

Justus-Liebig-University Giessen
Research Centre for Biosystems, Land Use and Nutrition
Department of Plant Breeding
Head of the Department. Prof. Dr. Dr. h.c. Wolfgang Friedt

**Identification of genomic regions of *Sorghum bicolor* (L.)
Moench linked to biofuel-related traits in grain x sweet
sorghum recombinant inbred lines**

Dissertation

Submitted for the degree of Doctor of Agricultural Science
Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management
Justus-Liebig-University Giessen

Submitted by
Amukelani Lacrechia Shiringani
From
Limpopo, Republic of South Africa
Giessen, 2009

Declaration

“I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and referenced all text passages that are derived literally from or are based on the content of published or unpublished work of others, and all information that relates to verbal communications. I have abided by the principles of good scientific conduct laid down in the charter of the Justus-Liebig-University of Giessen in carrying out the investigations described in the dissertation.”

Amukelani Lacreacia Shiringani

Place; Date

Acknowledgement

I would like to express my heartfelt gratitude to Prof. Wolfgang Friedt for allowing me to obtain my doctoral degree under his supervision. His professional advises, ideas, encouragements, and forever patience shaped my career. My special appreciation goes to Prof. Matthias Frisch for assisting and advising me to identify better method to analyse my data and for all the valuable inputs: Herzlichen Dank.

I am grateful to Mario Tolksdorf, Markus Kollmer, and Malte Luh for field technical assistance, and Swetlana Renner, Nelly Weiss, and Adriana Ochoa Fandinō for their assistance with the fibre content analysis. I am thankful to all my colleagues in the Institute of Crop Science and Plant Breeding I for sharing their experiences with me, assisting me, and creating the forever treasurable memory. Guys, we spent valuable time together, thumb up!

I am very grateful to the German academic exchange services (DAAD), the South African National Research Foundation and the Agricultural Research Council for financing my studies.

I would like to thank my parents, my sisters, my brother, and my nephew for giving me strength and support, and that laughter after a stressful day. My special thanks to my elder sister, who always gave me extraordinary moral support and valuable advises. *Inkomu Sesi*. I would like to express my heartfelt appreciation to Patrice Ahmann for all the encouragements and letting me draw strength from his fountain.

To all my friends, thanks, you were reliable, supportive, and helpful. Wubi, you were my colleague and became my friend, thanks for assisting me with opinion and brainstorming. It was I, you, 'my' *Sorghum bicolor* babies and fun.

If the African ancestry-god theory is true, then thanks *va-kokwana*. I owe to enhance your farming dreams. Finally, to GOD almighty, I am bringing appreciation 'oil' with my alabaster box or my Tsonga women beaded calabash to your dwelling place.

Table of contents

	<i>Page number</i>
Declaration	i
Acknowledgement	ii
Table of contents	iii
List of the tables and figures	iv
List of abbreviation	v
1. Chapter 1	
Introduction and Literature review	1
1.1 Origin and domestication of sorghum	1
1.2 Classification	3
1.3 Introduction to molecular markers	3
1.4 Overview of mapping quantitative traits	4
1.5 Sorghum linkage groups	6
1.6 Sorghum improvement and quantitative traits locus	7
1.7 Non-structural carbohydrate analysis	11
1.7.1 Carbohydrate assimilation in sorghum	11
1.7.2 C ₄ plant photosynthesis	11
1.7.3 Quantifying methods	13
1.8 Structural carbohydrates in sorghum	13
1.8.1 Cellulose biosynthesis	14
1.8.2 Hemicellulose biosynthesis	15
1.8.3 Lignin Biosynthesis	16
1.9 The scope of bioenergy	17
1.10 Objectives	20
1.11 Reference	21
2. Chapter 2	
Genotype-environmental analysis of RIL population segregating for sugar-related traits in <i>Sorghum bicolor</i> L. Moench	29
2.1 Introduction	29
2.2 Material and methods	31
2.2.1 Plant material	31
2.2.2 Experimental design	31
2.2.3 Phenotypic evaluation	32
2.2.4 Glucose and sucrose analysis	33
2.2.5 Statistical analysis	34
2.3 Results and discussion	34
2.3.1 Phenotypic evaluation	32
2.3.2 Analysis of the variance and variance components	36
2.3.3 Estimating traits heritability	37
2.3.4 Correlation	37
2.5 Conclusion	38
2.6 Acknowledgement	38
2.7 References	42

3. Chapter 3	
Genetic and heritability estimates of fibre-related traits in Grain x Sweet sorghum recombinant inbred lines	44
3.1 Introduction	44
3.2 Material and methods	46
3.2.1 Field experiment	46
3.2.2 Data collection	46
3.2.3 Fibre content analysis	47
3.2.4 Statistical analysis	47
3.3 Results and discussion	48
3.3.1 Phenotypic evaluation	48
3.3.2 Analysis of the variance and its contribution effects	49
3.3.3 Heritability	50
3.3.4 Correlation	50
3.4 Conclusion	51
3.5 Acknowledgement	51
3.6 References	54
4. Chapter 4	
Genetic mapping of sugar-related traits in a RILs population of <i>Sorghum bicolor</i> L. Moench.	56
4.1 Introduction	56
4.2 Material and methods	58
4.2.1 Mapping population	58
4.2.2 Genotyping	58
4.2.2.1 DNA extraction	58
4.2.2.2 SSR marker amplification	58
4.2.2.3 AFLP marker amplification	59
4.2.2.4 Electrophoresis	59
4.2.3 Linkage map construction	59
4.2.4 QTL analysis	60
4.3 Results	60
4.3.1 Trait performance	60
4.3.2 QTL detection	61
4.3.3 Epistatic effect and QTL x environmental interaction	64
4.3.4 Pleiotropic effect and co-localization	66
4.4 Discussion	66
4.5 Conclusion	68
4.6 Acknowledgement	68
4.7 References	73
5. Chapter 5	
Identification of QTL for fibre-related traits in grain x sweet sorghum	78
5.1 Introduction	78
5.2 Material and methods	81
5.2.1 Phenotyping analysis	81
5.2.2 Genotyping	81
5.3 Results	82
5.3.1 Phenotyping results	82

5.3.2 QTL analysis	82
5.3.3 QTL x environment interaction and epistasis	86
5.3.4 Co-localization and pleiotropic effect	87
5.4 Discussion	88
5.5 Conclusion	89
5.6 Acknowledgement	90
5.7 References	95
6. Summary	98
7. Zusammenfassung	101

List of tables and figures

	Page number
Table 1.1 Classification of sorghum	3
Table 2.1 Biophysical information during field experiments in GG and RH	33
Table 2.2 Performance of parents and RILs for sugar-related traits across environments	39
Table 2.3 Mean Squares of sugar-related traits analysed by TYPE III mixed model	40
Table 2.4 Variance components and Heritability estimates for sugar related traits	40
Table 2.5 Pearson correlation coefficient of sugar-related traits	41
Table 3.1 Performance of parents and RILs for fibre-related traits across environments	52
Table 3.2 Mean squares of fibre-related traits analysed by TYPE III mixed model	52
Table 3.3 Variance components and Heritability estimates for fibre-related traits	53
Table 3.4 Pearson correlation estimates of the fibre-related traits	53
Table 4.1 Descriptive statistics and analysis of the variance	69
Table 4.2 Digenic epistatic loci associated with sugar-related traits	72
Table 5.1 Descriptive statistics and analysis of the variance	91
Table 5.2 Digenic epistatic loci associated with fibre-related traits	94
Supplementary Table 4.3 and Table 5.3 Details of additive QTLs detected associated with sugar-related traits and fibre-related traits, respectively in an excel file.	
Figure 4.1 Genetic linkage map and significant additive QTLs for sugar-related traits	70
Figure 5.1 Genetic linkage map and significant additive QTLs for fibre-related traits	92

List of abbreviations

ADF, acidic detergent fibre
ADL, acidic detergent lignin
AFLP, amplified fragment length polymorphism
CIM, composite interval mapping
cM, centiMorgan
DAA, days after anthesis
DNA, deoxyribonucleic acid
EDTA, ethylenediaminetetraacetic acid
EST, expressed sequence tag
et al., et alii, (and coworkers)
F, filial generation after a cross
GG, Gross Gerau
GLM, general linear model
LOD, logarithms of odds
NDF, neutral detergent fibre
NIRS, near infrared reflectance spectroscopy
PCR, polymerase chain reaction
QTL, quantitative trait locus
REML, restriction maximum likelihood
RCBD, randomized complete block design
RH, Rauschholzhausen
RILs, recombinant inbred lines
SSR, simple sequence repeat
TE, TrisEdta

Chapter 1

Introduction and Literature review

Sweet sorghum (*Sorghum bicolor* [L] Moench) is a single-stemmed cereal grass with a plant height of more than 2 m. In many cases, it is taller than grain sorghum. The stem is anchored by fibrous and spreading root systems, which further consists of primary, secondary and supporting roots. The roots may extend 1.5 m below the soil surface, branching in all directions. Growth and development of the stem, stem diameter, grain, and plant height depend on genotype, soil nutrition, growing conditions, temperature, and day length. Sorghum is a short day plant and cultivars differ in their sensitivity to photoperiod. Sensitivity to photoperiod is a genetically controlled character and can be bred or selected for.

The crop has become more attractive to breeders due to its capacity to provide renewable energy products, industrial commodities, and food and animal feed. Because sweet sorghum is a multipurpose crop, it has potential to aid development in semi-arid regions. Sweet sorghum stalk are rich in sugars, which can be primarily used for biofuel production. It can produce up to 7000 L of ethanol/ha (Seetharam, 2005). Other utilization can include processing it into syrup, grains for human consumption, stillage fibre and animal feed. Sweet sorghum has major advantages compared with its competitive crops such as maize, sugar beet, and sugar cane because it can produce high yield under extreme environmental conditions. It can tolerate drought, waterlogging, and saline-alkali soils. These adaptations and tolerance is due to a number of physiological and morphological characteristics. It produces many roots compared to other cereals, can reduce transpiration due to less leaf area and a waxy coating, remain dormant during drought and can resume growth when conditions are favourable, can compete well with weeds in an advanced stage, and has high net photosynthesis compared to other cereals.

1.1 Origin and domestication of sorghum

The genus *Sorghum* comprises a high genetic diversity (Abu Assar et al, 2005; Uptmoor et al., 2003), and therefore, there is potential for crop improvement. The wild distribution and high genetic diversity of sorghum in Africa indicate that this crop must have originated in this continent. Kimber (2000) suggested that sorghum could have been introduced to Asia, in particular India and china, by human migrations around 1045 B.C. To date, Ethiopia and surrounding countries are considered the geographical area of origin (Dhillon et al., 2007;

Kimber, 2000; Vavilov, 1951), and these countries are well recognised in world sorghum improvement programs.

De Wet et al. (1970) reported that the wild representatives of sorghum are strictly African in origin. The genus *Sorghum* is classed into several wild and cultivated species (table 1.1) and the *Sorghum bicolor* subspecies *bicolor* is the most used during production. *Sorghum bicolor* exhibits great phenotypic variability (Deu et al, 2006; Doggett, 1970). House (1985) described the five different major races: *bicolor*, *caudatum*, *durra*, *guinea*, and *kafir* and their 10 intermediate races according to Harlan and de Wet (1972). Basically, these races can be differentiated by panicle and spikelet morphology. Besides, Deu et al. (2006) and de Alencar Figueiredo et al. (2008) showed the possibility of differentiating race cluster and its geographical distribution by using molecular markers. For example, Deu et al. (2006) distinguished Guinea cluster accessions from the West African cluster, Southern African cluster and Asian cluster, which was not accurately possible before.

These races are agreed to have emerged by migration and disruptive selection (Doggett, 1965), where the *caudatum* types emerged from the centre of origin and later spread in all directions (House, 1985). The *Durra* types are predominant in South Asia and Northeastern Africa and there is no evidence on where this race originated (de Alencar Figueiredo et al., 2008). The *Guinea* emerged in western Africa and the *Kaffirs* in Southern Africa. The existence of *Guinea* types in South Africa is probably due to germplasm movement or to a more complex origin of diverse types classified under this race.

Genetic changes during domestication show the genius of early farmers, the first plant breeders without resources. Kimber (2000) reported that the differences between the cultivated races and their intermediates could be linked back to their specific environments and the nomadic people that first cultivated them. Humans caused a major shift in the morphological traits of wild plants by selecting genes of interest (Vaughan et al., 2007). Early domestication of sorghum was associated with changing the small-seeded, shattering open panicles to larger, non shattering seeds and more compact panicles (Dhillon et al., 2007). Recently, plant breeders focus on improving sorghum to serve as food, feed, fuel and fibre (Laopaiboon et al., 2007; Tarpley and Vietor, 2007; Vermerris et al., 2007). This involves improving characters such as yield performance and stability, resistance to pests and pathogens, and grain and stem qualities.

1.2 Classification

Sorghum is difficult to classify because of its wide diversity. Over time, sorghum has been described and re-described by numerous taxonomists. Dhillon et al. (2007) summarised the time-line displaying the changes in *Sorghum* nomenclature over time.

Table 1.1. Classification of sorghum

Family	<i>Poaceae</i>
Tribe	<i>Andropogoneae</i>
Subtribe	<i>Sorghinae</i>
Genus	<i>Sorghum</i> Moench
Subgenera	<i>Sorghum</i> <i>Chaetosorghum</i> <i>Heterosorghum</i> <i>Parasorghum</i> <i>Stiposorghum</i>
Species of subgenera sorghum	<i>Sorghum propinquum</i> <i>Sorghum halepense</i> <i>Sorghum bicolor</i>
Subspecies of sp. <i>S. bicolor</i>	<i>Sorghum bicolor bicolor</i> <i>Sorghum bicolor drumondii</i> <i>Sorghum bicolor verticiliflorum</i>
Races of subsp. <i>S. bicolor. bicolor</i>	Bicolor Guinea Durra Kafir Caudatum
Crosses of races	10 intermediate races

Information compiled from House (1985); Kimber, (2000); de Wet and Harlan, (1972)

1.3 Introduction to molecular markers

Genetic markers represent genetic differences and act as signs for target genes between individual plant, lines, varieties or species (Collard et al., 2005). Today, DNA markers are the most used type of markers in comparison with morphological and biochemical types of markers. These DNA markers reveal abundance and neutral variation at the DNA sequence

level (Collard et al., 2005, Jones et al., 1997). As the rule of thumb, the following factors should be considered before selecting markers to be used: i) reliability and reproducibility, ii) easy technical procedure, iii) cost, iv) level of polymorphism, and v) required DNA quality and quantity. The DNA-based markers have been used for the genetic analysis and manipulation of important agronomic traits aiding conventional plant breeding (Ejeta et al., 2000). The use of molecular markers to analyse the inheritance of quantitative traits has a great potential for improving genetic gain and efficiency of superior genotypes selection through studying genetic and phenotypic basis of complex traits. Progress can be achieved through gathering precise information about the number of loci involved in trait expression, their location along the chromosome, and the relative contribution on trait expression of each locus (Maradiaga, 2003).

1.4 Overview of mapping quantitative traits

Traits that exhibit continuous variation are termed quantitative traits (House, 1985). These traits are largely influenced by the environment (House, 1985), may lie outside the mean range of the parents (Jones et al., 1997), are controlled by many genes or polygenes (Gupta, 2002), where most of these genes have minor or major effects. The loci can be detected by molecular markers, which should be segregating in a mendelian manner. Basically, the quantitative trait loci (QTL) analysis is based on detecting an association between phenotype and marker of a specific genotype. Tanksley (1993) reported that genes or QTL controlling traits can be detected through genetic linkage analysis, which is based on the principle of genetic recombination during meiosis. A linkage map indicates the position and relative distance between markers along chromosomes. Genes or markers that are close together or tightly-linked will be transmitted together from parent to progeny more frequently than genes or markers that are located further apart. Creating a linkage map requires three steps: production of mapping population, identification of polymorphism, and linkage analysis of markers. Construction of a well-saturated marker linkage map is the most fundamental step required for a detailed genetic study and marker-assisted selection (MAS) approach in any crop because the main objective of QTL mapping in plants is to utilize the information generated in MAS breeding programs.

Various populations may be developed for mapping purposes. The type of mapping population chosen depends on the objectives, infrastructure available, funds, labour, and time required. The parents selected to develop the mapping population should be diverse for one or

more traits of interest (Collard et al., 2005). However, parents should not be too diverse to avoid reduced vigour or even sterility of the progenies. The large population size is desired for increasing the accuracy of QTL mapping. In self-pollinating crops, mapping populations originate from crossing parents that are both highly homozygous. Selfing F₁ individual generates an F₂ population, while a backcross (BC) population is derived by crossing back the F₁ to the recurrent parent. Both populations are recombinant along each pair of homologous chromosome and are quick to develop. Moreover, an F₂ population has all possible combinations of parental alleles.

Double haploid (DH) populations are mapping populations that represent perfectly homozygous lines, are faster to develop than recombinant inbred lines (RILs), and present the same amount of recombination information as a backcross population. However, besides being labour intensive, this type of population may have somaclonal variation and aberrant segregation ratios due to tissue culture response. Near-isogenic lines (NILs) is another type of mapping population that can be developed by either backcrossing the progeny to one of the parents. This will results in a population in which the locus of interest differs among individuals with the same genetic background. NILs are time consuming, serve only to map one trait, and lack a segregating population to assess linkage relationships.

The most used type of mapping population is representing recombinant inbred lines (RILs), which are produced by inbreeding from individual F₂ plants. During the RIL population development, selection of the desired traits starts at the F₂ generation until homogeneity, i.e at least F₆. Then, the population consists of a series of largely homozygous lines, each containing a unique combination of chromosomal segments from the original parents. Mapping using RILs increases accuracy with high-resolution maps due to multiple meiotic events it undergoes in comparison with an F₂ population. Although the time required to develop this type of population is long, the lines produced are eternal (fixed), can be replicated across seasons and years, and can be exchanged between laboratories.

Collard et al. (2005) reported three statistical methods for detecting QTLs: simple-marker analysis, simple interval mapping, and composite interval mapping. Simple marker analysis considers the association between the trait and one marker locus at a time. The method includes t-test, analysis of the variance and linear regression, and does not require a complete linkage map. Simple interval mapping (Lander and Botstein, 1989) makes use of linkage

maps, intervals formed by the pairs of adjacent markers are taken as the unit of analysis and tested for the presence of single QTL by using information from flanking markers. The most recent and popular method is composite interval mapping (Zeng, 1994) which combines interval mapping with linear regression and includes additional genetic markers in the statistical model in addition to an adjacent pair of linked markers for interval mapping. Composite interval mapping is the most precise and effective method for mapping QTLs so far.

Different factors such as environmental effects, population size, experimental error and genetic properties of QTLs that control the trait, have been reported to influence the detection of QTLs in a segregating population (Asins, 2002; Tanksley, 1993). Genetic properties of QTLs controlling traits define the magnitude of the effect of individual QTLs, where only QTLs with large phenotypic effects will be detected and QTLs that are closely linked will usually be detected as a single QTL in population with less than 500 genotypes (Collard et al., 2005). Environmental effects can highly influence the expression of quantitative traits. Therefore, experiments should be replicated across years and locations to increase the accuracy. It has been reported that larger population size, increases the mapping precision (Haley and Anderson, 1997), increases the detection of QTLs with small effects (Tanksley, 1993), increases gains in statistical power, and reliably estimates gene effects and confidence intervals of the location of the QTLs (Beavis, 1998). The main sources of experimental error occurs due to mistakes in scoring genotypic data or by phenotypic evaluation. Errors and missing data can affect the distance between markers within the linkage maps and the accuracy of QTL mapping can be reduced by unreplicated experiments and inaccuracy in phenotyping.

1.5 Sorghum linkage groups

Genetic maps produced from molecular markers provide understanding of evolution, genetic diversity and phylogeny relevant to germplasm selection and improvement, and the co-dominant nature of markers allows thorough coverage of entire species' genome and can reveal the locations of QTLs influencing all measurable traits (Paterson, 1995). Several genetic maps have been developed in sorghum with Simple Sequence Repeats (SSR) markers, Amplified Fragment Length Polymorphism (AFLP), and Restriction Fragment Length Polymorphism (RFLP) (Bhatramakki et al., 2000; Boivin et al, 1999; Kong et al., 2000). The first genetic map of sorghum was created by Hulbert et al. (1990) using 55 F₂ plants and DNA

probes that had been used in maize, and obtained 8 linkage groups with a total length 283 cM. Ragab et al. (1994) constructed a linkage map with 15 linkage groups, with a length of 633 cM and average distance of 8.9 cM. The maps created by Pereira et al. (1994), Chittenden et al. (1994) and Peng et al. (1999) confirmed that sorghum consists of 10 linkage groups corresponding to its chromosome number. The information generated by a linkage map is required for plant improvement, linking species information from plant breeders and plant biologists about the traits of interest, positional cloning of genes, and broadening gene pools.

