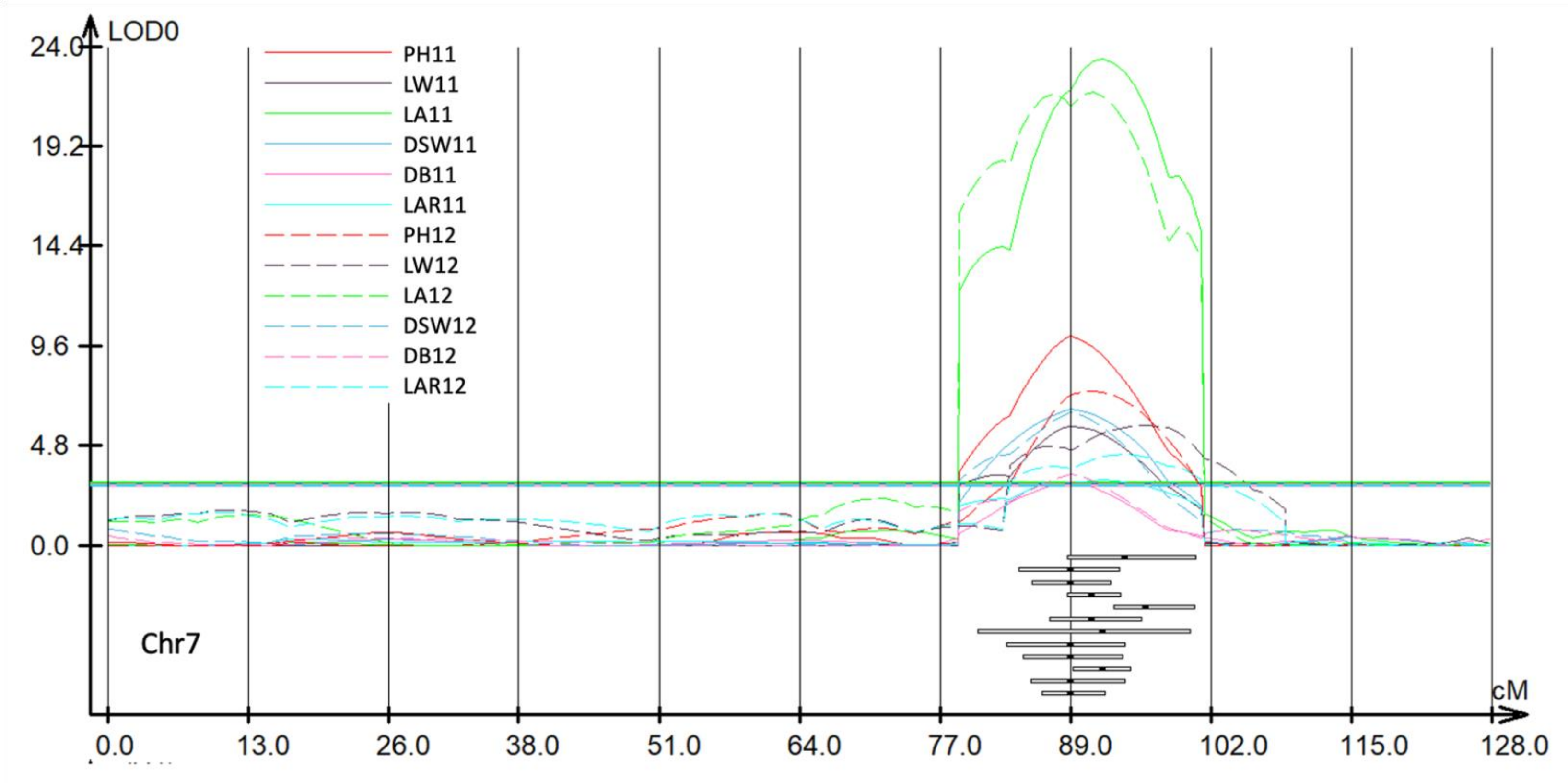

Figure 2.16. Twelve colocalized intervals for plant height, dry biomass, dry stalk weight, leaf area, leaf angle, and leaf width to locate near each other on chromosome 7 from 2011 and 2012 environments by QTL mapping.

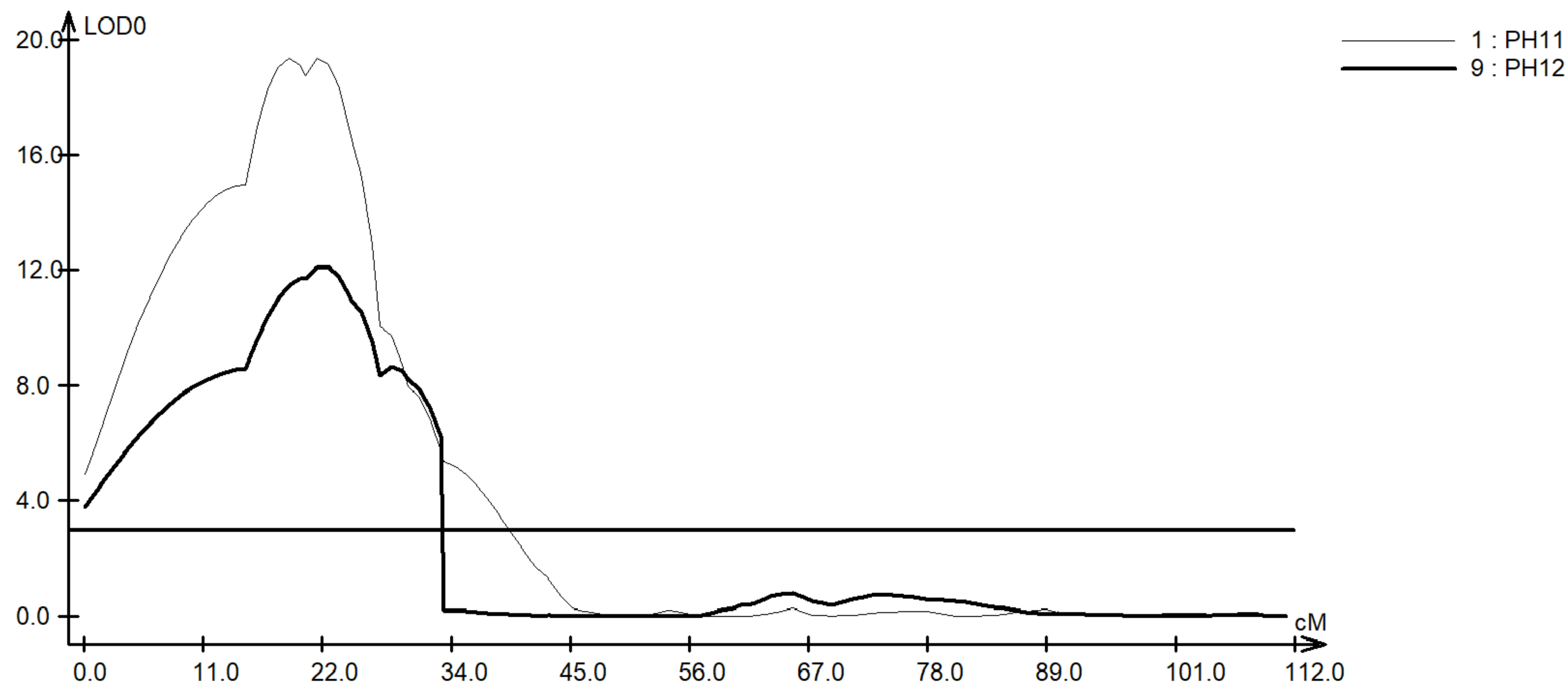

Figure 2.17. A significant stable QTL for plant height called *qPH6.1* at the physical distance around 42.4 - 45.5 Mb in the study with LOD score od 21, which accounted for 21.47% of the phenotypic variance from the ISRIL population.

CHAPTER 3

CONCLUSION

Increasing yield production has been a primary goal for improving sorghum cultivars, which is one of the most important cereal crops in the world. Leaves play a crucial part in the morphological and physiological functioning of plants. Leaves are the primary sites for photosynthesis, the process by which chlorophyll in leaves converts light energy into chemical energy. The leaf receives solar energy to oxidize water, generate sugar from carbon dioxide, and release oxygen into the atmosphere. Consequently, leaf architecture directly influences plant growth and yield. Yield production in sorghum involves other morphological traits, including plant height, leaf area, dry stalk and leaf weight, and dry biomass. Flowering time is also regarded as a significant characteristic because it controls the period of crop biomass accumulation. Identification of genetic regions affecting essential leaf morphology and yield-related characteristics will provide useful resources for a variety of applications, including marker-assisted selection in sorghum breeding programs.

This study attempted to discover genomic loci influencing leaf morphological attributes in two populations, IS and PQRIL, that represented the majority of the genetic variation present in 'eusorghums'. Both populations had a common ancestor, the elite inbred S. bicolor BTx623 that served as the basis of the sorghum reference genome. This parent was crossed with S. bicolor IS3620, representing race 'guinea' that is substantially divergent from BTx623, or *S. propinquum*, a sister species within the genus Sorghum, to produce IS and PQRIL populations, respectively. In both sorghum populations, genotyping data derived from SSR and genotyping-by-sequencing (GBS) were utilized to undertake quantitative trait locus (QTL) mapping and genome-wide association study (GWAS). GWAS typically produce high-resolution results, however it has a higher false-positive rate than QTL mapping. In contrast, QTL mapping has a low resolution but a decreased false positive rate. GWAS analysis will therefore complement QTL mapping. The strategy of combining Next-Generation Sequencing (NGS) technology and computational analysis in the study will enable the clarification of the relationship between leaf shape and genes in the two distinct populations. The traits of interest include leaf length, width, angle, and diameter of midribs regarding leaf morphology as well as plant height, day to flowering, dry stalk weight, dry leaf weight, dry biomass, and leaf area regarding yield-related traits. We identified a total of 101 QTLs in the ISRIL population, and 20 QTLs in the PQRIL population. The candidate genomic intervals for the finding of QTLs were identified through comparative analysis, resulting in more precise locations from which candidate genes that may contribute to the traits of interest. Several QTLs were identified in multiple populations and/or years. Some quantitative trait loci (QTLs) governing distinct qualities were found in overlapping chromosomal areas, indicating that their inheritance may be linked. The absence of overlap in other QTLs suggests that the trait is regulated by independent loci. The identification of these genes will boost genetic resources, enabling the identification of leaf morphological variants that are more resistant to climate change and disease attacks for crop improvement.

SUPPLEMENTARY FILES

Plant height

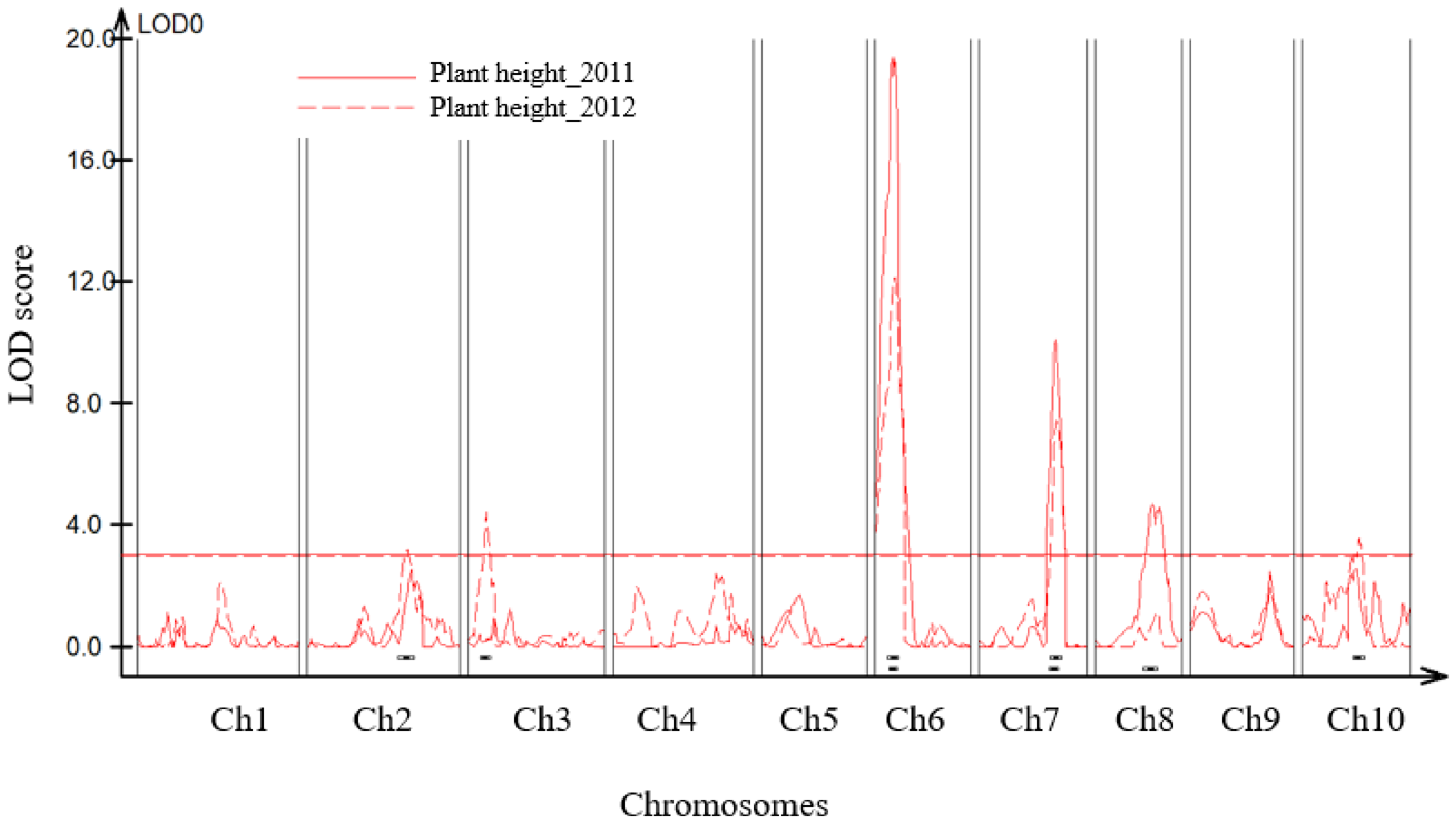

Figure S1: QTLs associated with plant height in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Dry leaf weight