According to Kong et al. (2000) a high level of polymorphism is important for genetic markers to be used in fingerprinting and marker-assisted selection. The first genetic maps were constructed using RFLP markers because of their stable and co-dominant nature. However, the technique requires large amount of DNA, is time consuming and can detect low polymorphism. RAPD markers are scarcely used to construct linkage maps because they are less stable and not reproducible (Hausmann et al., 2002). The utilization of SSR and AFLP have gained much popularity in recent years because they are reproducible, stable, and highly polymorphic (Zhi-Ben et al., 2006). Menz et al. (2002) constructed a map with 2,926 markers, 10 linkage groups, and 1713 cM with an average distance of 0.5 cM between markers using AFLP, SSR and RFLP markers.

The number of linkage groups in sorghum totals to 10 and can be visualized by FISH markers (Kim, 2005), where the designation of the chromosomes can be LG for linkage groups (Zhi-Ben, 2006) or SB for *Sorghum bicolor* (Kim, 2005) followed by either A, B,C, D or 1, 2, 3, 4. The arrangements of the 10 LGs or SBs, depend on chromosome length, ranging from 5.11 to 2.94 μm with an estimated DNA content ranging from 119.3 to 68.6 Mbp (Kim et al., 2005). A high density map, integrated with physical and cytological information, is essential for mapping and cloning target genes, sequencing genomes and assembling the contigs.

1.6 Sorghum improvement and quantitative trait loci

The success in improving crop characters over a number of generations depends on the mode of inheritance, breeding procedures, intensity of selection pressure and environmental factors (House, 1985). Sorghum is a self-pollinating crop with 2-20 % outcrossing. The genus *Sorghum* comprises 48 species, which constitute the primary gene pool. The world largest sorghum gene banks are found in India and the United States, where Africa contributes 60 %

of the total germplasm collected from 40-42 countries and Asia 20-40 % from 24-28 countries (Rai et al. 1999).

In Africa and Asia, sorghum remains a traditional food crop for resource-poor farmers. The production information has been going-on for years without documentation. The eco-geographical adaptation together with intensified selection by Africans has led to a high diversity in sorghum. However, poor productivity for grain yield of unimproved landraces is relative to their harvest index.

Sorghum breeders worldwide are aiming at improving various characteristics of the crop that can result in commercialization. Although the main trait of interest in sorghum for many years is grain yield, recently interest has also shifted from grain to stalk to manipulate biofuel-related traits: fibres, sugars, juice, and biomass. Grain yield and biofuel-related traits cannot be manipulated directly, because they are quantitatively inherited. Therefore to successfully improve these traits, the crop should be dissected into smaller morphological, physiological, and genetical components and then correlate the information acquired from these components.

Plant characteristics to be improved include high yield, appropriate height, environmental adaptation, non-lodging, good tillering, and appropriate time to maturity. Day length sensitivity in sorghum is an important criterion for environmental adaptation because the photoperiod insensitive varieties can be grown in all regions of the world. Understanding genetics behind photoperiod sensitivity will facilitate the improvement of landraces to high yielding varieties. In an effort to improve this trait, Chantreau et al. (2004) detected QTLs for photoperiod in linkage group 3, 6, and 8, and Lin et al. (1995) detected one in linkage group 2. Quinby (1973) suggested that flowering is controlled by four genetic loci, *Ma1*, *Ma2*, *Ma3*, and *Ma4*, where Aydin et al. (1997) added *Ma5* and *Ma6* genes.

Genomic regions for plant height and tillering have been identified in intra and interspecific crosses of sorghum (Hart et al., 2001; Lin et al., 1995). Cloning and utilization of these genes can have a positive impact on grain yield, grain harvesting techniques, and utilization of the crop for bio ethanol production.

African landraces are grown where weather is characterised by a greater degree of uncertainty e.g. rainfall distribution, which results in challenges to secure the level and quality of the

harvest. Due to its wide tolerance to harsh environments, sorghum has been widely studied to identify genomic regions associated with droughts. However, Hausmann et al. (2002) reported that the genetic and physiological mechanisms that condition the expression of drought tolerance are poorly understood since the traits are controlled by many genes, depend on the timing, and severity of moisture stress, and involve a number of physiological and biochemical pathways. Selection for drought tolerance is difficult because of inconsistency in testing environments and interactions between stages of plant growth and environments (Kebede et al., 2001). Significant progress of understanding the drought tolerance trait can be made by breaking the trait into specific phenological stages during pre-flowering and post flowering. Kebede et al. (2001) explained pre-flowering drought-stress response in sorghum as when plants are under significant moisture stress prior to flowering from panicle initiation until flowering while post-flowering drought stress causes premature leaf senescence leading to stalk lodging, stalk rot disease, and significant yield loss. The 'stay green' trait has been recognised as a major mechanism of post-flowering drought stress tolerance (Rosenow et al., 1996). Sanchez et al. (2002) referred stay green as drought resistance mechanism that gives sorghum resistance to premature senescence under soil moisture stress during grain filling. It results in greater functional photosynthetic leaf area during grain filling and even after physiological maturity. Rai et al. (1999) introduced stay green as a highly heritable trait. Many genomic regions involved in stay green and related agronomic traits have since been identified: Tunistra et al. (1996; 1998) mapped six regions associated with pre-flowering, two QTLs for stay green and yield, and Kebede et al. (2001) mapped nine QTLs for stay green, three QTLs for lodging tolerance, four QTLs pre-flowering drought resistance, two QTLs for flowering time, and two QTLs for plant height.

The major production constraints in sorghum are insect pest, diseases and the parasitic weed *Striga*. Rai et al. (1999) reported that biotic constraints have received relatively low priority in sorghum breeding programs in the semi-arid tropical regions of Africa and Asia, even though the lower yields in Asia and Africa are associated with pest damage. Common seed and seedling rot diseases in sorghum are caused by soil-and seed borne *Aspergillus*, *Fusarium*, *Pythium*, *Rhizoctonia* and *Rhizopus spp.* Until recently, there are no varieties with good resistant reported and therefore, fungicides are widely used to treat the seeds. Nearly 150 insects' species have been reported as pests on sorghum (Jotwani et al., 1980; Reddy and Davis, 1979). Such sorghum pests in tropical Africa are shoot fly (*Atherigona soccata*), stemborers (*Busseola fusca*, *Chilo partellus*, and *Sesamia calamistis*). Adoption of chemical

control method measures is not economically feasible for most of the sorghum-growing farmers. Therefore, utilization of host-plant resistance is the most realistic approach to reduce losses caused by sorghum insect-pests. Genetic manipulation of characters such as shoot fly resistance by conventional breeding methods faces many difficulties such as crossing barriers and transfer of undesired traits. Dhillon et al. (2006) reported that there is a high genetic diversity in shoot fly-susceptible parents and their hybrids in comparison to shoot fly resistant lines. This conclusion was based on high polymorphic information content values after screening 12 cytoplasmic male sterile (CMS) lines and maintainers, 12 restorer lines, and 144 F1 hybrids. Based on genetic and morphological distances of hybrids to CMS or restorer line, they suggested that CMS influences the expression of resistance to sorghum shoot fly.

Sorghum grains are used as staple food in several regions of Africa, China, and India, particularly in the semi-arid regions. Sorghum grain types, processing methods, and food habits vary considerably with the geographical area. Despite low yields, local cultivars are generally preferred by consumers due to their characteristics, and other food quality attributes. The most common products from the whole grain flour are porridges. In earlier work, I have found a correlation between milling quality characteristics and viscosity parameters with endosperm texture (Shiringani, 2005). Physical characteristics of the grain such as colour, size, pericarp thickness, endosperm texture, presence or absence of pigmented testa plays the major role in acceptability of the products. However, after the food has been tasted they become secondary to cryptic and palatability characteristics. Breeding for quality has been carried through a modified pedigree procedure and involves parents with good food quality traits, mould resistance, and a broad-spectrum resistance (House, 1985). Recently, research programs are focusing on food quality improvement through molecular markers. Rami et al. (1998) researched on QTLs of quality characteristics such as dehulling yield, kernel flouriness, friability and hardness, amylose content, protein content, lipid content, and mould during germination and after harvest. They detected a chromosomal segment on LG6 that play a major role in grain quality, and located four closely-linked QTLs for flouriness, dehulling yield, amylose content, and mould resistance and another four QTLs for kernel friability, kernel hardness and amylose content in RIL379 and RIL249 population respectively. Klein (2001) reported 5 QTLs for grain mould with significant effects on LG4, LG5, LG6, LG7, and LG9. The genetics of grain mould resistance is reported to be controlled by the minimum of four genes (Klein, 2001) while the B2/b2 gene known to control high-tannin testa layer in

sorghum seem related to QTL for mould sensitivity and kernel vitreousness (Rami et al., 1998).

1.7 Non-structural carbohydrates analysis in Plants

Plants carbohydrates can be grouped into two parts, structural and non-structural (Anglani, 1998). Starch, sugars, and fructans are examples of non-structural carbohydrates while cellulose, hemicellulose and pectin substances are structural carbohydrates. Carbohydrates are the main components of food. They can be classified according to the number of monomers that they contain as monosaccharides, oligosaccharides, and polysaccharides. The main aim of determining the type and concentration of carbohydrates includes standard identities, nutritional labelling, economic, and food quality and processing.

The sugar content of sorghum juice determines the commercial value of the crop. Sugar components play an important role in flavour characteristics and consumer acceptability in food products. Recently, renewable carbohydrate in sweet sorghum has become of interest for biological transformation into ethanol for use as fuel or fuel additive (Laopaiboon et al., 2007). The initial sugar concentrations correlate with the concentration of ethanol produce. The total amount of sugar is totally consumed during fermentation. The sugar accumulation is mainly influenced by varieties (pest resistance, period to maturity), planting density, fertilization, time of harvest, culm section and temperature (McBee and Miller, 1982).

1.7.1 Carbohydrate assimilation in sorghum

Glucose ($C_6H_{12}O_6$) is the most important monosaccharide in carbohydrates biosynthesis. It is a product of photosynthesis and the initiator of cellular respiration in eukaryotes and prokaryotes. D-glucose and L-glucose are two isomers of the aldohexose glucose, where D-glucose is the active form and L-glucose cannot be used by the cell.

1.7.2 C_4 plants photosynthesis

Plants are capable of synthesising a large number of different forms of sugars, as monosaccharides or disaccharides. Monosaccharides are simple carbohydrate with the general formula of $(CH_2O)_n$, for example glyceraldehydes-triose ($n=3$), erythrose-tetrose ($n=4$), ribose-pentose ($n=5$), or glucose-hexose ($n=6$). Disaccharides such as sucrose and lactose are compounds of two monosaccharides. Long branched or unbranched chains of

monosaccharides form polysaccharides such as starch, glycogen and cellulose. The production of sucrose is restricted to plants and cyanobacteria.

Most of the plants exhibit C_3 photosynthesis. However, major crop species such as maize, sugarcane, sorghum, millets, and many tropical grasses perform a C_4 pathway. The C_4 plants are generally characterised by high photosynthetic efficiency (particularly under elevated temperature) and growth, and they are the most productive crops. A C_4 pathway occurs in two photosynthetic cell types, mesophyll cells and bundle sheath cells. Carbon dioxide enters the mesophyll cells and is carboxylated by phosphoenolpyruvate (PEP) carboxylase into the organic acid oxaloacetate (OAA). The OAA is converted to malate or aspartate, which diffuses into the bundle sheath cells through plasmodesmata.

The C_4 acid is decarboxylated and the resulting CO_2 is re-fixed by Ribulose biphosphate carboxylase-oxygenase (Rubisco) in the bundle sheath cells. Decarboxylation for the C_4 acids can be done through chloroplastic NADP-malic enzyme that decarboxylates malate into pyruvate, mitochondrial NAD-malic enzymes and cytosolic phosphoenolpyruvate carboxykinase. The carboxylation of ribulose 1.5 bisphosphate (RuBp) produces 2 molecules of 3-phosphoglycerate (3-PGA), which is reduced to form triose glyceraldehyde-3-phosphate. The triose sugars (C_3) produced during photosynthesis are either used for further biosynthesis, i.e. starch, lipids or amino acids in the chloroplasts or sucrose and protein synthesis in the cytosol.

TrioseP is exported from the chloroplast to the cytosol in exchange of orthophosphate (P_i) and converted to fructose-1.6-bisphosphate by adolase, which is further hydrolysed to fructose-6-phosphate by the cytosolic fructose-1.6-bisphosphatase. Fructose-6-phosphate is converted to glucose-6-phosphate, glucose-1-phosphate, and Uridine diphosphate-glucose (UDP-GLc). UDP-GLc and ADP-GLc are the two major nucleotide sugars. Nucleotide sugars are monosaccharides that have a nucleotide bound to the carbohydrate chain through the linkage that can be used to form bonds to other sugars or proteins. ADP-GLc is formed by ADP-GLc pyrophosphorylase (*AGPase*) and is the sole precursor for starch formation. The sucrose-phosphate synthase (*SPS*) catalyses the synthesis of sucrose-6-phosphate from Uridine diphosphate-glucose, while sucrose-phosphate-phosphatase (*SPP*) hydrolyses sucrose-6-phosphate to produce sucrose and inorganic phosphate (P_i). Hawker and Smith (1984) suggested that *SPP* is unlikely to make a significant contribution to the control of

sucrose biosynthesis because it is present in larger quantities in plants than *SPS*. However, Lunn et al. (2000) found that the maximum potential activities of *SPP* and *SPS* are similar to each other in rice, wheat, and maize leaves and in developing wheat seeds using improved extraction and assay methods. Echeverria et al. (1997) have provided further evidence that *SPP* has a bigger role in controlling sucrose biosynthesis than *SPS*.

Unfortunately, there is no much information about *SPS*, *SPP*, and *Susy* in Sorghum. However, Castleden et al. (2004) outlined 62 sorghum *SPS* ESTs, within Family A, B, C, and D. These ESTs are expressed in seedlings, embryos, leaves, panicles, ovaries and pollen.

1.7.3 Quantifying methods

Various techniques are used to quantify specific sugars such as chromatographic methods, enzymatic analysis and electrochemical or spectrophotometer methods (Liu et al., 2006). Chromatographical methods are tedious, time consuming, require extensive sample preparation and expensive equipment (Liu et al., 2006; Velterop and Vos, 2001). Gomez et al. (2007) reported that analysis of sugars by High-performance Liquid Chromatography (HPLC) is characterised by non-specific detection of different sugars and difficulty in sample purification. Enzymatic analysis of sugar is based on the ability of enzymes to catalyze a specific reaction. These analyses are rapid, sensitive to low concentrations, and can measure different and independent end-points of sugars (McIntosh et al., 1998), while providing confidence for the data collected. Meanwhile, Gomez et al. (2007) reported that enzymatic analyses are costly because the enzymes involved are expensive and further suggested that the cost can be reduced by using a 96-well microplates.

1.8 Structural carbohydrates in sorghum

The plant cell wall is a dynamic, highly organised extracytoplasmic matrix consisting of various polysaccharides, structural proteins, and aromatic substances, which are constantly remodelled and restructured during growth and development. According to O'Neill and York (2003), the complex matrices define the shape and size of individual cells within the plant body and ultimately plant morphology. Other divergent and vital functions provided by the cell wall include tensile strength, defending the plant against pathogens and mediating the communication with symbionts and neighbouring cells.

Cosgrove in 2005 reported that plant cell grow by expanding their cell walls through a process of controlled polymer creep (slow, time-dependent, irreversible extension, in which the microfibrils and associated matrix polysaccharides slowly slide within the wall, therefore increasing its surface area). The growth is accomplished by the synthesis and integration of new polysaccharides into the wall and the expansion of the polysaccharides network. The primary cell wall is synthesized during expansion growth (Carpita, 1996) and the secondary cell wall begins in the primary cell wall when the cell is ceasing to expand (Dhugga, 2007). The primary cell walls contains mainly cellulose and hemicelluloses, pectins as well as many enzymes and structural proteins, whereas the secondary cell walls is rich in cellulose and also contain lignin and little protein or pectin. The main structural carbohydrates by content in the cell wall of sorghum are cellulose, hemicellulose and lignin (Anglani, 1998).

1.8.1 Cellulose biosynthesis

Cellulose is the most abundant polymer on earth and it the core of the plant cell wall (Lerouxel et al., 2006; Taylor, 2008). In primary cell wall, cellulose is important for load-bearing network and determining the orientation of cell expansion, while in secondary cell wall it provides mechanical properties that allow the cell to stand upright and facilitate the xylem vessels to function properly (Taylor, 2008). Cellulose is more abundant in secondary cell wall having 40% dry mass compared to in primary cell wall that has 15-30% dry mass (Sticklen, 2008).

Recently, high throughput genomic technologies made a significant contribution in understanding the molecular details involved in cellulose biosynthesis (Dhugga, 2007; Lerouxel et al., 2006). Genome sequencing revealed that all higher plants have multiple cellulose synthase catalytic subunit (CESA) proteins encoded by the CESA genes, which share a conserved structure. Plant genomes contain multiple CESA genes as part of the gene family, which may assemble to form a functional complex. Cellulose is made by a complex enzyme system (Dhugga, 2007; Hazen et al., 2002; Lerouxel et al, 2006) that is integrated into the plasma membrane after being assembled in the Golgi bodies.

Membrane-bound cellulose synthase enzyme complexes are visible as hexameric rosettes of approximately 25-30 nm in diameter (Kimura and Itoh, 1996; Lerouxel et al, 2006). These rosettes can move in the plasma membrane guided by cortical microtubules. Each rosette may synthesize six β - 1,4 glucan chains that coalesce as cellulose microfibrils in the cell wall. The

microfibrils are deposited in the cell wall in a controlled orientation that is the same as that of microtubules in the cytosol. This similarity in orientation suggests that microtubules control the orientation of microfibril deposition.

Besides the known CESA genes, other genes have been reported to be involved in cellulose biosynthesis. For example, COBRA genes encode protein hypothesised to orientate cellulose distribution, KORRIGAN encodes a β -1.4 glucanase, and PEANUT encodes the enzymes that are involved in glycosylphosphatidylinositol membrane anchors. However, few is known or validated about the biochemical involvement of these genes (Dhugga, 2007; Lerouxel et al., 2006; Taylor, 2008).

1.8.2 Hemicellulose biosynthesis

Unlike cellulose that is a branched polymer and contains anhydrous glucose, hemicellulose is an unbranched polymer containing glucose, xylose, mannose, arabinose, and galactose. These sugar components resulted in classification of hemicelluloses into xylogans that contains β -1,4-glucan backbone, mannans that contain β -1.4-linked mannose (glucose and mannose), glucuronoarabinoxylans that contain β -1.4-linked xylan backbone, and mixed linkage glucans that contain linkages of both β -1.4-glucan and β -1.3-glucan. Cosgrove (2005) reported that xyloglucan and arabinoxylan are two of the most abundant hemicelluloses.

Structural similarities between β -1.4-glucan chains of cellulose and the backbones of various β -linked hemicellulosic polysaccharides suggest that cellulose synthase like (CSL) genes might be involved in the biosynthesis of hemicelluloses (Lerouxel et al., 2006). The CSLs are a family of genes that have a sequence similar to the CESA genes and appear to be present in all plant genomes. Hemicelluloses are synthesized in the Golgi bodies, package into vesicles and then exported to the cell wall through exocytosis.

Unlike cellulose, hemicellulose of dicots and graminaceous monocots is different (Hazen et al., 2002). Dicots contain large amount of pectin and xyloglucan while cereals contain large amounts glucuronoarabinoxylan and some tissues have high mixed-linked glucan. These structural differences suggest that enzymes involved in hemicellulose biosynthesis CSL genes are divided into eight families: CSLa through CSLh (Vogel, 2008). The CSLf and CLSh families are unique to the grasses whereas CSLb and CSLg are unique to dicots. The

remaining families are in both dicots and monocots and might be responsible to synthesise glycans or lineage-specific glycans that occur in both dicots and monocots.