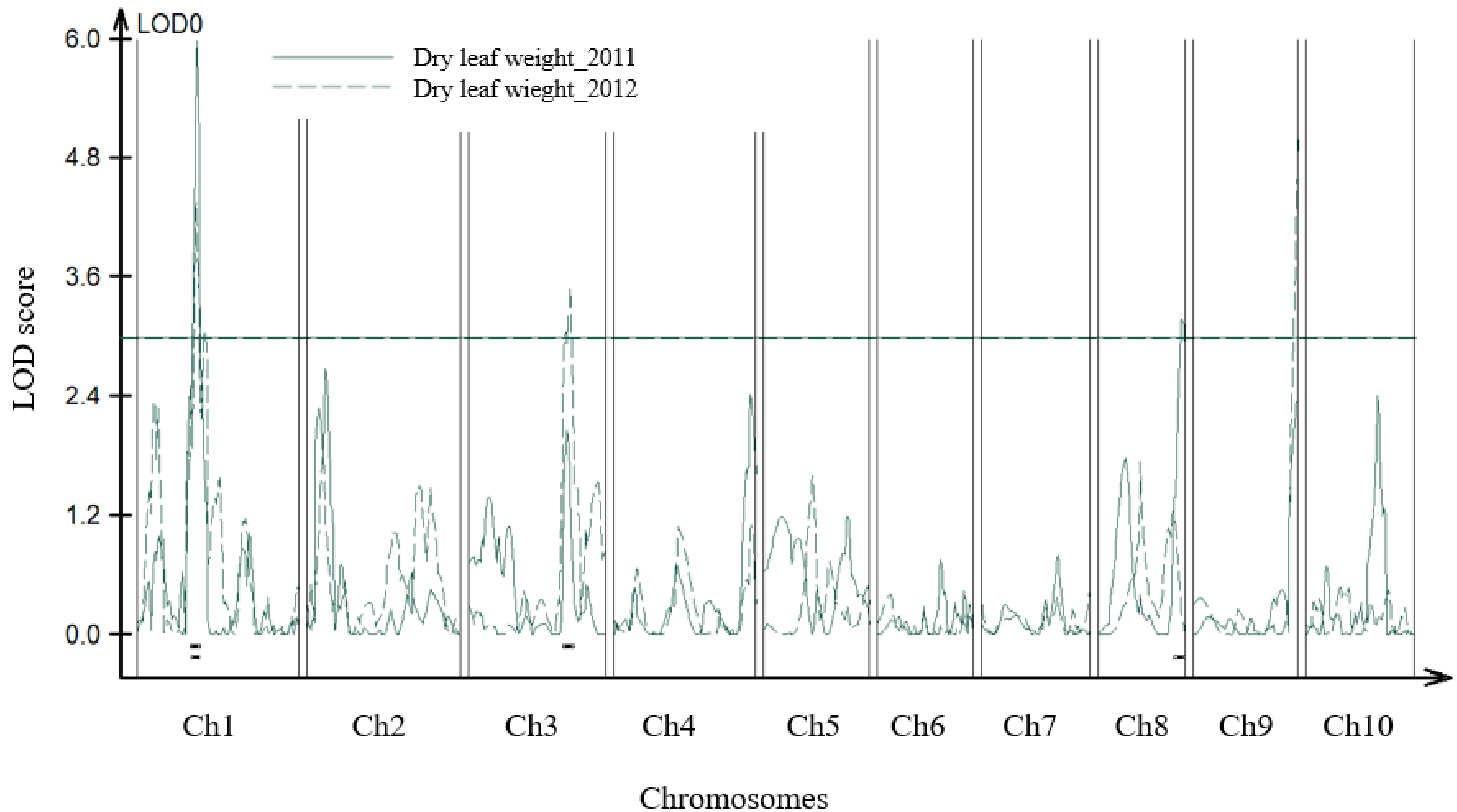

Figure S2: QTLs associated with dry leaf weight in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Dry stalk weight

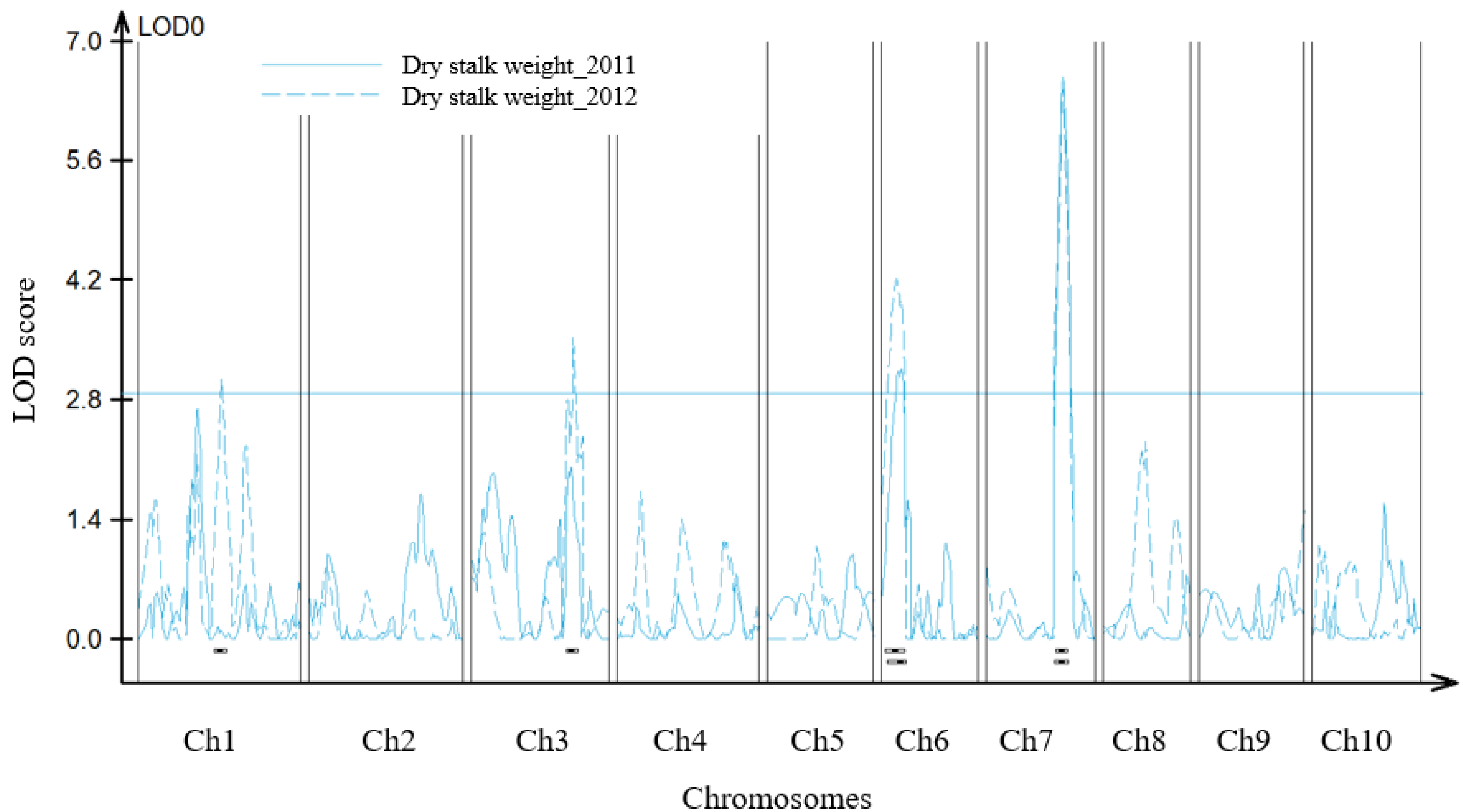

Figure S3: QTLs associated with dry stalk weight in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Dry biomass