1.8.3 Lignin Biosynthesis

Lignin is an amorphous polymer consisting of three aromatic alcohols (monolignols): *p*-coumaryl, coniferyl and sinapyl alcohols (Buranov and Mazza, 2008). The respective aromatic constituents of these alcohols in the polymer are *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) moieties (a part or half a molecule), which are transported from the cytosol to the apoplast. During lignification process, these monolignols produce a complex three-dimensional amorphous lignin polymer via linkages which lack the regular and ordered units such as those in cellulose or protein. Dixon et al. (2001) reported that the mechanisms that determine the relative proportions of these linkage types in a particular lignin polymer are currently not known.

Lignin content in grass cell wall ranges from 6-12 % of the dry weight (Jung et al., 1997), and together with lignin composition varies among major groups of higher plants and among species (Whetten et al., 1998). For example, wood lignin mainly contains guaiacyl and syringyl units, whereas those of herbaceous plants contain all three units in significant amounts with different ratios (Billa et al., 1997). The ratio of syringyl to guaiacyl in grasses increases as the plant ages (Grabber et al, 2004) and H occurs in grasses in a very low amount (Grabber et al, 2004). Lignin does not exist in plant tissue as an independent polymer but it is bonded with other polymers such as cellulose and hemicelluloses forming lignin-carbohydrates complexes (LCC). The LCC from herbaceous crops are structurally different from those in woods and contain ferulic bridges between lignin and carbohydrates (arabinoxylans) via ester-linked ferulic acids. Thus, these LCC are often referred to as lignin/phenolics-carbohydrates complexes. Ferulic acid is attached to lignin with ether bonds and to carbohydrates with ester bond.

Unlike in woody plants, lignin biosynthesis is not well studied in grasses; the structure and properties of lignin, and the precise interrelationships with other cell wall components are not well understood (Buranov and Mazza, 2008). Recently however, efforts are being made to understand lignin biosynthesis in grasses because of the growing interest in bioethanol production from renewable biomass. Among grasses, maize received the most attention in genetic and genomic studies related to cell wall lignification and degradability (Barrierè et al.,

2003), but with the complete sequence of rice and sorghum genome a better understanding of pathways and genes involved will be accumulated.

1.9 The scope of bioenergy

Besides that the adoption of bioenergy will mitigate the effects of elevated greenhouses gases in the environment, it will also diversify the energy portfolio, enhance energy security (Vermerris et al., 2007), reduce dependence on the oil producing countries which are known for their ‘political instability’, and may stimulate third world economic development. In a short run, bioethanol and biodiesel are considered the most promising transportation fuels to replace diesel, gasoline or petroleum. Biodiesel is produced from plant-based oils or fats that is produced from oil-containing seeds of either soybean [*Glycine max* (L) Merr.], sunflower [*Helianthus annus* L.] or canola [*Brassica napus* L.] and bioethanol can be produced by fermenting sugars mostly glucose and sucrose. These sugars can be obtained directly from juice of sweet sorghum or sugarcane, through conversion of starch of sorghum or maize grain, or through conversion of cellulose and hemicellulose from stovers of maize and sorghum. Genetic enhancement of plants for the production of renewable energy aims at altering plant biochemical composition to improve the processing characteristics, modifying plant stature and architecture, improving resistance to biotic and abiotic challenges and making potential overall juice sugar, grain starch and biomass yield a target for crop improvement. The choice of the method will depend in a part of the trait of interest or the biochemical process that is being targeted as well as the plant species being used.

In sweet sorghum, genetic improvement of sugars has not been intensively studied compared to that of sugarcane. In sugarcane, breeders selected successfully high sugar varieties by using refractometer-based °Brix for many years, the strategy that the sweet sorghum breeders have recently adopted. With the popular growing of molecular and genomics technology, plant breeders and biologists are aiming at increasing sugar content at the molecular and genetic level through understanding the differences in the regulation of genes related directly and indirectly to sucrose accumulation. Papini-Terzi et al. (2009) reported that genes for sucrose content and pathways are overlapping with drought and other abiotic stress response genes, ABA pathways, cell wall biosynthesis and plant growth and development genes, thus making sucrose yield a complex trait. This research was conducted in two populations of sugarcane and each having more than 500 individuals. Decreasing expression of cell wall metabolism-associated genes is reported to increase the expression of sucrose metabolism related genes

and increase sucrose content (Papini-Terzi et al., 2009). It seems like many of the sucrose-associated genes that are regulated during development are associated with drought response (Srinivas et al., 2008) or are modulated by sugars (Stitt, 1996) or ABA.

Wang et al. (2008) reported that ethanol yield from sorghum grain is comparable to that from maize grain. However, less research on starch fermentation to ethanol has been conducted in sorghum than in maize. Zhan et al. (2006) studied the effect of genotypes and location on ethanol production, several researchers (Rooney and Pflugfelder, 1986; Elkhailifa et al., 2004; Ezeogu et al., 2005; Wang et al., 2008) studied factors affecting starch digestibility, Beta and Corke (2001) and Beta et al. (2001) studied the starch properties and its isolation. Basically, the knowledge generated from starch-related research laid the foundation to understand the requirements of converting starch to ethanol. Recent work by Wang et al. (2008) focused on the effect of starch properties to ethanol yield. In this experiment, seventy genotypes and elite hybrids with extreme differences in starch properties were used. The starch content in genotypes ranged from 64-74% of grain dry weight, and they observed that genotypes with waxy and heterowaxy grain produce high ethanol yields than non-waxy grain at the same starch concentration. Starch can react with lipids or protein and form amylose-lipid complex or starch-protein complex, which cannot be converted into fermentation sugar, thus reducing ethanol yield. In this case, small starch granules may be embedded in the protein matrix during grain development and does not gelatinize during cooking, and some starch may remain embedded in the protein matrix due to protein denaturation during cooking. The embedded starch is inaccessible to hydrolytic enzyme degradation and cannot be converted into sugars for fermentation (Wang et al., 2008). Some researchers reported that fine grinding (0.5 mm sieve) compared to coarse grinding (5 mm sieve) of grains in sorghum and maize can improve gelatinization, enzymatic hydrolysis and fermentation, thus results in high ethanol yield (Naidu et al., 2007; Wang et al., 2008).

Lignocellulosic biomass such as stover from maize and sorghum, straw from wheat, bagasse from sugarcane, switchgrass, elephant grass, and trees such as poplar and eucalyptus offers an abundant and inexpensive source of fermentable sugars. Biomass quality is heavily influenced by cell wall composition and structure and can be determined by content and composition of lignin, cellulose and hemicellulose and how they are cross-linked, whereas biomass yield is determined by agronomic traits such as plant height and diameter, diseases and pest resistance and lodging susceptibility and can be manipulated through crop management practices.

Sugars obtained from cell wall polysaccharides (cellulose and hemicellulose), can be hydrolysed with hemicellulose or cellulose degrading enzymes; a process known as saccharification (Lin and Tanaka, 2006). Monosaccharides obtained are fed to microorganisms in fermenters similar to the systems used in with starch- or sugarcane-derived sugars. Although the energy balance of the stover-to-ethanol process has been debated, technical and economic analysis showed that the production of ethanol from lignocellulosics could result in a net gain. To make this technology economically competitive, cost-effective and efficient means of transportation and storage of biomass yield, and processing strategy should be developed. Besides, it is mandatory to understand the physical and chemical properties, biosynthesis and molecular genomics of plant cell walls and also how to manipulate them in the candidate crops.

The complex heterogeneous structure of the plant cell wall can influence the ability of cellulolytic enzyme complexes to digest plant biomasses to fermentable sugars. Lignin in particular appears to be very significant: It prohibits cellulose fibre swelling, which reduces surface area access to the enzyme, and cellulases irreversibly adsorb to lignin, thus preventing their action on the cellulose. This necessitates the use of more enzymes to saturate these non-productive sites adsorption sites on the surface of biomass and lead to high enzyme cost.

Mutation in the lignin biosynthesis pathway has been identified, in both maize (Vignols et al., 1995), sorghum (Bout and Vermerris, 2003), pearl millet (Cherney et al., 1988) and sudangrass (Fritz et al., 1981), as brown midrib mutation, which results in the formation of brown vascular tissue in the leaves and stems. In maize, this brown midrib mutation was spontaneous but in sorghum was generated via chemical mutagenesis with diethyl sulphate (Porter et al., 1978). Lignin content can be decreased and digestibility improved by genetic control of the lignification process through manipulation of the *bmr* trait. In 1975, Porter et al. obtained 19 *bmr* mutants sorghum lines by soaking sorghum seeds with diethyl sulphate and in 1988, Fritz et al. selected 3 mutants line and developed these lines using *bmr* sorghum genes: *bmr-6*, *bmr-12* and *bmr-18*. Bittinger et al. (1981) reported that *bmr-12* and *bmr-18* genes are allelic and Gupta (1995) indicated that *bmr-6*, *bmr-12*, and *bmr-18* are located at two independent loci. Recent research identified two separate enzymes exhibit reduced activity as the results of *bmr* mutations (Olivier et al., 2005). In sorghum, the *bmr-6* gene is linked to reduce lignin due to a decrease in CAD activity (Buchholz et al., 1980), and the

allelic *bmr-12* and *bmr-18* genes is suggested to decrease caffeic acid OMT activity (Bout and Vermerris, 2003).

1.10 Objectives

The main objective of the study was to identify molecular-markers linked to high biofuel-related traits in grain x sweet sorghum through:

1. Quantification of sugar and fibre components, assess the agronomic yields across environments
2. To estimate impact of genetic and environmental variance on these traits
3. To estimate their heritability and correlation coefficient
4. To identify additive QTLs, QTL x environment interaction, and QTL x QTL interaction.

1.11 References

- Anglani, C. 1998. Sorghum carbohydrates – A review. *Plant food for human nutrition*. 52:77-83
- Abu Assar, A.H., R. Uptmoor, A.A. Abdelmula, M. Salih, F. Ordon, and W. Friedt. 2005. Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assembled by simple sequence repeats (SSRs). *Crop Sci* 45:1636-1644
- Asins, M. 2002. Present and future of quantitative trait locus analysis in plant breeding. *Plant breed* 121:281-291
- Barrierè, Y., C. Gulliet, D. Goffner, and M. Pichon. 2003. Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. A review. *Anim. Res* 52:193-228
- Bhatramakki, D., J. Dong, A. Chhabra, and G. Hart. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* 43:988-1002
- Beavis, W. 1998. QTL analyses: power, precision, and accuracy. In A.H. Paterson (ed), molecular dissection of complex traits. CRC Press, Boca Raton. 145-162
- Beta, T., and H. Corke. 2001. Genetic and environmental variation in sorghum starch properties. *J. Cereal Sci* 34:261-268
- Beta, T., H. Corke, L.W. Rooney, and J.R.N. Taylor. 2001. Starch properties as affected by sorghum grain chemistry. *J. Sci Food and Agric* 81:245-251
- Billa, E., D.P. Koullas, B. Monties, and E.G. Koukios. 1997. Structure and composition of sweet sorghum stalk components. *Ind. Crops Prod* 6:297-302
- Bittinger, T.S., R.P. Cantrell, and J.D. Axtell. 1981. Allelism tests of the brown midrib mutants of sorghum. *J. Hered* 72:147-148
- 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* 98: 320-328
- Bout, S., and W. Vermerris. 2003. A candidate-gene approach to clone the sorghum Brown midrib gene encoding caffeic acid O-methyltransferase. *Mol. Genet. Genomics* 269:205-214
- Buchholz, D.L., R.P. Cantrell, J.D. Axtell, and V.L. Lechtenberg. 1980. Lignin biochemistry of normal and brown midrib mutant sorghum. *J. Agric. Food Chem* 28:1239-1245
- Buranov, A.U., and G. Mazza. 2008. Lignin in straw of herbaceous crops. *Ind. Crops Prod.* 28:237-259

- Carpita, N.C. 1996. Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:445-476
- Castleden, C.K., N. Aoki, V.J. Gillepsie, E.A. MacRae, W.P. Quick, P. Buchner, C.H. Foyer, R.T. Furbank, and J.E. Lunn. 2004. Evolution and function of the sucrose-phosphate synthase gene families in wheat and other grasses. *Plant physiol* 135:1753-1764
- Chanterreau, J., G. Trouche, J.F. Rami, M. Deu, C. Barro, and L. Grivet. 2004. RFLP mapping of QTLs for photoperiod response in tropical sorghum. *Euphytica* 120:183-194
- Cherney, J.H., J.D. Axtell, M.M. Hasen, and K.S. Anliker. 1988. Forage quality characterization of a chemically induced brown-midrib mutant in Pearl millet. *Crop Sci* 28:783-787
- 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
- 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 markers-assisted selection for crop improvement: The basic concepts. *Euphytica* 142:169-196
- Cosgrove, D.J. 2005. Growth of the plant cell wall. *Nat Rev Mol Cell Biol* 6:850–861
- De Alencar Figueiredo, L.F., C. Calatayud, C. Dupuits, C. Billot, J.-F. Rami, et al., 2008. Phylogeographic evidence of crop neodiversity in sorghum. *Genetics* 179:997-1008
- Deu, M., F. Rattunde, and J. Chanterreau. 2006. A global view of genetic diversity in cultivated sorghums using a core collection. *Genome* 49: 168-180
- De Wet, J.M.J., J.R. Harlan, and E.G. Price. 1970. Origin of variability in Spontanea complex of *Sorghum bicolor*. *Am. J. Bot* 57:704-707
- Dhillon, M.K., H.C. Sharma, B.V.S. Reddy, R. Singh, and J.S. Naresh. 2006. Inheritance of resistance to sorghum shoot fly, *Atherigona soccata*. *Crop Sci* 46:1377-1383
- Dhillon, S.L., F.M. Shapter, R.J. Henry, G. Cordeiro, L. Izquierdo, and L.S. Less. 2007. Domestication to crop improvements: Genetic resources for *Sorghum* and *Saccharum* (Andropogoneae). *Ann. Bot* 100:975-989
- Dixon, R.A., F. Chen, D. Guo, and K. Parvathi. 2001. The biosynthesis of monolignols: a "metabolic grid", or independent pathways to guaiacyl and syringyl units? *Phytochemistry* 7:1069-1084
- Dogett, H. 1965. The development of the cultivated sorghums. In J.B. Hutchison et al., (ed), *Essays on crop plant evolution* (ed). London: Cambridge University press. 50-69

- Doggett, H. 1970. Sorghum. London, New York: Longman publ. Wiley.
- Dhugga, K.S. 2007. Maize biomass yield and composition of biofuel. *Crop Sci* 47:2211-2227
- Echeverria, E., M.E. Salvucci, P. Gonzalez, P. Paris, and G. Salerno. 1997. Physical and kinetic evidence for an association between sucrose-phosphate synthase and sucrose-phosphate phosphatase. *Plant physiol* 115: 223-227
- Elkhalifa, A.O., B. Schiffler, and R. Bernhardt. 2004. Selected physicochemical properties of starch isolated from fermented sorghum flour. *Starch* 56:582-585
- Ejeta, G., M.R. Goldsborough, E.M. Tunistra, A. Grote., Y. Menkir, et al. 2000. Molecular marker application in sorghum. In: B.I.G. Hausmann et al., (ed), application of molecular markers in plant breeding, training manual for a seminar held at IITA, Ibadan, Nigeria, 16-17 August 1999. ICRISAT, Patancheru 502324, Andhra Pradesh, India: 81-89
- Ezeogu, L.I., K.G. Duodu, and J.R.N. Taylor. 2005. Effects of endosperm texture and cooking conditions on the vitro starch digestibility of sorghum and maize flours. *J. Cereal Sci.* 42:33-44
- Fritz, J.O., R.P. Cantrel, V.L. Lechtenberg, J.D. Axtell, and J.M. Hertel. 1981. Brown midrib mutants in sudangrass and grain sorghum. *Crop Sci* 21:706-709
- Fritz, J.O., K.J. Moore, and E.H. Jaster. 1988. Digestion kinetics and cell wall composition of brown midrib sorghum x sudangrass morphological components. *Crop Sci* 30:213-219
- Gomez, L., D. Bancel, E. Rubio, and G. Vercambre. 2007. The microplate reader: an efficient tool for the separate enzymatic analysis of sugars in plant tissues-validation of a micro-method. *J. Sci. Food Agric* 87:1893-1905
- Gupta, S.C. 1995. Allelic relationships and inheritance of brown midrib trait in sorghum. *J. Hered* 86:72-74
- Gupta, P.K. 2002. Molecular markers and QTL analysis in crop plants. *Curr. Sci* 83:113-114
- Grabber, J.H., J. Ralph, C. Lapierre, and Y. Barrierè. 2004. Genetic and molecular basis of grass cell-wall degradability. I. Lignin-cell wall matrix interactions. *C.R. Biologies.* 327:455-465
- Haley, C., and L. Anderson. 1997. Linkage mapping of quantitative trait loci in plants and animals. In: P. Dear (ed), genome mapping-A practical approach. New York, Oxford University Press:49-71
- Harlan, J.R., and J.M.J. de Wet. 1972. A simplified classification of cultivated sorghum. *Crop Sci* 12:172-176

- Hart, G.E., K.F. Schertz, Y. Peng, and N.Y. 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
- Hausmann, B.I.G., V. Mahalakshini, B.V.S. Reddy, N. Seetharama, C.T. Hash, and H.H. Geiger. 2002. QTL mapping of stay-green in two sorghum recombinant inbred populations. *Theor. Appl. Genet* 106:133-142
- Hawker, J.S., and G.M. Smith. 1984. Occurrence of sucrose phosphatase in vascular and non-vascular plants. *Phytochemistry* 23:245-249
- Hazen, S.P., J.S. Scott-Craig, and J.D. Walton. 2002. Cellulose Synthase-Like genes of rice. *Plant physiol* 128:336-340
- House, L.R. 1985. A guide to sorghum breeding. 2nd ed. ICRISAT. India
- 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. USA* 87:4251-4255
- Jones, N., H. Ougham, and H. Thomas. 1997. Markers and mapping: we are all geneticist now. *New Phytol* 137:165-177
- Jotwani, M.G., G.L. Teetes, and W.R. Young. 1980. Elements of integrated control of sorghum pests. *FAO plant production and protection paper*. FAO, Rome, Italy: 159
- Jung, H.G, D.R. Mertens, and A.J. Payne. 1997. Correlation of acid detergent lignin and Klason lignin with digestibility of forage dry matter and neutral detergent fibre. *J. Dairy Sci* 80:1622-1628
- 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* 103: 266-276
- Kim, J-S., P.E. Klein, R.R. Klein, H.J. Price, J.E. Mullet, and D.M. Stelly. 2005. Chromosome identification and nomenclature of *Sorghum bicolor*. *Genetics* 169: 1169-1173.
- Kimber, C. 2000. Origins of domesticated sorghum and its early diffusion to India and China. In W.C. Smith and R.A. Frederiksen (ed), *Sorghum: origin, history, technology and production*. John Wiley. New York: 3-98
- Kimura, S., and T. Itoh. 1996. New cellulose synthesizing complexes (terminal complexes) involved in animal cellulose biosynthesis in the tunicate *Metandrocarpa uedai*. *Protoplasma* 194:156-163

- Klein, R.R., R. Rodriguez-Herrera, J.A. Schlueter, P.A. Klein, Z.H. Yu, and W.L. Rooney. 2001. Identification of genomic regions that affect grain-mould incidence and other traits of agronomic importance in sorghum. *Theor. Appl. Genet* 102:307-319
- Kong, L., J. Dong, and G.E. Hart. 2000. Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple sequence repeats (SSRs). *Theor. Appl. Genet* 101:438-448
- Lander, E., and D. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199
- Laopaiboon, L., P. Thanonkeo, P. Jaisil, and P. Laopaiboon. 2007. Ethanol production from sweet sorghum juice in batch and fed-batch fermentations by *Saccharomyces cerevisiae*. *World J. Microbiol. Biotechnol* 23:1497-1501
- Lerouxel, O., D.M. Cavalier, A.H. Liepman, and K. Keegstra. 2006. Biosynthesis of plant cell wall polysaccharides-a complex process. *Curr. Op. Plant Biol* 9:621-630
- Lin, Y.R., K.F. Schertz, and A.H. Paterson. 1995. Comparative analysis of QTLs affecting plant height and maturity across *Poaceae*, in reference to an interspecific sorghum population. *Genetics* 141:391-411
- Lin, Y., and S. Tanaka. 2006. Ethanol fermentation from biomass resources: current state and prospects. *Appl. Microbiol. Biotechnol* 69:627-642.
- Liu Y., Y. Ying, H. Yu, and X. Fu. 2006. Comparison of the HPLC method and FT-NIR analysis for quantification of glucose, fructose, and sucrose in intact Apple fruits. *J. Agric. Food Chem* 54:2810-2815.
- Lunn, J.E., A.R. Ashton, M.D. Hatch, and H.W. Heldt. 2000. Purification, molecular cloning and sequence analysis of sucrose-6-phosphate phosphohydrolase from plants. *Proc. Natl. Acad. Sci. USA* 97:12914-12919
- Maradiaga, J.L.M. 2003. Quantitative trait loci affecting the agronomic performance of a *Sorghum bicolor* (L.) Moench recombinant inbred restorer line population. PhD dissertation, Texas A&M University.
- McBee, G., and F.R. Miller. 1982. Carbohydrates in sorghum culms as influenced by cultivars, spacing and maturity over a diurnal period. *Crop Sci* 22:381-385
- McItosh, K. A., S.A. Charman, L.A. Borgen, and W.A. Charman. 1998. Analytical methods and stability assessment of liquid yeast derived sucrose. *J. Pharm. Biomed Anal* 17:1037-1045

- Menz, M.A., R.R. Klein, J.E. Mullet, J.A. Obert, N.C. Unruh, and P.E. Klein. 2002. A high density map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. *Plant Mol. Biol* 48:483-499
- Naidu, K., V. Singh, D.B. Johnston, K.D. Raush, and M.E. Tumbleson. 2007. Effects of ground corn particle size on ethanol yield and thin stillage soluble solids. *Cereal Chem* 84:6-9.
- Oliver, A.L., J.F. Pederson, R.J. Grant, T.J. Klopfenstein, and H.D. Jose. 2005. Comparative effects of the sorghum *bmr-6* and *bmr-12* genes: II. Grain yield, stover yield, and stover quality in grain sorghum. *Crop Sci* 45:2240-2245
- O'Neil, M.A., and W.S. York. 2003. The composition and structure of plant primary cell walls. In JKC Rose, *Plant cell wall*. CRC Press :1-54
- Papini-Terzi, F.S., F.R. Rocha, R. ZN. Vêncio, J.M. Felix., D.S. Branco, et al., 2009. Sugarcane genes associated with sucrose content. *BMC Genomics* 10
- Paterson, A.H. 1995. Molecular dissection of quantitative traits: progress and prospects. *Genome Res* 5:321-333
- 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, and R. Whitkus. 1994. Construction of an RFLP map in sorghum and comparative mapping in maize. *Genome* 37:236-243
- Porter, K.S., J.D. Axtell, V.L. Lechtenberg, and V.F. Colenbrander. 1978. Phenotype, fiber composition, and in vitro dry matter disappearance of chemically induced brown midrib (*bmr*) mutants of sorghum. *Crop Sci* 18:205-208
- Porter, K.S., J.D. Axtell, V.L. Lechtenberg, and V.F. Colenbrander. 1975. The brown midrib gene (*bmr*) in sorghum and its effect on forage quality. In: P.R. Mayaguez, *proc Int Sorghum workshop*, pp 471-482
- Quinby, J.R. 1973. The genetic control of flowering and growth in sorghum. *Adv. Agron* 25:126-162
- Rai, K.N., D.S. Murty, D.J. Andrews, and P.J. Bramel-Cox. 1999. Genetic enhancement of pearl millet and sorghum for the semi-arid tropics of Asia and Africa. *Genome* 42: 617-628
- Rami, J.F., P. Dufour, G. Trouche, G. Fliedel, C. Mestres, F. Davrieux, P. Blancard, and P. Hamon. 1998. Quantitative trait loci for grain quality, productivity, morphological,

- and agronomical traits in sorghum (*Sorghum bicolor* L. Moench). *Theor. Appl. Genet* 97:605-616
- Ragab, R.A., S. Dronovalli, M.A. Maroof Saghai, and Y.G. Yu. 1994. Construction of a sorghum RFLP linkage map using sorghum and maize DNA probes. *Genome* 37:590-594
- Reddy, K.V.S., and T.C. Davies. 1979. Attractants for *Atherigona* spp. including sorghum shoot fly, Asocata Rond. ICRISAT, Patancheru:16
- Rooney, L.W., and R.L. Pflugfelder. 1986. Factors affecting starch digestibility with special emphasis on sorghum and corn. *J. Anim. Sci* 63:1607-1623
- Rosenow, D.T., G. Ejeta, L.E. Clark, M.L. Gilbert, R.G. Henzell, A.K. Borell, and R.C. Muchow. 1996. Breeding for pre- and post flowering drought stress resistance in sorghum. In: Proceedings of International Conference on genetic improvement of Sorghum and Pearl millet, Sept 23-27, Lubbock, Texas: 400-411
- Sanchez, A.C., P.K. Subudhi, D.T. Rosenow, and H.T. Nguyen. 2002. Mapping QTLs associated with drought resistance in sorghum (*Sorghum bicolor* L. Moench). *Plant Mol. Biol* 48:713-726
- Seetharam, A. 2005. Biofuel in Africa: A criteria to choose crop models for food and fuel. *ATDF Journal* 4:3-12
- Shiringani, A.L. 2005. Evaluation of hard endosperm, bird-proof sorghum [(*Sorghum bicolor* L. (Moench)] and its effect on food quality. MSc Dissertation. University of the Free State.
- Srinivas, G., K. Satish, S.M. Mohali, R.N. Reddy, R. Madhusudhana, D. Balakrishna, B.B. Venkatesh, C.J. Howarth, and N. Seetharama. 2008. Development of genic-microsatellite markers for sorghum staygreen QTL using comparative genomic approach with rice. *Theor. Appl. Genet* 117: 703-717
- Sticklen, M.B. 2008. Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. *Nat. Rev. Genet* 9:433-443
- Stitt, M. 1996. Metabolic regulation of photosynthesis. In: N. Baker (ed), *Advances in photosynthesis: Photosynthesis and the environment*. Dordrecht, Kluwer 50:151-190.
- Tanksley, S.D. 1993. Mapping polygenes. *Annu. Rev. Genet* 27:205-233
- Tarpley, L., and D.M. Vietor. 2007. Compartmentation of sucrose during radial transfer in mature sorghum culm. *BMC Plant Biol* 7:33
- Taylor, N.G. 2008. Cellulose biosynthesis and deposition in higher plants. *New Phytologist* 178:239-252.

- Tuinstra, M.R., G. Ejeta, and P.B. Goldsbough. 1998. Evaluation of near-isogenic sorghum lines contrasting for QTL markers associated with drought tolerance. *Crop Sci* 36:1337-1344
- Tuinstra, M.R., E.M. Grote, P.B. Goldsbough, and G. Ejeta. 1996. Identification of quantitative trait loci associated with pre-flowering drought tolerance in sorghum. *Crop Sci*. 36:1337-1344
- Uptmoor, R., W.G. Wenzel, W. Friedt, G. Donalson, K. Ayisi, and F. Ordon. 2003. Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions from southern Africa by RAPDs, AFLPs and SSRs. *Theor. Appl. Genet* 106:1316-1325.
- Vaughan, D.A., E. Balázs, and J.S. Heslop-Harrison. 2007. From domestication to Super domestication. *Ann. Bot* 100:893-901
- Vavilov, N.I. 1951. The origin, variation, immunity and breeding of cultivated plants. Translated by K. Start. *Chron. Bot* 13:1-366
- Velterop J. S., and F. Vos. 2001. A rapid and inexpensive microplate assay for enzymatic determination of glucose, fructose, L-malate and citrate in Tomato (*Lycopersicon esculentum*) extracts and in Orange Juice. *Phytochem. Ann* 12:299-304
- Vermerris, W., A. Saballos, G. Ejeta, N.S. Mosier, M.R. Ladisch, and N.C. Carpita. 2007. Molecular breeding to enhance ethanol production from corn and sorghum stover. *Crop Sci* 47:S143-S153
- Vignols, F., J. Riggau, M.A. Torres, M. Capellades, and P. Puigdomenech. 1995. The *brown midrib3 (bm3)* mutation in maize occurs in the gene encoding caffeic acid O-methyltransferase. *Plant Cell* 7:407-416
- Vogel, J. 2008. Unique aspects of the grass cell wall. *Curr. Op. Plant Biol* 11:301-307
- Wang, D., S. Bean, J. McLaren, P. Seib, R. Madl, M. Tuinstra, Y. Shi, M. Lenz, X. Wu, and R. Zhao. 2008. Grain sorghum is a viable feedstock for ethanol production. *J. Ind. Microbiol. Biotechnol* 35:313-320
- Whetten, R.S., J.J. Mackay, and R.R. Sederoff. 1998. Recent advances in understanding lignin biosynthesis. *Ann. Rev. Plant Physiol. Plant Mol. Biol* 49:585-609
- Zeng, Z.-B. 1994. Precision mapping of quantitative trait loci. *Genetics* 136:1457-1468
- Zhan, X, D. Wang, S.R. Bean, X. Mo, X.S. Sun, and D. Boyle. 2006. Ethanol production from supercritical-fluid-extrusion cooked sorghum. *Ind. Crop Prod.* 23:304-310
- Zhin-Ben, Y., S. Yi, L. Xiao-Hong, Z. Wei-Jun, Y. Min, and C. Li-Xia. 2006. Advances in genetic mapping of the sorghum genome. *Chin. J. Agr. Biotechnol* 3:155-161

Chapter 2

Genotype-environmental analysis of RIL population segregating for sugar-related traits in *Sorghum bicolor* L. Moench.

Amukelani Shiringani & Wolfgang Friedt

Abstracts: Sweet sorghum has the potential of storing sugar in its stem and its commercial value is based on exploiting these sugars. A population of 213 recombinant inbred lines derived from a cross between grain and sweet sorghum were used in our study to estimate the genetic variance, and genotype by environmental interaction on sugar components and sugar-related agronomic traits, estimate their heritability and understand their relationship. The field experiments were carried out at two locations in 2007 and 2008 growing season. The data was analysed using SAS[®] 9.1 software. The mixed model TYPE III and restriction maximum likelihood (REML) method were used to analyse the variance and to estimate the variance components, respectively. The variance components were used to calculate heritability of the traits. Parents performed as expected where the sweet sorghum parent SS79, was taller, and produced higher sugar components than grain sorghum parent M71. Mean squares obtained by TYPE III model showed genotypes, and genotypes by environment were a significant source of variation in most traits except for estimated juice weight whereas location and year were not a significant source of variation. The results of REML showed that the contribution of genetic variance was significant in all traits but small, while the contribution of location variance was high but not significant. The interaction among genotype-location-year showed a significant influence in eight out of ten traits studied. In most of the traits, the heritability and correlation estimates were moderate to high suggesting the possibility of genetic enhancement and that gene association exist among these traits.

2.1 Introduction

Sorghum bicolor (L.) Moench is an important crop in the semi-arid tropics of Africa and southern Asia, and grown substantially by small-holder farmers. Sorghum is closely related to sugarcane (Tarpley and Vietor, 2007), and the commercial value of these crops is based on exploiting their stems which contains high amount of sugars (Billa et al., 1997). Sorghum and sugarcane are C₄ species, characterized by high photosynthetic efficiency. However, compared to sugarcane, sorghum is a species showing extensive genetic variation, including drought and heat tolerant varieties, which would enable the usage of marginal land that is not

suitable for sugarcane cultivation. Besides, the growth period of sorghum is 3-5 months compared to 18 months in sugarcane. Because the complicated genome of sugarcane hinders the progress of its manipulation, *Sorghum bicolor* serves as the role model of the crop.

The success of bioethanol production in Brazil has received much of the world attention and different countries are starting to make regulations and draw strategic plans to adopt the environmental friendly fuel. Europe has set the biofuel usage target of 5.76% in 2010, which can be amended to 10%, if the promotion of ‘high quality fuels’ can function ([europa.eu/legislation_summaries/internal market](http://europa.eu/legislation_summaries/internal_market)). However, the success and commitments varies by country, and Germany is reported the major contributor. South Africa used bioethanol from sugarcane between the 1920s and the 1960s. The industry failed due to low international crude oil prices (National department of Energy, South Africa, strategic draft, 2005). Recent high prices of crude oil, political instability of oil countries, cutting its dependence on foreign countries, the needs to address rural unemployment and socio-economic problems have caused the South African government to re-visit the bioethanol strategic plan.

As the industry takes shape, there is a lot that needs to be done such as developing appropriate harvesting method, storage of the juice and fermentation technology, and above all to have crops that can supply us with adequate, reliable, and stable sugar and juice yield across generations or location and years. As mentioned before, sweet sorghum has the potential to serve as biofuel crop. The total juice components in sweet sorghum stalk consist of 23.33% of different sugars (Anglani, 1998). The main sugars include sucrose (9.4%), glucose (3.4%) and fructose (3%). Other sugars components are available in quantity less than 1%. However, the concentration percentage differs with varieties and growing conditions. Sugars are found in the phloem and pith of the sorghum stalk (Billa et al., 1997) or simply culm parenchyma (Fortmeier and Schubert, 1995).

Increasing sugar and juice content is one of the objectives of plant breeders and the success will depend on the ability to identify and select the best genotypes across environments, to understand the inheritance of sugars in sweet sorghum, to estimate the genetic variance from the total phenotypic variance, and to distinguish the parameters that cause rank changes in a multi-environmental trial. Murray et al. (2008) and Ritter et al. (2008) reported that stem sugar content in sorghum is quantitatively inherited with additive effect without a noticeable

dominance effect. Although a multi-environmental trial is the best way to assess the performance of the genotypes, changes and inconsistency in expressions of a trait by a genotype across environments or genotypes-environmental interactions is a major problem in plant improvements. Several statistical approaches have been proposed to analyze complicated traits (Burgueno et al., 2007; Denis et al., 1997; Piepho, 1997, and Yang, 2002; 2007). These methods aim at partitioning the phenotypic values of the components traits into contribution due to different effects. The objectives of the study were to 1) estimate the genetic parameters and test the significance of G-E interaction of sugar related traits in sorghum 2) Estimate the H^2 of these traits 3) Estimate the relationship between these traits.

2.2 Materials and methods

2.2.1 Plant material

A population of 213 recombinant inbred lines derived from a cross between SS79 (sweet sorghum) and Macia-SA denoted as M71 (grain sorghum), developed in the Agricultural Research Council of South Africa (ARC), was used for this study. The female parent SS79 was collected from a traditional farmer in Limpopo province of South Africa. It is approximately 300 cm tall with long internodes, have thin but sweet juicy stalks. It has open and loose panicles with light brown seeds. The male parent, M71, originate from ICRISAT-Bulawayo in Zimbabwe bred under a Sorghum and Millet improvement program. It is an open pollinated variety (OPV), early maturing and high grain yielding, with white grains. It has short internodes and is approximately 140 cm tall. The stems are juicy but the juice is insipid. After the introduction of SS79 and M71 at the ARC, a single ear was selected from each variety and selfed until homozygosity was reached. The mapping population of our study was developed using conventional single seed descent method until $F_{5:6}$ generation.

2.2.2 Experimental design

A total of five experiments were planted in 04 May 2007 and 14 May 2008 at three different locations: Gross Gerau (2 trials), Rauischholzhausen (2 trials) and Giessen (1 trial in 2008). For balanced analysis, the Giessen trial was excluded because of poor germination. The characteristics of the locations are presented in table 2.1. All experiments were laid in a Randomized Complete Block Design (RCBD) with three replications, where Gross Gerau (GG) had field experiments and Rauishholzhausen (RH) had pot trials. At GG, a one-row plot of 5 m row length, and a three-row plot of 3 m row length were planted in 2007 and 2008,

respectively. In both years, the inter-row was 0.3 m spaced and seeds were sown at 0.03 m depth and at 0.15 m interval. In the pot trial of RH, two seeds were sown per pot with a diameter of 0.20 m at 0.02 m depth. Best conditions for plant growth were implemented in both locations.

The trials were harvested manually by cutting the plant at the base by a scissor in all the three locations. In the three field experiments three middle plants were selected at random for all data collection and later the average of the three plants was used. However, in the pot trials both plants per pot were harvested and averaged. Some lines were photoperiod sensitive, and resulted in differences in flowering dates of up to 2 months. To harvest the plots at the same physiological stage, each line was harvested from 40 days after anthesis (DAA), which estimated the hard dough stage. At hard dough stage, sugar accumulation has been reported to reach the optimum. Because the plants developed tillers at early stage in both locations, the main stem and its tiller were treated as a single plant during harvest.

2.2.3 Phenotypic evaluation

Sampling was done in the following sequence: Flowering date was recorded when 50% of the plot had flowered. The length of the plant from the ground to the panicle tip was measured to estimate plant height. Stem diameter was measured 20 cm above ground and the sum of tillers per plant was counted to estimate number of tillers per plant. Panicles were cut at the flag leaf, weighed to estimate fresh panicle weight per plant. The juice weight was estimated as the difference between fresh stripped stalk mass and dry stripped stalk mass. The fresh stripped stalks were milled and then dried in an oven at 105 °C for dry mill. This difference gives the similar juice weight as that collected by roller miller in the same material (Wenzel, personal communication). The fresh main stalk was pressed and 2-3 droplets of juice were collected on a sucrose-sensitive refractometer to measure the Brix and another juice was squeezed to fill 2 ml tubes for future glucose and sucrose analysis.

Table 2.1: Biophysical information during field experiments in GG and RH

Data	Gross Gerau		Rauischholzhausen	
	2007	2008	2007	2008
Averaged accumulative temperature [°C]	85.7	84.6	82.0	81.9
Accumulative rainfall [mm]	311.1	445.9	275.6	481.2
Latitude	49° 45 min North		50° 45 min North	
Longitude	8° 29 min East		8° 52 min East	
Altitude (m)	90.8		200.0	
Soil type	Sandy Loam		Sandy Loam	
Soil pH	6.8	6.4	5.2	5.0

Accumulative temperature and rainfall from planting to the last harvest (May to end September). The Averaged accumulated temperature was obtained by averaging the temperature per month and adding the averages.

2.2.4 Glucose and sucrose analysis

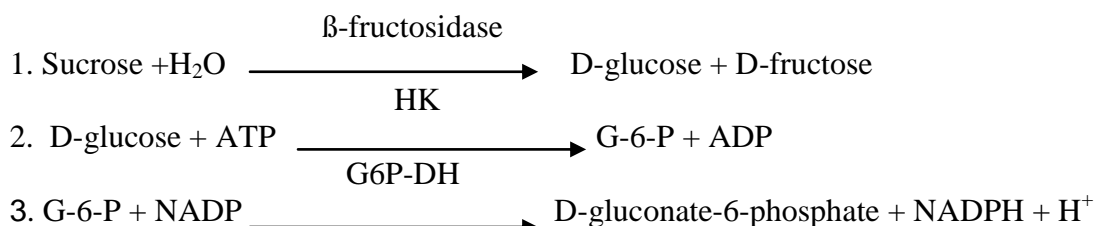
Carbohydrate quantification was done only in the main stem. The samples were frozen at -20 °c until the end of harvest. The three plants per replication were pooled to obtain an average sample extract. The pooled juice was diluted to 1:600 $\mu\text{L/ml}$ ratios and then filtered. The enzymatic Bioanalysis kit produced by Boehringer Mannheim/ r-biopharm was used for the analysis. Instead of using the cuvettes as per manufacturer manual, the 96-well microtitre plates were used for the analysis. Reagents were adjusted to make 200 μl total reaction volume. Measurements were taken at 340 nm wavelength using the Tecan, sunrise spectrophotometer. A dilution series for provided glucose and sucrose standards was done as references of linear regression to use in concentration calculation. The difference in absorbance at 340 nm is directly proportional to sugar concentration in each well, with reference to standards of varying concentration using a calibration curve (Henderson et al., 1998; Viola and Davies, 1992). The results were verified using the manufacturer's formula:

$$C = (V \times MW / \epsilon \times d \times v \times 1000) \times \Delta A \text{ [g/l]}$$

Where V is the final volume, v is the sample volume, MW= molecular weight of the substance to be assayed, d= light path and

ϵ is the extinction of NADPH at 340 nm= $6.3 [1 \times \text{mmol}^{-1} \times \text{cm}^{-1}]$.