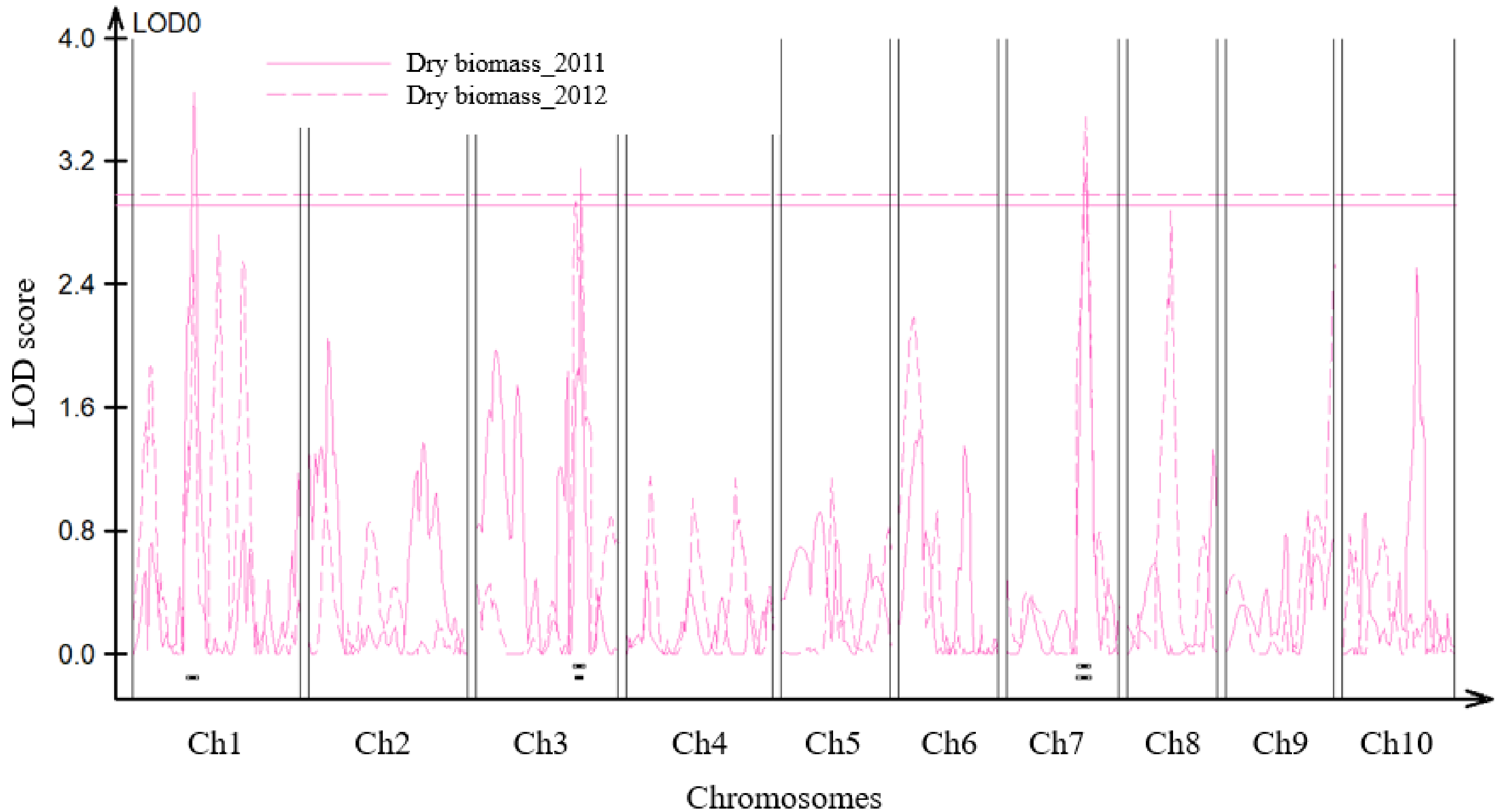

Figure S4: QTLs associated with dry biomass in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Leaf area

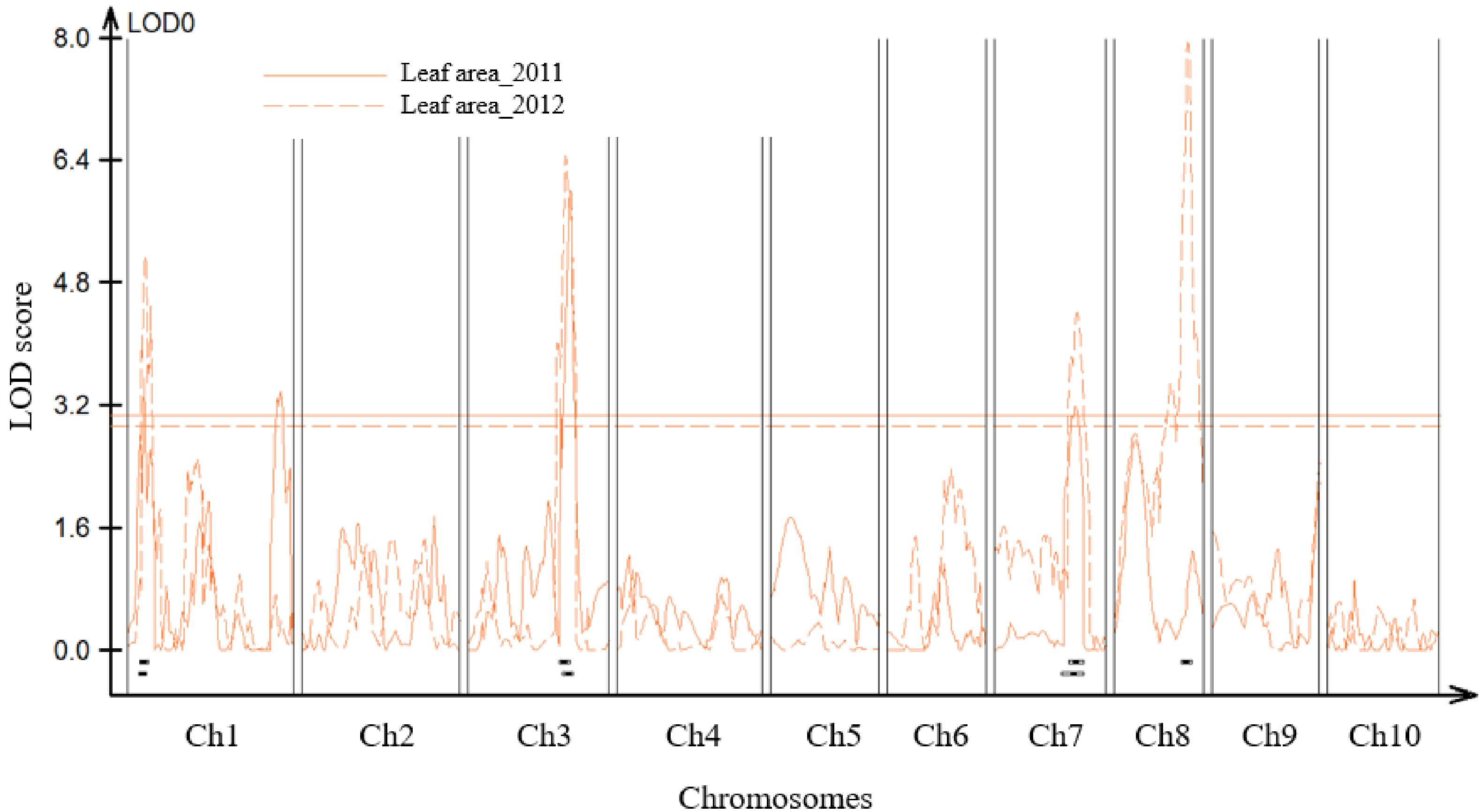

Figure S5: QTLs associated with leaf area in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Leaf length