The principle of measurement is based on the following steps:



The end-point of the assay is based on the reduction of NAD to NADH.

Sugar content was calculated according to Ritter et al. (2008). The concentration of fructose was replaced by concentration of glucose because fructose was not quantified in our study. Glucose and fructose can give similar content in sorghum (Godwin Ian, personal communication) and sugarcane (Robertson et al., 1996).

2.2.5 Statistical analysis

The SAS[®] procedure PROC MIXED from SAS[®] 9.0 software was used for all the statistical analysis. The analysis of the variance was done using the TYPE III model. All variance parameters such as genetic (G), location (loc), genotype x location (G x L), year (Y), genotype x year (G x Y), year x location (Y x L), which was sometimes referred as the environments in results, genotype x year x loc (G x Y x L, or G x E), and harvesting date were treated as random. There were no differences between the replications, therefore replications and the interactions between genotypes and replication were omitted during the whole analysis. Restricted maximum likelihood (REML) approach was used to estimate the genetic parameters and to test the significance of the different causes of genotype-environmental interactions. As reported by Gravois and Bernhardt (2000), Littell et al. (2006), and Yang (2002), the general model to estimate the variance components in a mixed model can be defined as: $Trait_{ijkl} = \mu + G_i + L_j + GL_{ij} + Y_i + GY_{il} + YL_{ij} + GE_{ijl} + H_{ijk}l + E_{ijkl}$

Where $Trait_{ijkl}$ was trait of the i^{th} genotype (G) in the k^{th} replicate (R) within the i^{th} year (Y) and j^{th} location (L), the μ was the overall mean, $H_{ijk}l$ was the harvest date and E_{ijkl} was the residual error.

Secondly, genotypes were treated as fixed effects to obtain simple statistics results and the corrected trait means. This was done with least significance mean commands. The corrected trait means was further used to estimates the correlation effects among traits. Broad-sense heritability for each trait was calculated as follows:

$$H^2 = [\sigma^2_G / (\sigma^2_G + \sigma^2_{GY/y} + \sigma^2_{GL/l} + \sigma^2_{GYL/yl} + \sigma^2_{E/ryl})] \quad (\text{Littell et al., 2006})$$

2.3 Results and discussion

2.3.1 Phenotypic evaluation.

The results of the performance of parents and RILs are presented in table 2.2. In GG, the average performance of parents and the recombinant inbred lines was better in almost all traits than in RH. The RIL parents performed as expected based on their characteristics. For example, SS79 the sweet sorghum parent, was taller than M71 in all field trials, although

there were rank changes over environments. Besides, it had high sucrose, glucose and estimated sugar content and less panicle weight. Both parents are known to be juicy although the juice has not been quantified before. Genotypes flowered between 73-165 days after planting across the environments. The female parent, SS79 was consistent across environment flowering at 111 or 112 days after planting, while the male parent M71 flowered between 90-100 days after planting. However, M71 flowered earlier than SS79. RH2007 was in average the late flowering trial compared to the other three. The average height of the RILs in GG was 56.66% higher than in RH. The height of most of the lines was within the range of the parents, however, an average of approximately 5% showed transgression segregation. The RILs parents were not diverse enough in terms of stem diameter. However, the M71's stem was thicker than SS79 in most of the trials except for the trial in GG2008. For stem diameter, more than 59.80% of the RILs had thicker stem outside the range of the parents. SS79 resulted in more tillers than M71 in all the environments except for GG2008, and the most tillers (average of 4 tillers per plant) were observed in GG2007. Genotypes were highly significantly different, with tillers among RILs ranging from 0-6 across the environments. For better comparison, the average tiller numbers were subjected to decimals. The ability of sorghum to tiller can have a significant advantage on the crop if most of the tillers will be harvestable or mature with the main stem. In GG2007 and GG2008, tillers in most genotypes appeared 2-3 weeks after the emergence of the main stem, thus matured at the same time. The ability of sorghum to regrow, can solve the critical concerns of biofuel of leaving the soil bare, thus reduce erosion and provide additional organic matter.

The GG2008 had higher average juice weight compared to the other three trials. The less accumulation of juice was observed during plant development in RH2008. However, the main reason for the decrease is not known. The average brix ranged from 5.67 - 22.67% across the locations, was 11% less in RH than that of GG, and 30% of the genotype produced more than sucrose-based soluble solids (brix) than both parents. The brix mean and range was higher and wide respectively compared to what was reported by Ritter et al. (2008). The sweet parent used in this cross was having higher brix of 17.60% than theirs of 12.60%. In all environment combined, glucose, sucrose and estimated sugar content ranged from 3.17 – 28.70 g/l, 31.73 – 216.14 g/l, and 26.91 – 307.6 g/l, respectively. In RH, concentration of glucose, sucrose and estimated sugar content in stem juice was 26.72%, 20.43% and 25.98% lower than that of GG, respectively. The decrease in fresh panicle weight, estimated sugar content, sucrose and glucose in RH is probably due to lines flowering late when the temperature started decreasing.

This might have tempered with the physiology of the plants. Fortmeier and Schubert (1995) reported that the crop starts accumulating these sugars immediately after anthesis until physiological maturity. During this stage it undergoes two different types of sugar storage: temporary stage, a stage where sugar is later transported from the stalks to the grain and the permanent storage where sugar is forever kept in the stalks.

2.3.2 Analysis of the variance and variance components

The results of analysis of the variance are presented in table 2.3. Mixed model TYPE III mean squares showed that genotypes were highly significantly ($P < 0.001$) different in all the traits except for estimated juice weight. However, there was no significant difference observed between location and year. The interaction between lines and location ($G \times L$) was significant for flowering dates and stem diameter, the interaction between genotypes and year ($G \times Y$) was significant for flowering dates, glucose, sucrose and estimated sugar content, and the interaction between genotype-location-year ($G \times L \times Y$) was highly significant ($P < 0.001$) for flowering dates, plant height, stem diameter, brix and estimated sugar content. Harvest dates showed significance at $P < 0.05$ in brix, glucose, sucrose and sugar content. Estimates of variance components by restricted maximum likelihood (table 2.4) indicated that locations were not a significant source of variation for all the traits studied. However, the location effect was a greater source of variation in terms of percentage contribution to the total variance in most traits. Unlike location effect, genetic effect was low but did significantly influence all traits. Year showed to be a highly significant source of variation to flowering dates contributing at least 54.9% of the total variance, while year, as well as $G \times Y$ and $G \times L$ had a negligible effect on plant height estimated at zero. Besides, both year and location had a non-significant and negligible effect to glucose. It can also be noted that all variance parameters studied had a significant contribution to flowering date. Quinby and Karper (1945) reported that the major flowering gene in sorghum is regulated by photoperiod while minor genes are influenced by temperature. Although the variance estimate of harvesting dates was low, it showed to be a significant source of variation to brix, fresh panicle weight, tiller numbers per plants and glucose.

Breeding for high sugar content in the stem is challenging because of high environmental effects. Although the parents were stable in most of the traits, the RILs were showing rank changes across the environments that resulted from the interaction between genotype-location-year. Besides, the differences in the performance of these sugar-related traits could

be attributed to differences in biophysical data (table 2.1) and other factors such as planting dates and cultural practices, which were not investigated. These rank changes would have a serious effect on selection progress for improving sugar-related traits.

2.3.3 Estimating traits heritability

The results of broad-sense heritability are presented in table 2.4. Broad-sense heritability was high in stem diameter (60%), number of tillers per plants (68%), fresh panicle weight (69%) and plant height (81%). This indicates that genotype x environmental interaction contribution was smaller than that of genetic contribution in these traits. A moderate heritability estimate was obtained in glucose content (47%), estimated sugar content (50%), sucrose content (52%) and brix (59%), and flowering dates and stem juice showed low heritability of 33% and 17%, respectively. Most of the traits in this study had similar heritability estimates to those reported by Brown et al. (2006), Murray et al. (2008), and Ritter et al. (2008). However, the heritability of the flowering dates obtained was relatively lower than in their studies. This can be due to the adaptation of our lines to South African photoperiod which is different from German longer summer days.

2.3.4 Correlation

Table 2.5 presents Pearson correlation coefficients of sugar-related traits. Flowering dates correlated significantly with all the traits. The strongest positive correlation was observed between flowering dates and stem juiciness ($r = 0.445$), followed by stem diameter ($r = 0.369$) and brix ($r = 0.251$) the strongest negative correlation was observed between flowering dates and fresh panicle weight ($r = -0.626$). Flowering dates showed a significant, but weaker correlation with plant heights ($r = 0.183$), which is higher than that reported by Ritter et al. (2008) of $r = 0.164$ but lower than what Murray et al. (2008) and Wenzel (1990) reported of $r = 0.29$ and $r = 0.5$, respectively. This is probably due to environmental effect to flowering dates. Fresh panicle weight has negatively and significantly correlated with all traits, except for a positive, non significant correlation with number of tillers and a moderate significant positive correlation with sugar content. Number of tillers only correlated with estimated juice content ($r = 0.20$). Brix and sucrose showed a strong positive correlation ($r = 0.606$), even though a higher correlation estimates was expected. Ritter et al. (2008) reported correlation estimates of 0.946. The reduction in the correlation of brix and glucose might be that our juice samples were not frozen at liquid nitrogen during harvest as they did. The strong correlation between sucrose and brix suggest that brix can be used for selection instead of the laborious

sucrose quantification. Glucose correlated negatively and significantly with sucrose ($r = -0.334$), but correlated positively with brix ($r = 0.122$). The estimated sugar content correlated with all traits except with juice weight. The moderate and strongest correlation was observed between sugar content and sucrose ($r = 0.988$), glucose ($r = 0.556$), brix ($r = 0.386$), and stem diameter ($r = 0.320$)

2.4 Conclusion

Our results show that locations have the biggest effect on most of the sugar traits even though it was not significant. The interaction between genotype-location-year showed to have a significant contribution and this reduces the heritability and correlation effects, and also affects the precision performance of the genotypes. Genetic variance showed less but significant contribution, but the small or negligible contribution by genotype-location, and genotype-year resulted in moderate to high heritability estimates in most of the traits. These suggest that it is possible to manipulate these traits through plant breeding. Besides, sugar-related traits enhancement can be quicker because most of these traits have significantly correlated with each other indicating that gene association or pleiotropism exists among these traits.

2.5 Acknowledgement

The authors would like to thank the excellent technical assistance of Mario Tolksdorf, Markus Kollmer and Malte Luh. This work has been financially supported by the National Research foundation and the agricultural research foundation of South Africa, and Deutsche akademische Austausch Dienst (DAAD).

Table 2.2: Performance of parents and RILs for sugar- related traits across environments

	Gross Gerau										Rauischolzhausen									
	2007					2008					2007					2008				
	SS79	M71	RIL mean	SD	Range	SS79	M71	RIL mean	SD	Range	SS79	M71	RIL mean	SD	range	SS79	M71	RIL mean	SD	range
FLD	111	100	104.61	12.22	75-151	112	90	99.44	12.89	73-132	112	92	114.06	11.52	91-165	113	97	103.50	14.03	79-140
PH (cm)	240.42	151.67	177.02	57.02	87.5-308.75	332.54	130.63	195.42	74.62	56.75-398.25	210.12	110.6	142.57	43.86	50-260.50	170	80	94.95	30.90	50-180.17
Diameter (cm)	1.76	1.94	2.13	0.28	1.40-4.50	2.15	2.08	2.18	0.27	0.93-3.23	1.37	1.50	1.44	0.17	0.60-2.10	1.36	1.52	1.30	0.24	0.31-1.93
Tillers /plant	4	1	2.42	0.65	0-5	1	2	1.29	0.65	0-4.00	3	2	3	0.75	0-6.00	3	2	2.24	0.80	0-6.00
PW g/plant	43.70	69.86	90.30	35.04	8.75-250.79	80.00	211.44	102.54	45.65	13.33-370.00	8.19	35.00	20.40	9.94	1.9-66.04	NA	NA	NA	NA	NA
Juice g/plant	384.82	229.38	308.57	151.97	11.43-990.16	1242.00	255.76	597.83	296.94	124.00-2135.00	223.42	194.64	193.96	81.06	5.13-577.40	102.52	89.63	68.18	28.34	4.15-155.65
Brix %	17.44	12.78	17.194	2.88	5.67-22.67	17.60	9.57	14.80	3.52	3.90-21.33	16.36	9.30	14.37	2.53	6.70-20.40	16.53	11.20	13.45	2.21	6.4-19.00
Glucose g/l	16.38	9.65	12.09	2.84	4.37-28.70	13.52	6.18	6.24 2.02	2.02	2.35	10.12	2.97	3.49	1.76	1.76-20.08	9.5	3.17	9.95	2.80	3.17-18.94
Sucrose g/l	168.99	85	152.35	44.80	37.83-240.78	121.62	47.14	113.83	40.31	21.97-209.59	90.04	32.56	93.23	36.06	31.73-216.14	82.63	30.86	104.83	28.89	21.20-215.69
Sugar content g/l	201.75	105.99	177.25	42.10	57.61-307.6	133.99	54.18	126.36	42.39	26.91-225.85	95.98	39.22	100.13	37.00	37.37-222.02	105.86	37.20	124.60	33.13	28.83-246.93

FLD: Flowering dates PH: Plant height Diameter: Stem diameter Tillers: number of tillers per plant PW: Fresh panicle weight Juice: estimated juice weight Glucose: Glucose content Sucrose: Sucrose content Sugar content: Estimated sugar content SD: standard deviation NA: Not assayed.

Table 2.3: Mean squares of sugar-related traits analysed by TYPE III mixed model

Source of variation	FLD		Height (cm)		Diameter (cm)		Juice (g/plant)		PW (g/plant)		Brix %		Glucose (g/l)		Sucrose (g/l)		Sugar content (g/l)	
	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS
Genotypes	213	68.49**	213	12408**	213	0.21***	213	652438	213	1496.68**	213	16.86**	213	9.81***	213	2219.81***	213	2325.33***
Location	1	67.98	1	1.6E-4	1	3.74E-10	1	539663			1	149.52	1	17.30	1	14522.00	1	6.18E-4
Year	1	693.639	1	6.76E-9	1	5.77E-12	1	16378			1	2.01	1	8.96	1	1246.42	1	3.46E-8
G*L	208	38.15**	209	1869.66	209	0.09*	209	672903			208	8.01	210	5.41	210	1498.82	210	1655.33
G*Y	209	29.08*	209	2415.74	209	0.06	209	690536			208	7.35	209	7.99**	209	1912.30*	209	7859.76**
L*Y	1	30.78	1	1.51E-4	1	1.49E-11	1	5411107	2	7687.46***	1	47.68*	1	155.79**	1	5897.03*	1	3.5E-6**
G*L*Y	161	21.81**	156	54706**	156	0.07***	155	810620	376	1065.39***	156	11.38**	159	5.43	159	1365.62*	160	45358**
HD	44	NA	44	506.19	44	1.332	44	475737	31	649.63	44	3.33*	44	4.88*	44	741.83*	44	831.95*
Residual	1548	8.74	1521	1098.50	1522	61.21	1515	701090			1503	4.68	1539	5.39	1538	1132.36	1534	1097.70

MS: Mean squares DF: Degree of freedom FLD: Flowering dates Height: Plant height Diameter: Stem diameter PW: Fresh panicle weight Juice: Estimated juice weight Sugar content: Estimated sugar content Glucose: Glucose content Sucrose: Sucrose content G: Genotype L: Location Y: Year HD: Harvesting dates. * Significant at the 0.05 probability level ** significant at the 0.01 probability level *** significant at the 0.001 probability level. Fresh panicle weight was quantified in three environments and for balanced analysis, environments were used instead of across locations and years. Therefore, Y*L represents environments and G*L*Y represents genotype *environment.

Table 2.4: Variance components and Heritability estimates for sugar-related traits

	Heritability	Genetic	Location	Year	G*L	G*Y	L*Y	G*L*Y	HD	Residual
FLD	0.33	0.042***	0.009	0.549***	0.064**	0.029**	0.049***	0.092***	NA	0.168
Diameter (cm)	0.60	0.036***	0.806	0	0.019***	0	0.013	0.020***	0	0.106
Brix %	0.59	0.106***	0.142	0.083	0	0	0.028	0.164***	0.068**	0.409
Juice (g/plant)	0.18	0.016	0.023	0	0	0	0	0.054	0.008	0.907
PW (g/plant)	0.69	0.022***	0.643	5.94E-14	0.006	0.017	0	0.016	0.103***	0.202
Tillers/plant	0.68	0.056***	0.209	0.310	0	0.025*	0.052	0.046***	0.006*	0.297
PH (cm)	0.81	0.212***	0.345	0	0	0	0.114	0.139***	0.002	0.190
Glucose content (g/l)	0.47	0.035***	0	0	0.001	0.019*	0.693	0.004	0.010**	0.230
Sucrose content (g/l)	0.52	0.092***	0	0.170	0.012	0.069**	0.177	0.058**	0.005	0.360
Sugar content (g/l)	0.50	0.048***	0.144	0	0.011	0.0520***	0.292	0.073***	0.007	0.372

FLD: Flowering dates PH: Plant height Diameter: Stem diameter PW: Estimated panicle weight Tillers: number of tillers per plant Juice: Estimated juice weight Sugar content: Estimated sugar content HD harvesting dates G: Genotype L: Location Y: Year NA: Not assayed * significant at the 0.05 probability level ** significant at the 0.01 probability level *** significant at the 0.001 probability level Fresh panicle weight was quantified in three environments and for balanced analysis, environments were used instead of across locations and years. Therefore, Y*L represents environments and G*L*Y represents genotype *environment.

Table 2.5: Pearson correlation coefficient of sugar-related traits

	FLD	PH	Diameter	PW	Tillers	Juice	Brix	Glucose	Sucrose
PH	0.183**								
Diameter	0.369***	-0.277***							
PW	-0.626***	-0.128	-0.141*						
Tillers	-0.148*	-0.003ns	-0.111ns	0.024ns					
Juice	0.445***	0.741***	0.016ns	-0.239***	0.195***				
Brix	0.251***	0.626***	-0.056ns	-0.290***	0.080ns	0.636***			
Glucose	0.147*	0.088 ns	-0.173*	-0.133***	0.051 ns	0.239***	0.122**		
Sucrose	0.152*	0.427***	-0.051 ns	-0.141***	0.046ns	0.363***	0.606***	-0.334***	
Sugar content	-0.142**	0.214***	0.320***	0.256***	-0.052*	0.011ns	0.386***	0.556***	0.988***

FLD: Flowering dates PH: Plant height Diameter: Stem diameter PW: Fresh panicle weight Tillers: number of tillers per plant Juice: Estimated juice weight Glucose: Glucose content Sucrose: Sucrose content Sugar content: Estimated sugar content ns: non significant * significant at the 0.05 probability level ** significant at the 0.01 probability level *** significant at the 0.001 probability level

2.6 References

- Anglani, C. 1998. Sorghum carbohydrates-A review. *Plant foods for human nutrition* 52:77-83
- Billa, E., D.P. Koullas, B. Monties, and E.G. Koukios. 1997. Structure and composition of sweet sorghum stalk components. *Ind. Crops Prod* 6:297-302
- Brown, P.J., P. Klein, E. Bortiti, C. Acharya, W. Rooney, and S. Kresovich. 2006. Inheritance of inflorescence architecture in sorghum. *Theor.Appl.Genet* 113:931-942
- Burgueno, J., J. Crossa, P.L. Cornelius, R. Trethowan, G. McLaren, and A. Krishnamachari. 2007. Modelling additive x environment and additive x additive x environment using genetic covariances of relatives of wheat genotypes. *Crop Sci* 47:311-320
- Denis, J.-B., H.-P. Piepho, and F.A. van Eeuwijk. 1997. Modelling expectation and variance for genotype by environment data. *Heredity* 79:162-171
- Fortmeier, R., and S. Schubert. 1995. Storage of non-structural carbohydrates in sweet sorghum [*Sorghum bicolor* (L.) Moench.]: Comparison of sterile and fertile lines. *J.Agronomy & Crop Science* 175:189-193
- Gravois, K.A., and J.L. Bernhardt. 2000. Heritability and genotype x environmental interactions for discoloured rice kernels. *Crop Sci* 40:314-318
- Henderson, S.K., C.A. Fenn, and J.D. Domijan. 1998. Determination of sugar content in commercial beverages by density. A novel experiment for general chemistry courses. *J.Chem. educ* 75:1122-1123
- Littell, R.C., G.A. Milliken, W.W. Stroup, R.D. Wolfinger, and O. Schabenberger. 2006. SAS[®] for mixed models. Second edition. SAS Institute Inc., Cary, NC, USA.
- Murray, S.C., A. Sharma, W.L. Rooney, P.E. Klein, J.E. Mullet, S.E. Mitchel, and S. Kresovich. 2008. Genetic improvement of sorghum as biofuel feedstock I: QTL for stem sugar and grain non-structural carbohydrates. *Crop Sci* 48:2165-2179
- Piepho, H.-P., 1997. Analyzing of genotype-environment data by mixed models with multiplicative effects. *Biometrics* 53:761-766
- Quinby, J.R., and R.E. Karper. 1945. The inheritance of three genes that influence time of floral initiation and maturity date in Milo. *J. Am. Soc. Agron* 37:916-936
- Ritter, K.B., D.R. Jordan, S.C. Chapman, I.D. Godwin, E.S. Mace, and C.L. McIntyre. 2008. Identification of QTL for sugar-related traits in sweet x grain sorghum (*Sorghum bicolor* (L.) Moench) recombinant inbred population. *Mol. Breed* 22:367-384

- Robertson, M.J., R.C. Muchow, A.W. Wood, and J.A. Campbell. 1996. Accumulation of reducing sugars by sugarcane: effects of crop age, nitrogen supply and cultivar. *Field Crop Research* 49:39-50
- Tarpley, L., and D.M. Vietor. 2007. Compartmentation of sucrose during radial transfer in mature sorghum culm. *BMC Plant Biol* 7:33
- Viola, K., and H.V. Davies. 1992. A microplate reader assay for rapid enzymatic quantification of sugars in potato tubers. *Potato research* 35:55-58
- Yang, R.-C. 2002. Likelihood-based analysis of genotype-environment interactions. *Crop Sci* 42:1434-1440
- Yang, R.-C. 2007. Mixed-model analysis of crossover genotype-environment interactions. *Crop Sci* 47:1051-1062
- Wenzel, W.G. 1990. Inheritance and interrelationship of quantitative traits in sorghum. *Sorghum newsletter* 31:1-3

Chapter 3

Genetic and heritability estimates of fibre-related traits in Grain x Sweet sorghum recombinant inbred lines.