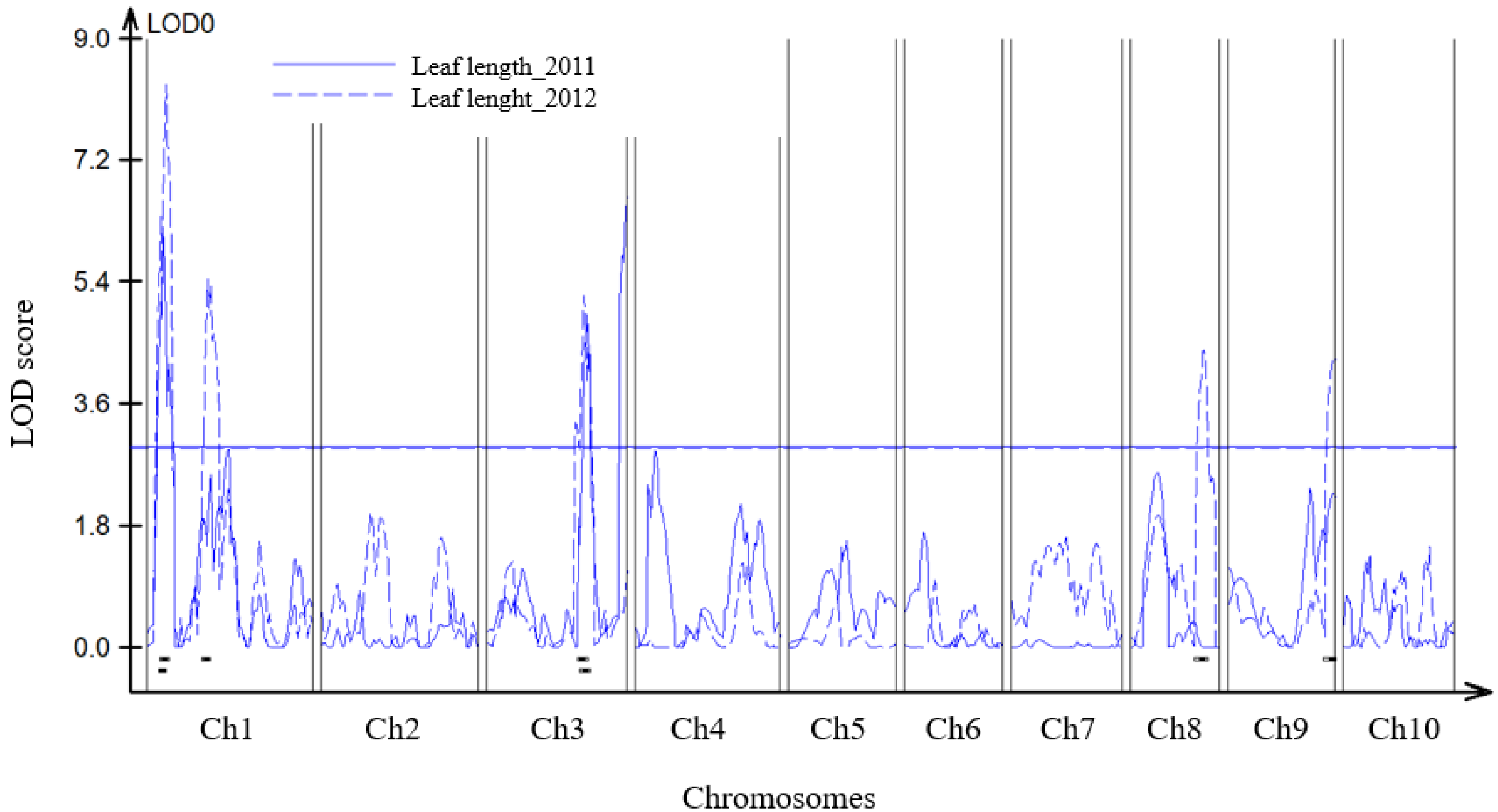

Figure S6: QTLs associated with leaf length in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Leaf length

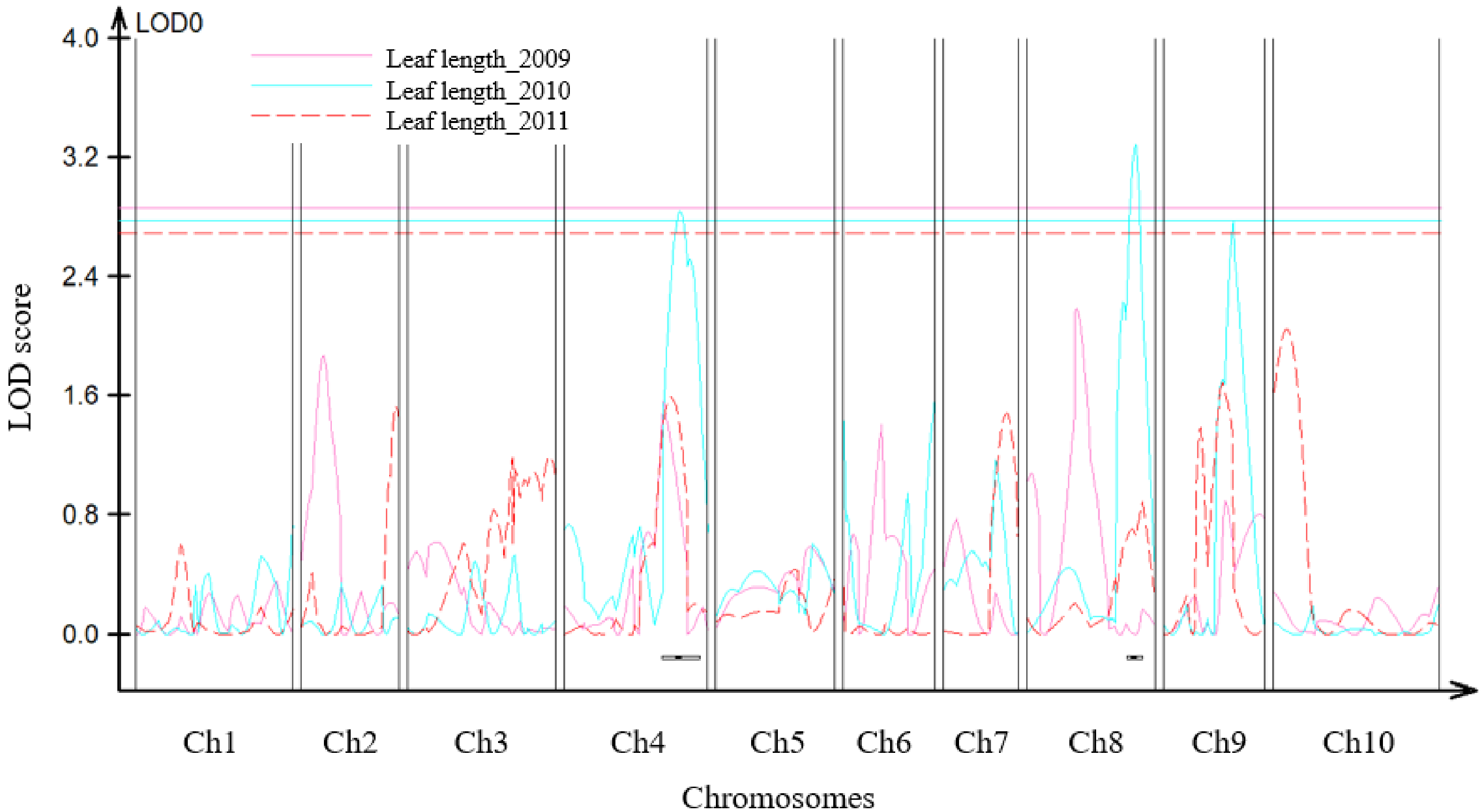

Figure S7: QTLs associated with leaf length in the *S. bicolor* BTx623 X *S. propinquum* population.