Amukelani Shiringani and Wolfgang Friedt

Abstract: There is a growing interest of using *Sorghum bicolor* (L) Moench as a dedicated bioenergy crop, where cell wall polysaccharides can be hydrolysed with hemicelluloses and celluloses degrading enzymes to produce sugar. Recently, sorghum breeders are aiming at increasing fibre quality and quantity to meet the global demand of fuel. The objectives of this study were to estimate the genetic parameters and test the significance of genotype by environment interactions affecting fibre and related agronomic traits, to estimate the heritability and also the relationship among these traits. A grain x sweet sorghum RIL population segregating for analysis fibre-related traits was planted across 4 environments for fibre and related agronomic traits. Phenotypic data was analysed using Mixed model of SAS[®] 9.0 version. Analysis of the variance showed that genotypes and environments were significantly different in all the traits except for dry stalk weight. A significant genotype by environment interaction was observed. REML showed that genetic variance was a higher significant contributor in fibre traits than in agronomic traits. Environments were not a significant contributor in all traits although the contribution in some traits was high. There was a significant genotype by environment interactions in all the traits. Heritability estimates were moderate to high in all traits except in dry biomass and dry stalk mass. These traits are dependant in sampling and can results in dependent error. The correlation coefficients were positively significant among fibre traits and among agronomic traits but were negatively correlated among agronomic and fibre traits.

3.1 Introduction

Lignocellulosic biomass such as stover from maize and sorghum offers an abundant and inexpensive source of fermentable sugars. Carpita and McCann (2008) reported that biomass quality is heavily influenced by cell wall composition and structure and can be determined by content and composition of lignin, cellulose and hemicellulose and how they are cross-linked, whereas biomass yield can be determined by agronomic traits such as plant height and diameter, diseases and pest resistance, and lodging susceptibility.

Sugars obtained from cell wall polysaccharides (cellulose and hemicellulose) can be hydrolysed with hemicellulose or cellulose degrading enzymes; a process known as saccharification (Lin and Tanaka, 2006). The monosaccharides obtained are fed to microorganisms in fermenters similar to the systems used in starch- or sugarcane-derived sugars. Although the energy balance of the stover-to-ethanol process has been debated, technical and economic analysis showed that the production of ethanol from lignocellulosics could result in a net gain. To make this technology economically competitive, cost effective and efficient, means of transportation and storage of biomass yield, and processing strategy should be developed. Besides, it is mandatory to understand the physical and chemical properties, biosynthesis and molecular genomics of plant cell walls and also how to manipulate them in the candidate crops.

Cosgrove (2000) reported that plant cells grow by expanding their cell walls through a process of controlled polymer creep (slow, time-dependent, irreversible extension, in which the microfibrils and associated matrix polysaccharides slowly slide within the wall, therefore increasing its surface area). The growth is accomplished by the synthesis and integration of new polysaccharides into the wall and the expansion of the polysaccharides network. The primary cell wall is synthesized during expansion growth (Carpita, 1996) and the secondary cell wall begins in the primary cell wall when the cell is ceasing to expand (Dhugga, 2007). The primary cell wall contains mainly cellulose and hemicelluloses, pectins as well as many enzymes and structural proteins, whereas the secondary cell walls is rich in cellulose and also contain lignin and little protein or pectin. The main structural carbohydrates by content in the cell wall of sorghum are cellulose, hemicellulose and lignin (Anglani, 1998).

Lignin does not exist in plant tissue as an independent polymer but it is bonded with other polymers such as cellulose and hemicelluloses forming lignin-carbohydrates complexes (LCC). The LCC from herbaceous crops are structurally different from those in woods and contains ferulic bridges between lignin and carbohydrates (arabinoxylans) via ester-linked ferulic acids (Grabber et al., 2004). The complex heterogeneous structure of the plant cell wall can influence the ability of cellulolytic enzyme complexes to digest plant biomasses to fermentable sugars (Carpita and McCann, 2008). Lignin in particular appears to be very significant: It prohibits cellulose fibre swelling, which reduces surface area access to the enzyme, and cellulases irreversibly adsorb to lignin, thus preventing their action on the

cellulose. This necessitates the use of more enzymes to saturate these non-productive adsorption sites on the surface of biomass and lead to high enzyme cost.

As researchers are collecting information describing the physiological and biochemical processing related to fibre component, quality and quantity, the mechanisms by which the environment interact with these processes is not simple or well documented in sorghum. Improving lignocellulosic biomass value depends on successful selection for traits that influence ethanol yield, thus the right proportion of cellulose-hemicellulose to lignin and related agronomic traits. As it is already known, differences in environmental conditions among years can mask any genetic effects, thus hinder improvement. Quantifying the genotypic and environmental variation in sorghum stems cell wall polysaccharides would give information on the possibility to alter cell wall components concentration through breeding. Besides, it will also give knowledge on heritability of the traits. The objectives of this study were to 1) estimate the genetic parameters and test the significance of G-E interactions of fibre-related traits in sorghum 2) estimate the H^2 of these traits 3) and to estimate the relationship among these traits.

3.2 Materials and methods

1. Field experiment

Information about the plant material and field experiments layout is presented in chapter 2.

2. Data collection

Similar to chapter 2, the trials were harvested manually by cutting the plant at the base by a scissor. In Gross Gerau, three middle plants were selected at random for all data collection and latter the average of the three plants was used. However, in the pot trials both plants per plot were harvested and averaged. Some lines were photoperiod sensitive, and resulted in differences in flowering dates of up to 2 months. To harvest the plots at the same physiological stage, each line was harvested from 40 days after anthesis (DAA), which estimated the hard dough stage. Everything above ground was weighed to estimate the fresh biomass. The leaves and panicle of the plants were stripped from the stalks. Total fresh leaf weight was recorded per plant followed by weighing the stripped stalk to obtain the stripped stalk mass. Panicle, leaves and stripped stalk mass were milled and then dried in an oven at

105 ° C for two to three days. After drying, the samples were weighed and their sum was used to estimate the dry stripped stalk mass and the dry biomass.

3. Fibre content analysis

Following the harvesting procedure above, the grounded-dried stripped stalks were further milled to pass 0.5 or 1 mm sieve in Gross Gerau (GG) or Rauischholzhausen (RH) for fibre content analysis. From the milled sample, 130 genotypes from GG and RH were selected due to extreme agronomical traits differences to quantify acidic detergent fibre (ADF), neutral detergent fibre (NDF) and acidic detergent lignin (ADL). The procedure was done according to van Soest et al. (1991) with some small modifications. The NDF measures cellulose, hemicellulose and lignin, while ADF measures cellulose and lignin, and ADL only measures lignin. Therefore, NDF minus ADF yields hemicellulose content, and ADF minus ADL yields cellulose.

Approximately 1.5 g of the milled sample were scanned in a near infrared reflectance spectroscopy system (NIRSystem Foss model, Silver spring, MD), using a 3.7 cm diameter sample holder over a wavelength of 400-2500 nm with 2 nm interval. The selected 130 genotypes were split into two subsets: calibration set, in which the experimental values were used to develop the best calibration equations through WinISI II software (Infrasoft international, LLC, PA), and the validation set, which was used to test the ability of the calibration equations to predict the trait value. The best equation should have low standard error of calibration (SEC), standard error of cross validation (SECV), systematic difference between the two sets, but higher coefficient of determination in the calibration (R^2). After obtaining the best equations, a total of 213 genotypes x 4 environments were scanned to quantify the fibres of interest.

4. Statistical analysis

Combined phenotypic analysis for all traits was done by the mixed model (MIXED) using SAS[®] 9.1 software. Unlike the General Linear model (GLM) method that uses the least square to fit the general linear models of means, the MIXED fits the model based on means, variances and covariances. Restricted maximum likelihood (REML) and TYPE III method were used for all the analysis and the best method was selected by Bayesian and Akaike's information criterion, which measures the goodness of the model fit. In this case, the method with lower values of AIC and BIC is the best model, because it gives the correct balance

between the fit to the data and model complexity. TYPE III method was used to analyse the variance while REML method was used to analyse the covariance effects and their significance. For analysis of variance and to estimate the covariate effects, all parameter: genetic, harvesting date, environment and genotype x environmental interactions were treated as random effects. The significance of the covariates was obtained by entering the COVTEST option which provides the estimates of the standard errors of the variance components, Z-scores, and the associated p values. The output of the REML method was further used to calculate broad-sense heritability calculation as follows:

$$H^2 = \sigma^2_G / (\sigma^2_G + \sigma^2_{GXE}/E + \sigma^2_{error}/ER)$$

where σ^2_G is genotypic variance, σ^2_{GXE} is the variance of genotype x environment interaction, σ^2_{error} is the error variance, E is number of environments and R is the number of replications.

The REML procedure was repeated where genotypes were treated as fixed effects to obtain simple statistics results and the corrected trait means. This was done with least significance mean commands. The corrected trait means was further used to estimates the correlation effects among traits using the Proc Corr command.

3.3 Results and discussion

3.3.1 Phenotypic evaluations

The average performance of fibre traits and related agronomic traits is presented in table 3.1. Differences in environmental conditions between years showed a potential to alter the plant's performance. For example, in 2007, RILs performance in RH decreased by more than 50% than that of GG in terms of fresh biomass, stripped stalk mass, fresh leaf mass, and dry stalk weight. This is not surprising because the reduction in stem juice was observed in most lines during plant development. Still on the same traits, SS79 weighed more than M71 in all environments. In terms of dry biomass and dry stalk weight, SS79 reduced its weight tremendously compared to its fresh biomass and stripped stalk mass, showing the high amount of juice within its stalks. At most 30% transgressive segregation of the RILs was observed in fresh biomass, dry biomass, stripped stalk mass, fresh leaf mass, and dry stalk weight.

The male parent M71 had more fibre per dry weight compared to the female parent SS79 in all environments. Although not verified, it was expected because of the higher sugars found in the sweet sorghum parents. Cellulose was the abundant fibres followed by hemicelluloses and

Lignin respectively. The average combined content cellulose, hemicelluloses, ADL and NDF, was 26.33 g/dry weight, 18.42 g/dry weight, 7.17 g/dry weight, and 52.05 g/dry weight respectively. The average content of each fibre trait showed consistency across environments. Amaducci et al. (2004) reported consistent lignin content in sorghum across growing conditions. However, parents had more fibre content in RH compare to GG.

3.3.2 Analysis of variance and its contribution effects

Table 3.2 shows the results of the degrees of freedom and the mean squares obtained from the TYPE III mixed model. Genotypes were significantly different among each other in most of the traits studied. Actually TYPE III showed all the source of variation to be significant in most traits. Even though all source of variation showed to be highly significant to ADF and dry biomass, none showed to be significant to their components, hemicelluloses and dry stalk mass. Zerbini and Thomas (2003) reported that there is limited information on the significance of genetic variation on fibre contents especially NDF and ADF. Besides, there is no information on how much of the variation can be attributed to genotype, environment, or genotype x environment interactions. In this study we estimated how much is the contribution of the genotype, environment and genotype x environment interactions to the traits using restrictions maximum likelihood (table 3.3). Environment showed to be the major contributor in terms of percentage to ADL (67.9%), fresh leaf mass (57.9%), stripped stalk mass (59.1%), fresh biomass (52.8%) and dry biomass (81.5%). However, in all the traits studied, environment showed no significant impact. Murray et al. (2008) found environmental variation to have a significant impact on the stem fibre related traits but not in some leaf fibre-related traits. There is limited information in this topic to conclude in the findings. Genotypic effect was observed in all the traits and higher contribution was observed in cellulose (32.4%), ADF (31.1%), NDF (17.5%) and stripped stalk mass (12.4%). The higher genetic variation indicates that it is feasible to breed and select for high or low fibre content in sorghum. The lowest non-significant genetic effects were observed in dry stalk weight and dry biomass, this might be a multi-environmental error. There was a significant interaction between genotype and environment, and the highest was observed in ADF (13.9%) and cellulose (15.8%). Harvest dates showed to be a significant contributor to most of the traits. This results support the findings of Jung and Castler (2006) in hybrid maize that sampling dates is significant for all cell wall traits. In this study harvest dates varied in genotypes, however there were no correlations done between late or early harvest dates to fibre yield.

3.3.3 Heritability

The results of broad-sense heritability are presented in table 3.3. Genetic variation was more important in explaining the fibre content. Heritability estimates were found to be higher in ADF (84%), cellulose (82%), NDF (70%), and ADL (67%). Several studies in forage grasses reported moderate to high heritability estimates in ADF and NDF (Jensen et al., 2006, Ray et al., 2006). There are literature references for genetic variation and the measurements of heritability for fibre-related traits. However, most of the information is in temperate grasses. Fresh leaf mass and fresh biomass showed moderate heritability estimates of 51% and 34%, respectively. Heritability estimates for dry stalk weight was found to be non-existent or conflicting, precluding efficient selection against this traits. However, care should be taken when milling the stalks for dry matter analysis, which could have promoted the error.

3.3.4 Correlation

Most of the fibre-related traits studied correlated significantly with each other (table 3.4). The fibre quality traits (ADF, NDF, ADL, cellulose and hemicelluloses) showed a strong and positive reflection on each other. The strongest correlations was observed between ADF and cellulose ($r = 0.881$), NDF and hemicelluloses ($r = 0.841$), and ADF and NDF ($r = 0.667$). Marvin et al. (1995) and Cardinal et al. (2003) working on maize recombinant inbred lines reported related moderate to high correlation estimates on the same fibre quality. In this study however, negative and non-significant correlation was observed between cellulose and ADL ($r = -0.014$). Agronomic traits (fresh leaf mass, stripped stalk mass, fresh biomass, and dry biomass) also showed positive and strong correlation among each other. The strongest correlation was observed between stripped stalk mass and fresh biomass ($r = 0.709$), stripped stalked mass and dry biomass ($r = 0.625$), and fresh biomass and fresh leaf mass ($r=0.607$). This was expected because of the sampling method. However, the poor reflection of stripped stalk mass and dry stalk mass was not expected. Dry stalk mass showed poor and non-significant correlation with all the characteristics except for fresh leaf mass ($r = 0.316$). Fibre quality correlated negatively and significantly with agronomic traits. The strongest correlations were observed between ADF and ADL with agronomic traits and the weakest correlations was observed between hemicelluloses and agronomic traits. The strong and moderate correlation between fibre traits and agronomic characters indicates that these traits can be used for indirect selection.

3.4 Conclusion

The nature of the genetic variation in various characteristics and its genetic relationship are important for planning a successful breeding program. This information can be expressed in terms of heritability, genetic and environmental correlations, genotype x environment interactions (Aastiveit and Aastiveit, 1990). In this study, biomass yield showed to be more influenced by the environment than biomass quality, although the effect was not significant. However, in both biomass yield and quality a strong significant G x E was observed. Moderate and high heritability estimates were found in most traits, which give the evidence that these traits can be improved through plant breeding. Besides, good correlation coefficients were observed among most of the traits. This will facilitate multiple traits selection for lignocellulosic biomass in sorghum.

3.5 Acknowledgement

The authors would like to thank Mario Tolksdorf, Markus Kollmer, and Malte Luh for field technical assistance and Nelly Weiss, Swetlana Renner, and Adriana Ochoa Fandiño for fibre-content analysis. This work was financially supported by the Deutsche Akademische Austausch Dienst.