Leaf width

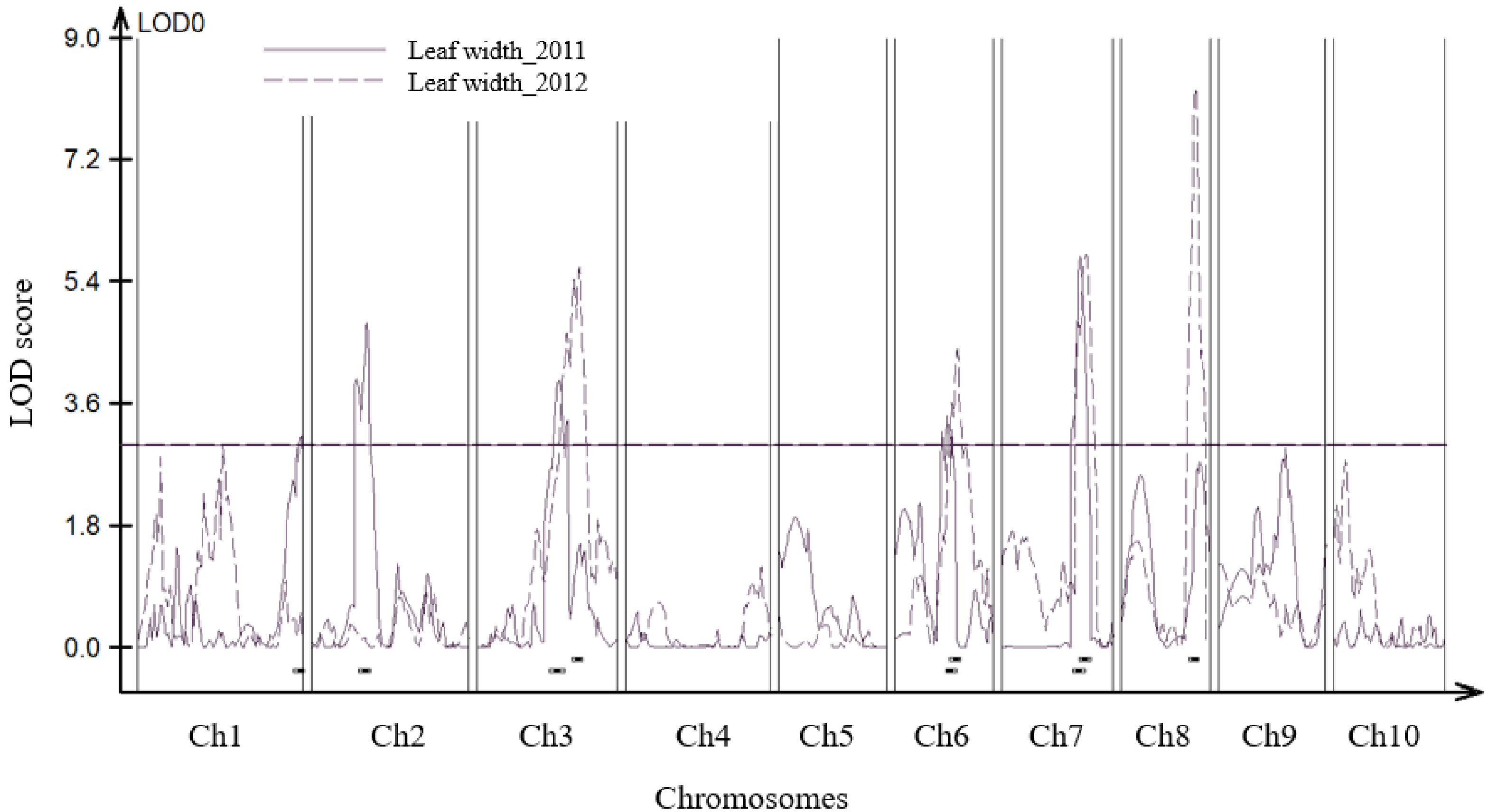

Figure S8: QTLs associated with leaf width in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Leaf width

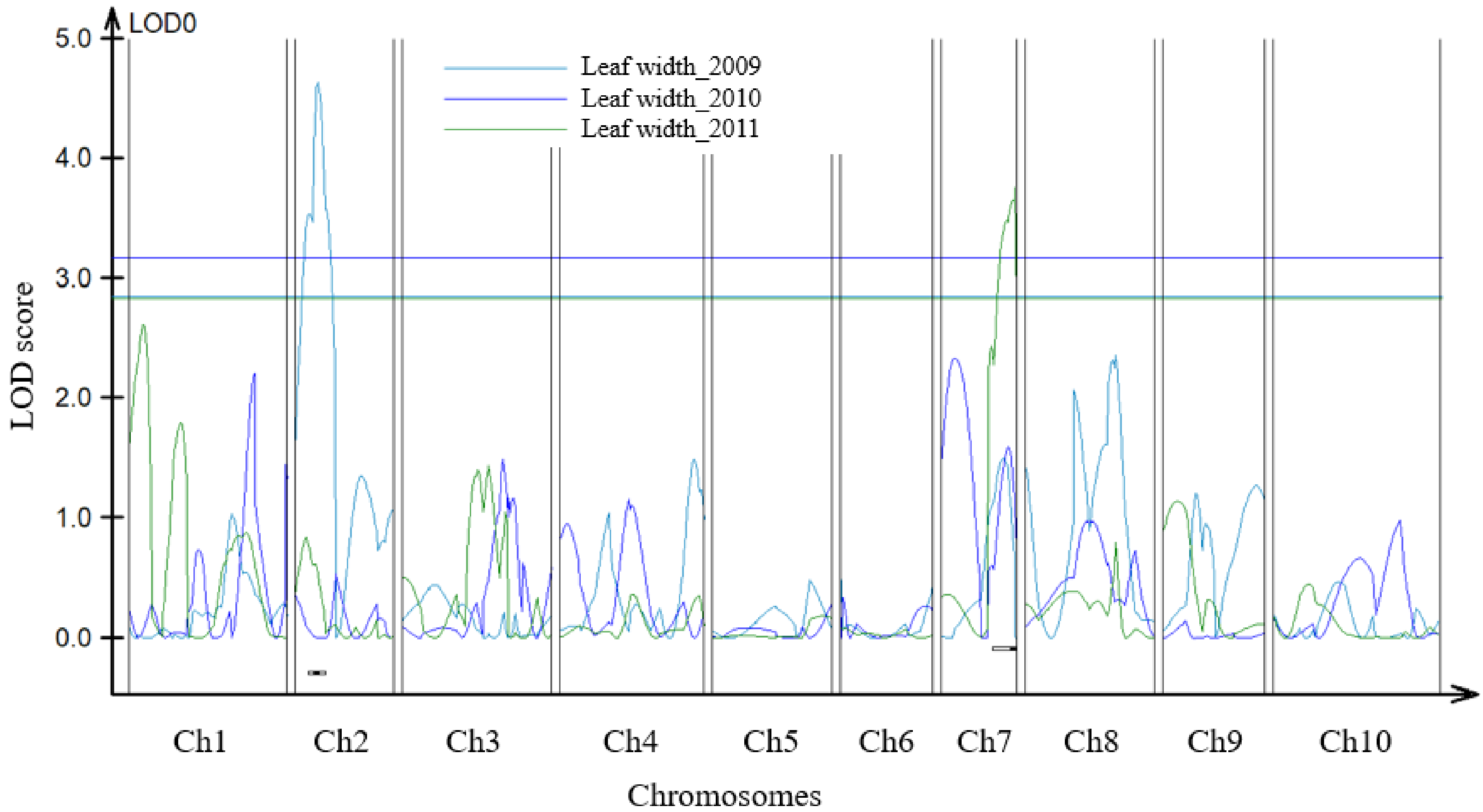

Figure S9: QTLs associated with plant height in the *S. bicolor* BTx623 X *S. propinquum* population.

Leaf angle

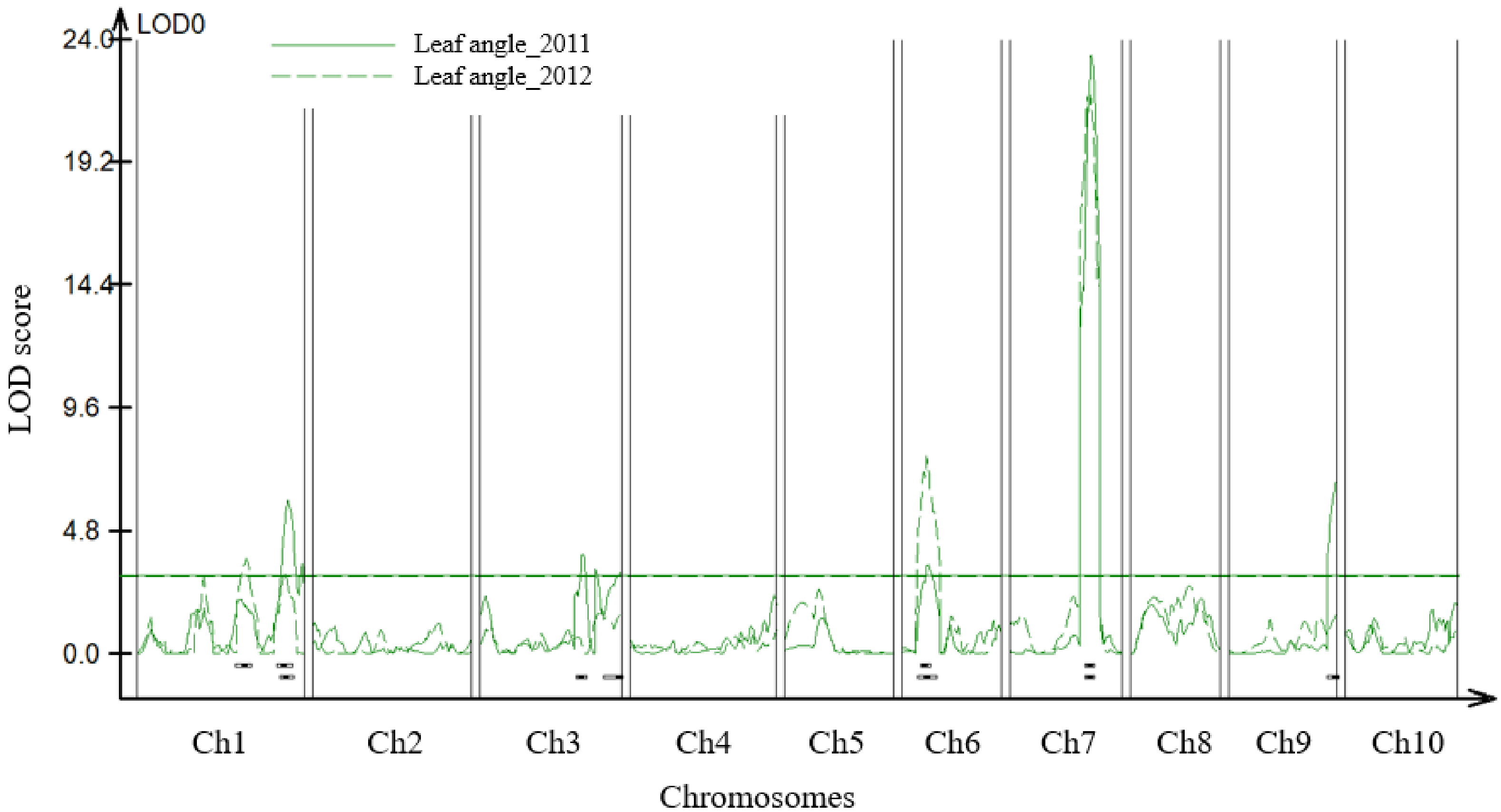

Figure S10: QTLs associated with leaf angle in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Leaf angle

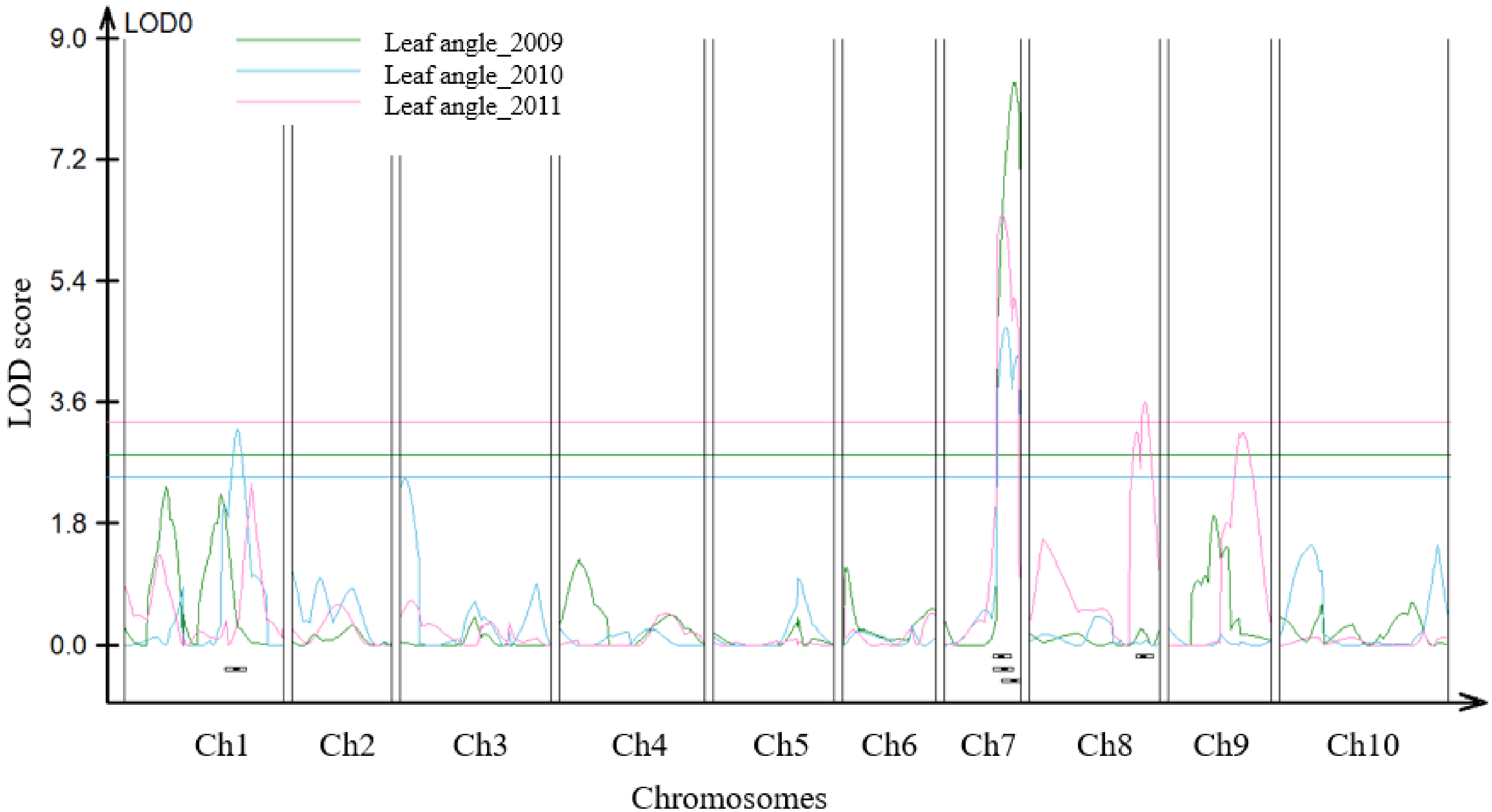

Figure S11: QTLs associated with leaf angle in the *S. bicolor* BTx623 X *S. propinquum* population.