Table 3.1: Performance of parents and RILs for fibre-related traits across environments

Trait	Gross Gerau										Rauischolzhausen									
	2007					2008					2007					2008				
	SS79	M71	RIL mean	SD	Range	SS79	M71	RIL mean	SD	Range	SS79	M71	RIL mean	SD	range	SS79	M71	RIL mean	SD	range
ADF g/dry weight	28.05	31.05	31.23	3.00	24.40-42.65	30.45	39.95	35.20	9.53	26.60-45.43	31.36	36.40	34.19	3.72	24.55-41.28	29.02	37.02	34.01	2.69	26.18-41.65
ADL g/dry weight	3.80	4.73	4.99	0.90	2.00-7.45	6.07	8.42	7.22	0.97	2.00	6.01	8.53	7.73	1.27	4.85-11.40	6.87	8.60	8.72	1.07	5.55-10.9
NDF g/dry weight	44.97	50.97	49.71	10.28	39.50-63.43	45.88	59.67	52.23	4.90	41.37-68.00	48.89	61.20	52.72	5.35	40.80-66.00	46.56	60.10	53.50	2.92	46.07-61.38
Cellulose g/dry weight	24.25	26.32	26.24	2.41	19.75-36.25	24.38	31.56	27.48	3.34	19.13-39.25	25.35	29.03	26.45	3.15	18.25-36.5	22.15	30.00	25.15	2.98	25.61-35.7
HC g/dry weight	16.55	19.92	18.12	1.40	14.55-25.62	15.43	19.69	17.52	1.37	13.87-22.10	16.63	19.76	18.53	2.00	13.05-26.00	17.51	20.23	19.52	1.10	15.6-25
FLM g/plant	159.44	70.72	168.77	46.04	46.47-421.59	142.89	68.59	136.85	66.06	18.33-853.33	65.1	40.23	58.40	19.66	8.02-181.97	NA	NA	NA	NA	NA
SSM g/plant	927.11	262.62	550.01	259.25	187.1-1603	1265.28	277.72	623.31	300.00	131.97-2157	247.78	120.3	249.24	106.17	9.45-752.09	129.98	105.63	85.37	33.80	7.14-192.54
DSM g/plant	190.72	28.11	113.50	58.93	22.56-359.88	23.27	21.97	22.79	2.93	7.97-45.38	53.15	40.60	55.28	26.53	4.32-192.21	27.46	50.42	17.85	7.94	1.08-41.64
FBM g/plant	1499.00	683.18	761.88	296.81	166.83-2027.00	1488.17	504.57	856.70	329.12	303.67-2867.00	321.06	230.42	328.03	122.26	17.47-903.06	NA	NA	NA	NA	NA
DBM g/plant	268.63	71.47	197.33	71.82	64.54-496.45	87.24	90.56	87.74	9.31	56.27-118.32	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

SD: standard deviation ADF: acidic detergent fibres ADL: acidic detergent lignin NDF: Neutral detergent lignin HC: Hemicellulose FLM: fresh leaf mass SSM: striped stalk mass DSM: dry stalk mass FBM: fresh biomass DBM: dry biomass NA: not assayed

Table 3.2: Mean squares of fibre-related traits analysed by TYPE III mixed model

Source of variation	ADF g/dry weight		ADL g/dry weight		NDF g/dry weight		Cellulose g/dry weight		HC g/dry weight		FLM g/plant		SSM g/plant		DSM g/plant		FBM g/plant		DBM g/plant	
	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS	DF	MS
G	213	39.36* **	213	3.27** *	213	89.43* **	213	29.60* **	213	25.59	213	3261.1 1***	213	12912 8***	213	4984.7 2	213	11076 2**	211	3089.0 1***
E	3	236.90 ***	3	209.97 ***	3	232.66 ***	3	62.30* **	3	48.63	2	58086 ***	3	46226 95***	3	1961.8 2	2	14691 72***	2	72996 3***
Harvest date	44	19.12* **	44	2.04** *	44	44.68* **	44	11.50* **	44	7.99	30	362.77	44	2149.5 1	44	132.92	31	6953.2 1	21	134.91
G x E	575	9.50** *	575	1.40** *	575	39.37* **	575	7.54** *	575	24.83	372	2228.0 5**	574	39123 ***	575	73638 3	373	86108 ***	366	3060.2 1***
Residual	1528	4.22	1528	0.54	1528	30.57	1528	3.30	1528	25.47	1185	1769.9 2	1522	17542	1517	70044	1201	48049	1148	906.51 9

DF: degree of freedom MS: Means squares G: genotype E: environment G X E: genotype x environment ADF: acidic detergent fibres ADL: acidic detergent lignin NDF: Neutral detergent lignin HC: hemicelluloses FLM: fresh leaf mass SSM: striped stalk mass DSM: dry stalk mass FBM: fresh biomass DBM: dry biomass * significant at the 0.05 probability level ** significant at the 0.01 probability level *** significant at the 0.001 probability level

Table 3.3: Variance components and Heritability estimates for fibre-related traits

	H ²	Variance components				
		Genetic	Environment	Harvest date	G X E	Residual
ADF g/dry weight	0.84	0.311***	0.134	0.105***	0.139***	0.311***
NDF g/dry weight	0.70	0.175***	0.054	0.044***	0.079***	0.648***
ADL g/dry weight	0.67	0.066***	0.679	0.026***	0.083***	0.146***
Cellulose g/dry weight	0.82	0.324***	0.079	0.089***	0.158***	0.350***
HC g/dry weight	0.24	0.025*	0.024	0.007	0.018	0.926***
FLM g/plant	0.51	0.041***	0.579	0.010	0.052***	0.317***
SSM g/plant	0.79	0.124***	0.591	0.043**	0.080***	0.161***
DSM g/plant	0.0005	0.0004	0.0009	0	0.0006	0.998***
FBM g/plant	0.34	0.028**	0.528	0.019*	0.111***	0.313***
DBM g/plant	0.13	0.004	0.815	0.008*	0.084***	0.089***

ADF: acidic detergent fibres ADL: acidic detergent lignin NDF: Neutral detergent lignin HC: hemicelluloses FLM: fresh leaf mass SSM: striped stalk mass DSM: dry stalk mass FBM: fresh biomass : dry biomass * significant at the 0.05 probability level ** significant at the 0.01 probability level *** significant at the 0.001 probability level

Table 3.4: Pearson correlation estimates of the fibre-related traits

	ADF	NDF	ADL	Cellulose	HC	FLM	SSM	DSM	FBM
NDF	0.667***								
ADL	0.559***	0.398***							
Cellulose	0.881***	0.574***	-0.014						
HC	0.159***	0.841***	0.122***	0.121***					
FLM	-0.242***	-0.166***	-0.451***	-0.072**	-0.052*				
SSM	-0.348***	-0.324***	-0.520***	-0.120***	-0.177***	0.578***			
DSM	-0.039	-0.024	-0.018	-0.037	-0.003	0.316***	0.013		
FBM	-0.263***	-0.188***	-0.349***	-0.153***	-0.065**	0.607***	0.709***	0.040	
DBM	-0.387***	-0.270***	-0.746***	-0.012	-0.09***	0.340***	0.625***	0.033	-0.128***

ADF acidic detergent fibres ADL acidic detergent lignin NDF Neutral detergent lignin HC Hemicellulose FLM fresh leaf mass SSM striped stalk mass DSM dry stalk mass FBM fresh biomass DBM dry biomass * significant at the 0.05 probability level ** significant at the 0.01 probability level *** significant at the 0.001 probability level

3.6 References

- Anglani, C. 1998. Sorghum carbohydrates-A review. *Plant foods for human nutrition* 52:77-83
- Aastiveit, A.H., and K. Aastiveit. 1990. Theory and application of open pollinated and polycross in forage grass. *Theor. Appl. Genet* 79:618-324
- Amaducci, S., A. Monti, and G. Venturi. 2004. Non-structural carbohydrates and fibre components in sweet and fibre sorghum as affected by low and normal input techniques. *Ind. Crops Prod* 20:111-118
- Cardinal, A.J., M. Lee, and K.J. Moore. 2003. Genetic mapping and analysis of quantitative trait loci affecting fibre and lignin content in maize. *Theor. Appl. Genet* 106:866-874
- Carpita, N.C. 1996. Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol* 47:445-476
- Carpita, N.C., and M.C. McCann. 2008. Maize and sorghum: genetic resources for bioenergy grasses. *Trends in Plant Science* 13: 415-420
- Crosgrove, D.J. 2000. Expansion growth of plant cell walls. *Plant physiol. Biochem* 38:109-124
- Dhugga, K.S. 2007. Maize biomass yield and composition of biofuel. *Crop Sci* 47:2211-2227
- Grabber, J.H., J. Ralph, C. Lapierre, and Y. Barrierè. 2004. Genetic and molecular basis of grass cell-wall degradability. I. Lignin-cell wall matrix interactions *C.R. Biologies* 327:455-465
- Jung, H.G., and M.D. Castler. 2006. Maize stem tissues: Cell wall concentration and composition during development. *Crop Sci* 46:1793-1800
- Jensen, K.B., J.G. Robins, B.L. Waldron, and M.D. Peel. 2006. Genetic variation in dry matter production and nutritional characteristics of meadow bromegrass under repeated defoliation. *Crop Sci* 46:1948-1954
- Lin, Y., and S. Tanaka. 2006. Ethanol fermentation from biomass resources: current state and prospects. *Appl. Microbiol. Biotechnol* 69: 627-642
- Marvin, H.J.P., C.F. Krechting, E.N. van Loo, C.H.A. Snijders and O. Dolstra. 1995. Relationship between stalk cell wall digestibility and fibre composition in maize. *J.Sci food Agric* 69:215-221
- Murray, S.C., W.L. Rooney, S.E. Mitchel, A. Sharma, P.E. Klein, J.E. Mullet, and S. Kresovich. 2008. Genetic improvement of sorghum as biofuel feedstock. II. QTL for stem and leaf non-structural carbohydrates. *Crop Sci* 48:2181-2193

- Ray, I.M., J.F. Karn, and S.T. Dara. 2006. Heritabilities of nutritive quality factors and interrelations with yield in tetraploid crested wheatgrass. *Crop Sci* 36:1488-1491
- Van Soest, P.J., J.B. Robertson, and B.A. Lewis. 1991. Methods for dietary fibre, neutral detergent fibre, and nonstarch polysaccharides in relation to animal nutrition. *J. Dairy Sci* 74: 3583-3597
- Zerbini, E., and D. Thomas. 2003. Opportunities for improvement of nutritive value in sorghum and pearl millet residues in South Asia through genetic enhancement. *Field Crops Research* 84:3-15

Chapter 4

Genetic mapping of sugar-related traits in a RILs population of *Sorghum bicolor* L. Moench.

Amukelani Shiringani & Wolfgang Friedt

Abstract: Genetic mapping and characterization of quantitative traits loci is the first and major step in crop improvement. Unlike Mendelian traits that are controlled by individual genes, the phenotypic variations of complex traits results from segregation of alleles at multiple chromosomal regions that can be attributed to epistasis, QTL x environment and more complex higher order interactions. Sugar components and sugar related agronomic traits were quantified at four different environments using a population of 213 recombinant inbred lines. The phenotypic results per trait were averaged to obtain the combined environmental mean and analysed by composite interval mapping (CIM) of PLABQTL software at LOD score >2.5 and the identified QTLs were tested for significance in each environment and for QTL x environment interaction. The model AA was used to obtain QTL x QTL interaction. CIM identified more than five additive QTLs in most traits explaining a range of 6.0 to 26.1% phenotypic variation. Digenic epistasis seems to be abundant through out the genome because 24 pairs were identified in seven traits; supporting that QTL analysis without epistasis analysis can be biased. QTLs showing multiple effects were identified, where the QTL in SBI-06 was significantly associated with almost all the traits. A total of four out of ten traits studied showed a significant QTL x environmental interaction. The stability of the QTLs across environments is important for marker-assisted selection.

4.1 Introduction

Sorghum bicolor is a diploid species ($2n=20$) that is closely related to sugar cane, a polyploid species. The cultivated sorghum belongs to the genus *sorghum* of the family *Poaceae* as sugarcane does. The sweet sorghum type stores high amount of sugar in its stem, and produces high amount of juice. Sweet sorghum stalks have been processed to syrup, sugar and molasses for years. Recently, Sorghum breeders worldwide are aiming at improving various characteristics of the crop that can results in commercialization. Although for many years the main trait of interest in sorghum has been grain yield, recently interest has also shifted from grain to stalk to manipulate biofuel-related traits: fibres, sugars, juice and biomass. As in grain

yield, biofuel-related traits cannot be manipulated directly, because are quantitatively inherited traits. Therefore to successfully improve these traits, the crop should be dissected into smaller morphological, physiological and genetical components and then analyse the information acquired from this components.

Sorghum like sugarcane is a C₄ plant characterised with high photosynthetic efficiency. Sugar yield is the major breeding objective in sweet sorghum. Aitken et al. (2006) defined sugar yield as the function of cane yield and recoverable sucrose content from the harvested cane. Milligan et al. (1990) reported that genetic correlations of sugar traits are a better strategy for development of selection in sugarcane. After screening a population in 2 years at five locations, they found that stalk number has high correlation with cane yield ($r = 0.77$) and concluded that selection for sucrose should emphasize cane yield and attention can be given to stalk numbers. A high correlation between stalk weight and stalk diameter ($r = 0.70$) was also observed and brix was highly correlated to sucrose. In sugarcane, stalk diameter and stalk number are highly heritable with more than 80% reported (Aitken et al., 2008) and a number of QTLs influencing these traits has been reported (Ming et al., 2002). In sorghum, Murray et al. (2008) reported a weak but significant correlation between sugar in dry stem and stem diameter ($r = 0.11$) and tillers ($r = -0.09$). Total sugar yield had no significant correlation with diameter but significant correlation with tillers ($r = 0.16$), while juice sucrose had no correlations with tillers number but had a negative, weak but significant correlation with stem diameter ($r = -0.06$). These weak genetic correlations make indirect selections for sugar traits or sucrose using morphological traits less successful.

Ritter et al. (2008) reported QTLs for sugar-related traits in a population of 186 recombinant inbred lines derived from grain x sweet sorghum. They used 247 AFLP, sorghum and sugarcane SSR markers. Using composite interval mapping with a LOD score of ≥ 3.00 , they detected 11 markers, 8 markers, and 10 markers distributed in 5 different linkage groups for sucrose content, sucrose yield and sugar yield respectively. Natoli et al. (2002) reported several QTLs of sugar related traits in a population of 129 F₂ progenies derived from unrelated sweet sorghum parents. Yun-long et al. (2006) detected several QTLs in 207 F₂ individuals derived from a cross between grain and sweet sorghum. Natoli et al., (2002) detected QTLs on SBI-01, SBI-03, and SBI-05, Yun-Long et al. (2006) detected QTLs on SBI-04, SBI-10, Ritter et al. (2008) and Murray et al. (2008) detected QTLs for sugar components on SBI-01, SBI-03, SBI-05, SBI-06, SBI-07, SBI-10, but with extra SBI-02 for

Murray et al. (2008). The distribution of these QTLs among chromosome and the less phenotypic variation explained suggest that sugar traits are controlled by more loci with complicated interactions. The objectives of the study were to 1) identify additive QTLs linked to sugar related traits in sorghum 2) detect possible additive x additive interaction among these QTLs 3) identify QTLs showing pleiotropic effect.

4.2 Materials and methods

4.2.1 Mapping population

The population of Recombinant inbred lines (RIL) derived from a cross between grain sorghum (M71) and sweet sorghum (SS79) discussed in chapter two was used for this study. Details about experimental trials and data collection have been discussed in chapter two. The data was analyzed using SAS[®] 9.0 software. Descriptive statistics was analysed using the LSmeans command and the analysis of the variance was done using the TYPE III method of the general linear model.

4.2.2 Genotyping

4.2.2.1 DNA extraction

The same RIL population was planted in small pots in the green house in three replications with two pots per replication. Fresh leaf samples were collected 20 days after planting from replication one and three as A and B samples respectively. Tissues were placed in eppendorf tubes and frozen immediately in liquid nitrogen and then stored in -80 °C freezer. DNA extraction was performed according to Doyle and Doyle (1990).

4.2.2.2 SSR markers amplification

A total of 200 SSR primer pairs reported by Bhatramakki et al. (2000), Dean et al. (1999), Kong et al. (2000), Taramino et al. (1997), and Schloss et al. (2002) were used to screen parents of the mapping population. Few EST-SSRs published by Srinivas et al. (2008; 2009) were also included. The reaction protocol and thermal cycler touch-down protocol were carried according to Hasan et al. (2006) and Xu et al. (2005). Most of the SSR primers were labelled with M13 tailing as describe by Berg and Olaisen (1994). The fluorescently labelled universal M13 primer 5'-AGGGTTTTCCAGTCACGACGTT-3' is added to the PCR reaction, and the forward primer of each SSR is appended with the sequence 5'-

TTTCCCAGTCACGACGTT-3'. After the first cycle of PCR, fragments are successively amplified with the labelled primer (Hasan et al., 2008; Rygulla et al., 2007).

4.2.2.3 AFLP markers amplification

AFLP markers were amplified according to Vos et al. (1995) using Gibco AFLP core reagents. A total of 150 ng (6 µl x 25 ng) genomic DNA was restricted using 2 µl EcoRI/MseI in 5 µl x 5x reaction buffer and 12 µl AFLP-grade water. The solution was incubated for 2 hrs in 37 °C followed by 15 min in 70 °C in Perkin Elmer 9700 thermal cycler. After incubation, a ligation master mix consisting of 24 µl adapter ligation solutions and 1 µl T4 DNA ligase was added to the solution, which was again incubated for 2 hrs at 20 °C. Ligation was followed by a pre-amplification step (+0) using the 5 µl of 1:10 diluted ligation with non-selective primers in order to minimize background noise at electrophoresis and to receive more distinct bands. The +0 sequence is complementary to the adapters and to the enzyme cutting-sequence. The following +1 pre-amplification was carried out using 5 µl of 1:10 dilutions of +0 pre-amplification prior to amplification (+3) using selective primer combinations. The sequence of the primers +0 EcoRI, +0 MseI, +1 EcoRI, and +1 MseI were 5'-GAC TGC GTA CCA ATT C-3', 5'-GAT GAG TCC TGA GTA A-3', 5'-GAC TGC GTA CCA ATT CA-3', and 5'-GAT GAG TCC TGA GTA AC-3 respectively as reported by Uptmoor et al. (2003).

4.2.2.4 Electrophoresis

A total of 10 µl fusion dyes was added to the PCR products, the samples were denatured at 94 °C for 5 min and then 2 µl samples was loaded in 8% polyacrylamide gel. Amplification products were separated and visualized using LI-COR model 4200 (MWG Biotech, Ebersberg, Germany) automated DNA sequencing system. This system can detect amplification products that were labelled with two different dyes: IRD-700 nm and IRD-800 NM.

4.2.3 Linkage map construction

A total of 215 scored markers and 213 RILs were used to construct linkage maps using the Joinmap® 4.0 software. The map did not include 59 markers that were excluded because of missing values or divergence from the expected segregation ratio using a chi-square test with one degree of freedom at $\alpha = 0,05$. The tentative linkage groups were grouped with the minimum LOD score greater than 3.0 and the Kosambi mapping function was used to

calculate map distances in cM from the recombination frequencies. The linkage maps were compared to sorghum map published by Bhatramakki et al. (2000) and Mace et al. (2009), and assigned according to a recent sorghum chromosome nomenclature published by Kim et al. (2005).

4.2.4 QTL analysis

QTLs were detected from the combined environmental means using the composite interval mapping (Utz and Melchinger, 1994) of the PLABQTL software. Due to missing values, only 188 lines and 157 markers comprising of SSR and AFLPs markers were used for the analysis. A model 'AA' was used because no reliable dominance effect is expected in the lines after F₅ in self pollinating crops. A LOD score of 3.0 was used to detect QTLs. A QTL was considered suggestive if detected between a LOD score of 2.5 and 3.0 and had significantly influenced the phenotype in a certain environments. When two peaks in the same chromosome were detected for a single trait, they were considered as two QTLs only if the distance between the QTLs was greater than 20 cM otherwise only the higher peak was considered for better estimation of the QTL position (Parth et al., 2008; Ungerer et al., 2002).

4.3 Results

4.3.1 Trait performance

The phenotypic performance of parents and recombinant inbred lines is presented in table 4.1. The parents performed as expected for sugar related traits, where SS79 was taller than M71 in all locations. Besides, SS79 was late flowering, had high sugar content, and low fresh panicle weight. Trait means of the recombinant inbred lines were intermediate, although trans-segregation was observed.

According to the general linear model, genotypes were significantly different ($P > 0.0001$) in all the traits studied except for estimated juice weight. Although the estimated juice weight showed segregation among the RILs, both parents are known to have juicy stems. In this study, SS79 showed to be juicier than M71. Environments were significantly different in all traits and high genotype by environment interaction was observed in all traits except for estimated stem juice weight. Harvest dates showed no significant different in all traits except for flowering dates. Most of the lines were photoperiod sensitive resulting in differences in

harvest dates, therefore harvest dates was used to estimate the same physiological state of the lines. Harvesting was done 40 days after anthesis for each line.

There were acceptable repeatability values in most of the traits ranging from 0.17 to 0.81 and the heritability values ranging from 0.18 to 0.81. Juice showed a lower repeatability ($r^2=0.17$) values and heritability values ($r^2=0.18$), showing the impact of the environment in this trait. Although the heritability of flowering dates was low ($H^2 = 0.33$), its repeatability value was high ($r^2 = 0.75$). The moderate to high heritability in these sugar-related traits shows that breeding for these traits is possible.

4.3.2 QTL detection

Figure 4.1 shows the results of the identified significant QTLs either in each environment or from the mean of the combined environments or both. Table 4.3 of the supplementary results shows the details of the identified additive QTLs. The QTLs that significantly affected the traits using combined environmental means or mean of each single environment are reported as significant. Environments 01, 02, 03, and 04 throughout the chapter represent the field experiment in Gross Gerau 2007, Gross Gerau 2008, Rauschholzhausen 2007, and Rauschholzhausen 2008, respectively.

4.3.2.1 Flowering dates

CIM detected four QTLs and one putative QTL on SBI-04a, SBI-06, SBI-07 and SBI-08 associated with flowering dates at LOD score ≥ 2.5 . The QTLs on SBI-03, SBI-04a and SBI-06 were found significant in 4 environments and also in the overall environment. The phenotypic variance explained by these significant QTLs was 33.40% and their additive effects were 3.90, 0.62 and -2.72 days on SBI-03, SBI-04a and SBI-06, respectively. Other studies have detected several QTLs on SBI-01, SBI-04, SBI-06, and SBI-10 (Ritter et al., 2008), SBI-06, SBI-09 (Murray et al., 2008), and SBI-01, SBI-03, SBI-06, and SBI-08 (Feltus et al., 2006). Chantreau et al. (2004) detected few QTLs controlling photoperiod respond in linkage group C, F and H that correspond to SBI-01, SBI-02, and SBI-10 in our study.