Diameter of midribs

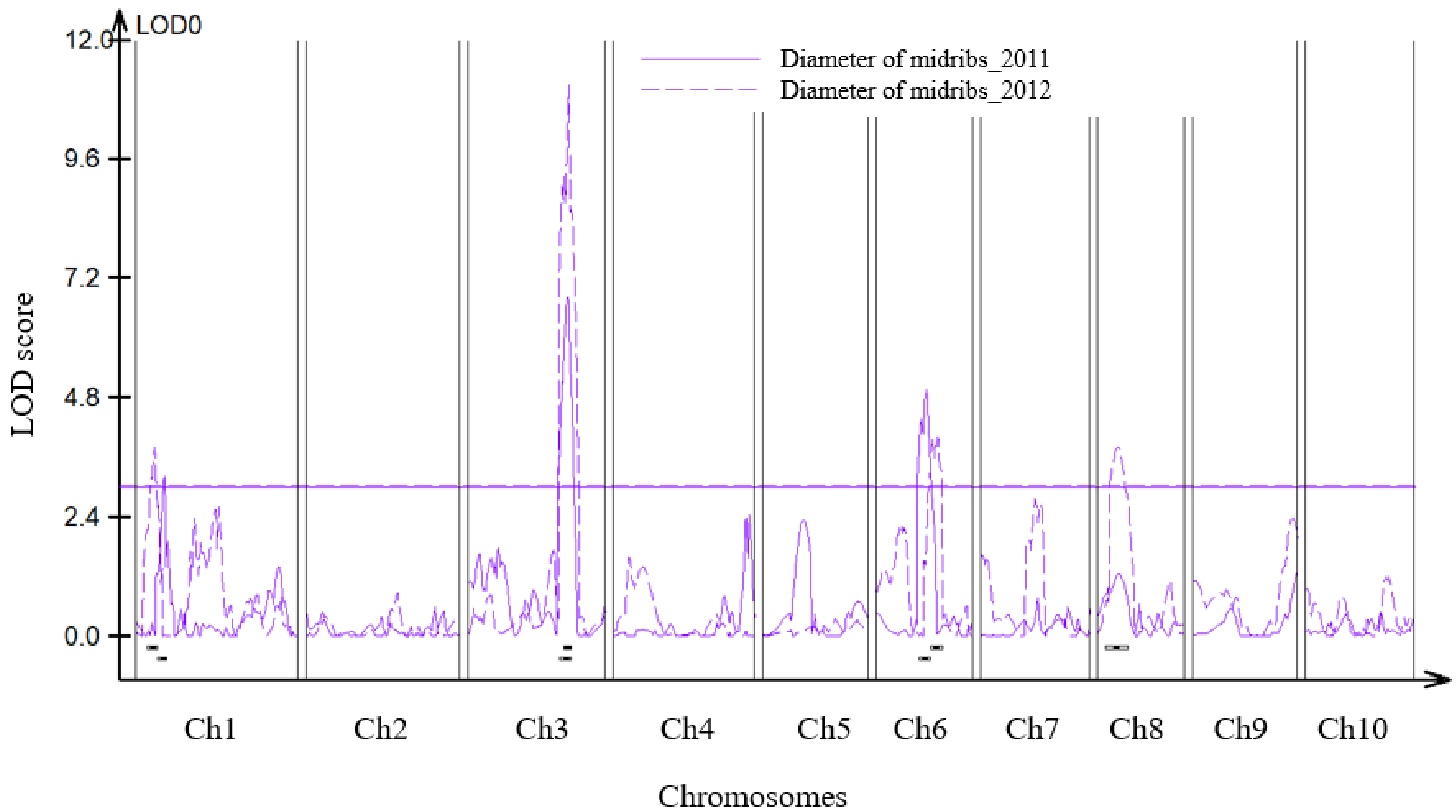

Figure S12: QTLs associated with the diameter of midribs in the *S. bicolor* BTx623 X *S. bicolor* IS3620C population.

Days to flowering

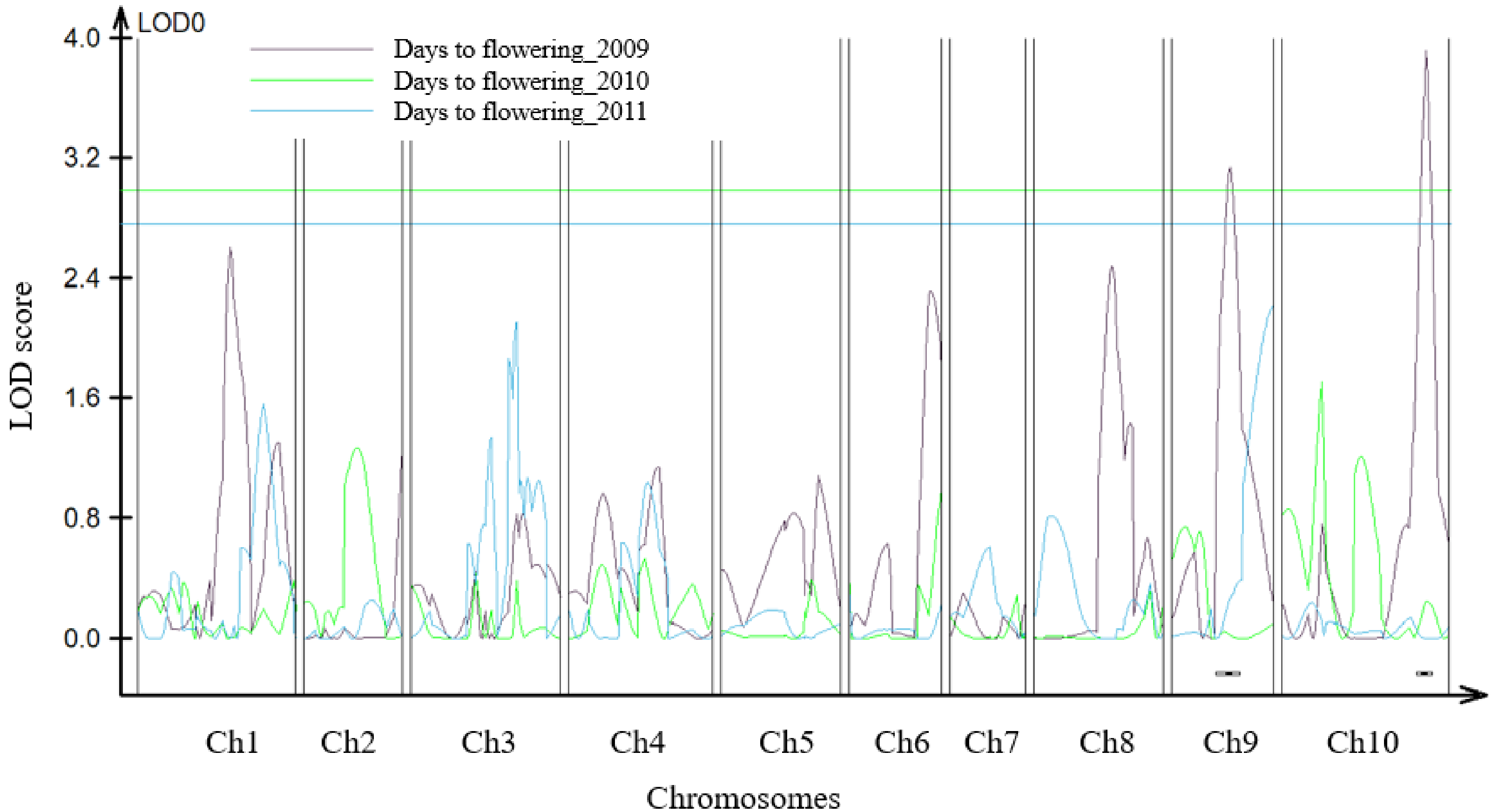

Figure S13: QTLs associated with days to flowering in the *S. bicolor* BTx623 X *S. propinquum* population.