4.3.2.2 Plant height

Seventeen QTLs were detected in nine chromosomes associated with Plant height at LOD score ≥ 3.0 . A total of nine significant QTLs were detected on SBI-01, SBI-02, SBI-04a, SBI-4b, SBI-06, SBI-07, SBI-08, SBI-09 and SBI-10. The phenotypic variation explained by the

detected QTLs ranged from 7.5 to 26.1% whereas the additive effects ranged from -18.79 to -0.197 cm for SS79 and 3.05 to 9.479 cm for M71 parent. Most of the QTLs (61%) had negative additive effects suggesting that the parent SS79 had a higher contribution of tall alleles. Pereira and Lee (1995) detected eight QTLs by interval mapping and single factor analysis on four different linkage groups, which corresponded to SBI-06, SBI-07, SBI-09, and SBI-10. Supporting our study, their QTLs accounted for 9.2 to 28.7% phenotypic variation

4.3.2.3 Stem diameter

Thirteen QTLs and one putative QTL influencing stem diameter were distributed in eight chromosomes. Nine out of thirteen QTLs were significant across the environments. The QTL on SBI-03 at position 102 cM seem to be the major QTL because it was significant in 3 environments and explaining 13.3% of phenotypic variation. The additive effect of this QTL was -0.042 cm suggesting that the QTL allele favours the female parent. Besides, the QTL in SBI-07 at position 30 also seem to be a major QTL because of its high additive effect (0.069 cm), detected at high LOD score of 5.03 and explaining 11.6% of the phenotypic variation.

4.3.2.4 Numbers of tillers per plant

Six significant QTLs were detected at $LOD \geq 2.5$ which were associated with number of tillers per plant. A QTL on SBI-06 at position 64 was found significant in all the environments and was detected at LOD score of 2.90 using the overall mean of the data set. In total, fourteen QTLs were detected in nine chromosomes explaining a phenotypic variation ranging from 6.7 to 15.7% and their additive effect ranging from -0.09 to 0.092 number per plant. Only three of the detected QTLs had positive additive effects indicating that the increase in tiller numbers was contributed by SS79.

4.3.2.5 Fresh panicle weight

At a LOD score of 5.62, a major significant QTL associated with fresh panicle weight was detected on SBI-06. This QTL explained 12.9% of phenotypic variation and showed additive effect of 4.87 g/plant. Srinivas et al. (2009) detected a major single QTL on SBI-06 in 168 F₇ RILs of sorghum explaining 14.7% of the phenotypic variation at a LOD score of 4.90. In our study, four other significant QTLs were detected at a LOD score ≥ 2.5 on SBI-01, SBI-02, and SBI-04a. Together, they explained 20.3% of the phenotypic variation. This study supports some findings of Murray et al. (2008) where seven QTLs were detected in various environments on SBI-04, SBI-06, SBI-09 and SBI-10.

4.3.2.6 Stem juice weight

For stem juice weight, one QTL was detected on SBI-02 between 12 and 26 cM with LOD score of 3.54. This QTL explained 7.9% of the phenotypic variation and had additive effect of - 17.603 g / plants. The negative additive effect shows that the allele was derived from the SS79. Stem juiciness was not segregating well in the population because both parents are juicy and also significantly controlled by the environment.

4.3.2.7 Brix

Four out of fourteen detected QTLs significantly influenced the total soluble solids in sorghum stem juice. These significant QTLs were distributed on three chromosomes where chromosome four had two significant QTLs. Two major QTLs were identified on SBI-02 and SBI-06 at position 82 and 16, respectively. These QTLs explained 21.9 and 22.3% of the total phenotypic variation, respectively. Besides, they had the LOD score ≥ 10.00 and the additive effects of -0.167 and -0.797, respectively. The negative signs of both alleles show that these alleles were derived from the SS79 parent with high sugar. Ritter et al. (2008) reported 5 QTLs on SBI-05 and SBI-06 in two environments. Murray et al. (2008) reported five QTLs on SBI-03, SBI-06 and SBI-07 that explained 4.3 to 10% phenotypic variation with the additive effect ranging from -0.71 to -0.51.

4.3.2.8 Glucose content

A total of ten QTLs and one suggestive QTL were detected for glucose on seven chromosomes. Other studies reported QTLs on SBI-02, SBI-05, and SBI-09 (Murray et al., 2008), and SBI-07 (Ritter et al., 2008). Out of these eleven QTLs in our study, four QTLs were found significant among the four environments or the mean of the entire environments. The QTL detected at position 70 of SBI-07 seem to be the major QTL. This QTL had LOD score of 5.99 explaining 13.7% phenotypic variation with additive effect of -0.351 g/l. This QTL was significant in overall environment and two single environments, and its negative additive effects show the contribution of the female parent in glucose content. In general, both parents contributed alleles for increasing and decreasing glucose content.

4.3.2.8 Sucrose content

For sucrose content, seven QTLs and two putative QTLs were detected on seven chromosomes where four QTLs were found to be significant. These significant QTLs explained 42% of the phenotypic variation where a QTL on SBI-06 explained the highest

phenotypic variation (10.6%) compared to the others. Ritter et al. (2008) detected eleven QTLs on SBI-01, SBI-05, SBI-06, and SBI-10. In our study, no QTL was detected on SBI-05. Both negative and positive additive effects were observed showing the contribution of both parents to sucrose content.

4.3.2.9 Sugar content

A total of fifteen QTLs and two putative QTLs were detected that controls sugar content in stem juice on seven different chromosomes. Out of these seventeen QTLs a total of nine QTLs were found significant in different environments and overall environment. Ritter et al. (2008) detected ten QTLs in four linkage groups at two environments where linkage group six was found with the highest cluster of four QTLs. In our study, all detected QTLs explained from 6.6 to 12.5% phenotypic variation. The additive effects of the QTLs were either negative or positive showing that the QTLs were detected from all parents.

4.3.3 Epistatic effect and QTL x environment interaction

The results of digenic epistatic effects are presented in table 4.2. A total of twenty-four significant digenic epistatic interactions (AA) were detected for seven traits using the overall environments suggesting that two locus interactions were widespread in the entire genome. A total of five digenic QTLs interaction were found in fresh panicle weight, four were found in plant height and glucose content, and three were found in number of tillers, sucrose and sugar content and two were found in flowering dates. There were no significant additive x additive interactions observed for estimated juice weight, brix and stem diameter. Purcell and Sham (2004) concluded that epistasis is difficult to detect due to most of the experimental variation being assigned to main effects. According to Whitlock et al. (1995) detecting epistasis has limitation due to genotype x environmental interactions, and linkage disequilibrium. In flowering dates, one significant additive x additive interactions could be detected between SBI-04 and SBI-08 with the effect of -1.437 together explaining 3.0% of the phenotypic variance. Ritter et al. (2008) detected additive x additive effect between SBI-04 and SBI-08 with the effect of 0.979. These interactions in both studies favour the late flowering parents.

Epistasis between the same loci affecting sucrose and sugar content was found between SBI-01 and SBI-4b, and SBI-01 and SBI-06. In other traits additive x additive interactions differed per traits. The total phenotypic variation explained by interactions for flowering dates was 8.7% from two digenic pairs, plant height was 21.4% from four digenic pairs, fresh panicle

weight was 25.4% from five digenic pairs, tillers were 18.7% from three digenic pairs, glucose content was 21.7% from four digenic pairs, sucrose content was 19.1% from three digenic pairs, and sugar content was 17.8% from three digenic pairs. In general, the total phenotypic variations explained by the interactions were lower than that of corresponding main effects. The importance of additive x additive interaction effects in total genetic effects may be trait-dependent.

Uncertainty in terminology regarding epistasis often arises because of the differences in definitions and types of epistasis being reported. The epistatic interaction effect was not consistent in our study. For example, digenic pairs in plant height on SBI-01 and SBI-09 had a negative effect separate and produced a positive effect together, in sucrose content the interaction between SBI-01 and SBI-4b showed a strong negative effect while they showed small effects with different signs separately. The simplest statistic interactions are duplicate and complementary interactions. Duplicate interactions occur when two or more loci perform the same function. If at least one locus is functioning, the gene product will be produced. The failure of one locus to produce its product will not result in an abnormal phenotype. Complementary interactions results when more than one gene is necessary for the final product. For example, if two or more genes are involved in the biosynthetic pathways, both genes must be functioning and their products must be present in order to produce the final product. If either gene product is absent, the pathway will cease to function and the final product will not be produced. Although it is known that biosynthetic pathways of sugar traits involve series of genes, no conclusions can be made on the type of epistasis reported.

Analysis of epistasis is often done in combination with QTL mapping, but this result in confusing the epistatic effect with background genome segregation and QTL x environmental interactions (Holland, 2001). Routman and Cheverud (1997) reported that epistasis in both negative and positive direction is important for the breeding populations even if the net effects are not observed. In this study, interaction between QTL and environments was analyzed for all traits (table 4.3, supplementary results). A significant QTL x environmental interaction was observed in sugar content on SBI-01, fresh panicle weight on SBI-1a and SBI-06, tillers on SBI-07 and SBI-10 and flowering dates on SBI-06 and no significant QTL x environment interactions was observed in other traits.

4.3.4 Pleiotropic effect and co-localization

To identify a linkage group as a hotspot, we considered QTL clusters that have at least one QTL explaining more than 5% phenotypic variance with both paternal and maternal effects as reported by Rae et al. (2009). QTLs for all traits co-localized in all linkage groups except for linkage group 3 and 5 (Supplementary table 4.3, Figure 4.1). The hotspots can be declared on SBI-01, SBI-02, SBI-06, SBI-07, and SBI-09 because all traits co-localized on those linkage groups and more than one QTL has been found per traits.

Several QTLs that influenced multiple traits were observed in this study (table 4.3, figure 4.1). The major QTL with multiple effects was detected in linkage group six at position interval 12-18 cM. This QTL has been detected in all traits except for glucose and number of tillers per plant, and was significant in five out of six traits. The additive effects of this QTL differed per trait where it favoured the female parent in flowering dates, plant height, total soluble solids, and sucrose but favoured the male parent in fresh panicle weight and estimate of sugar content. It has been reported in chapter two that fresh panicle weight was negatively correlated with flowering dates, plant height, total soluble solids and sucrose, and positively correlated with total sugar content.

Other QTLs with pleiotropic effect has been detected in various linkage groups such as E44M48-140 on SBI-04a, E35M49-212 on SBI-07, Xtxp51 on SBI-04a and Stgnhsbmnsbm36 on SBI-02. Most of the QTLs detected for sucrose were also detected either in Brix, or total sugar content or in both.

4.4 Discussion

The growing importance of bioenergy-derived sugar crops such as sorghum and sugarcane is likely to play a role in environmental and economical challenges of fossil fuel usage. Similar to sugarcane, our lines differed in sucrose accumulation as reported by Lingle (1997). However, in our lines it is not known if the accumulation mechanisms differ per genotype during growth as in sugarcane. As sugar accumulation and content is controlled by polygenes and significantly affected by the environment, success in breeding strategies can be based on partitioning the agronomic traits and sugar components into various studies as it is done in drought tolerance studies. In the current study, agronomic traits related to sugar content such as flowering dates, plant height, number of tillers, stem diameter, fresh panicle weight and

stem juice, and sugar components such as glucose, sucrose, brix, and total sugar content were investigated.

Six maturity genes are known to alter flowering in sorghum: *Ma*₁, *Ma*₂, *Ma*₃, *Ma*₄, *Ma*₅, and *Ma*₆ (Aydin et al., 1997; Quinby, 1967). The first four genes cause long days to inhibit flowering but allow early flowering under short days while the last two can inhibit floral initiation regardless of day length. Collinearity between rice chromosomes seven and sorghum chromosome two (in our study) has been observed by Paterson et al. (1995; 2005). In rice chromosome seven has been reported to carry photoperiod sensitive genes *Se-2* and *El* (Laurie 1997; Yano et al., 1997). In our study, linkage group 3 harboured significant QTL associated with flowering dates in three environments and with the overall mean.

Quinby and Karper in 1954 reported that plant height is controlled by four independent genes: *Dw1*, *Dw2*, *Dw3*, and *Dw4*. The dwarf gene *Dw2* has a pleiotropic effect to grain yield and panicle length (Graham and Lessmann, 1966), and *Dw3* has a pleiotropic effect to number of tiller and panicle size (Casady, 1965). The QTLs detected for plant height in this study co-localized with QTLs in other morphological and sugar traits. Pereira and Lee (1995) and Rami et al. (1998) observed co-localization of QTL for plant height and stem diameter, with other morphological traits and concluded that these QTLs can be *Dw* loci. Although the markers reported in these two studies are different from ours to verify, this hypothesis is a possibility in our studies.

Several alignments between *Saccharum* and *Sorghum* chromosomes has been done in recent years (Dufour et al., 1997; Guimarães et al., 1997; Ming et al., 1998, 2001), which is important to determine collinearity or orthologs of sugar-related QTLs. Al Janabi et al., (1994) suggested that sorghum and sugarcane diverged approximately five million years ago. The map created by Ming et al. (2002) showed that six of the thirteen homology groups identified in the consensus map of sugarcane is closely correspondence to that of sorghum while the remaining showed varying level of synteny. Jordan et al. (2004) reported that stalk number and suckering (underground branching or tillering) influence sugar yield in sugarcane. They aligned the maps of sugarcane generated by Grivet et al. (1996) with those of sorghum generated by Paterson et al. (1995) and Boivin et al. (1999). Using the sugarcane RFLP probes, they identified QTLs for rhizome number, rhizome distance, regrowth, and seedling tillers on SBI-01, SBI-02, SBI-03, SBI-06, and SBI-07 based on the chromosome

nomenclature of sorghum by Kim et al. (2005). The biggest cluster of these QTLs was found on chromosome 1. In the current study, QTLs associated with number of tillers were detected on all chromosomes except on chromosome eight.

Ming et al. (2001) detected 36 sugar content QTLs in sugarcane using two interspecific segregating populations derived from *S. officinarum* and *S. spontaneum*. The sugar content meant sucrose at 96% purity, calculated based on Brix and Pol values. Through alignment, the detected QTLs corresponded to eight regions of the sorghum genome of linkage groups A, B, C, D, F, G and I. By searching for the linked markers in www.gramene.org and also using the maps of Ming et al. (1998), these linkage groups seem to be SBI-01, SBI-02, SBI-03, SBI-04, SBI-06, SBI-09, and SBI-10. Although in our study the QTLs for brix, sucrose, and sugar content were detected in these linkage groups, the exact positions of the QTLs cannot be aligned for comparative analysis. However, as Ming et al. (2001) concluded, the observed QTL between the two species may be accounted for by a much smaller number of ancestral genes that could have been multiplied by the rapid duplication of chromosomes that characterized sugarcane genome evolution since its divergence from a common ancestor shared with sorghum.

4.5 Conclusion

The composite interval mapping detected more than five additive QTLs per traits. As reported by Murray et al. (2008) and Ritter et al. (2008), sugar related traits in sweet sorghum are mainly influenced by additive genes. Seven out of ten traits studied showed that additive x additive effects exist between QTLs associated with sugar-related traits. However, the failure of detecting digenic epistasis in other traits does not mean that it does not exist. QTLs showing multiple effects were detected, while the major one was found at 14 cM on SBI-06. Enhancing sugar-related traits can be faster if pleiotropic effects favour the positive alleles.

4.6 Acknowledgement

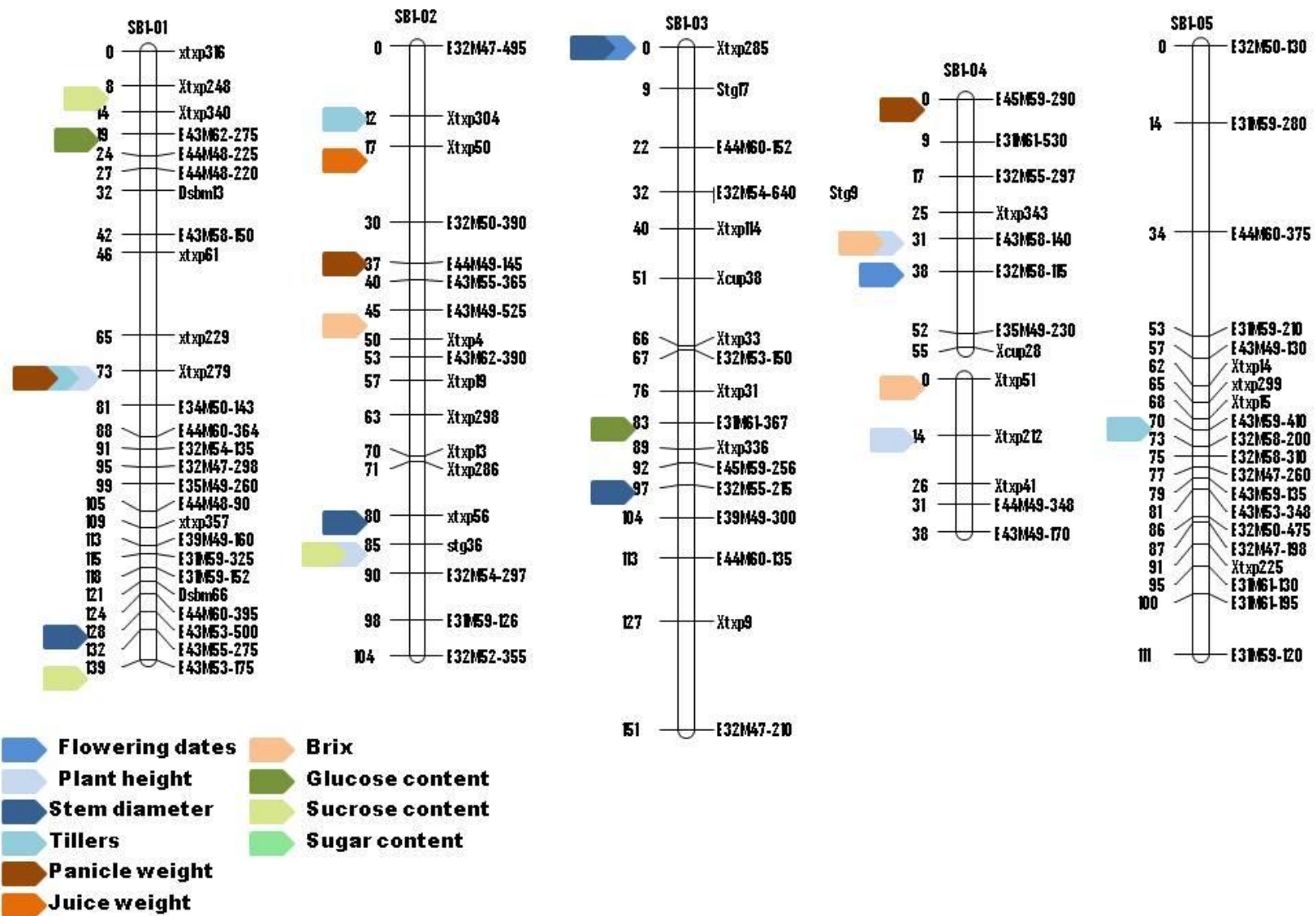
The authors would like to thank Mario Tolksdorf, Markus Kollmer and Malte Luh for field technical assistance. This work has been primarily supported by the DAAD, Agricultural Research Council and National Research Foundation of South Africa.

Table 4.1: Descriptive statistics and analysis of the variance

Trait	RECOMBINANT INBRED LINES						Anova with GLM						
	Parents		Mean	std dev	Mini	Max	Genotype	Env	G*E	Hardate	CV	H ²	R ²
Flowering dates	112.00	95.00	105.85	6.78	91.69	142.73	***	***	***	***	2.81	0.33	0.75
Plant Height cm	238.27	118.23	152.61	38.89	69.25	251.63	***	***	***	NS	21.31	0.81	0.48
Stem Diameter cm	1.66	1.76	1.77	0.16	1.33	2.26	***	***	***	NS	11.19	0.6	0.81
Number of Tillers per plant	3.00	2.00	2.23	0.35	1.22	5.40	***	***	***	NS	26.44	0.68	0.7
Fresh Panicle weight g/plant	43.96	105.43	57.97	19.00	2.57	124.53	***	***	***	NS	37.37	0.69	0.56
Estimated Juice weight g/plant	488.19	192.35	296.8	93.28	148.63	623.08	NS	***	NS	NS	261.67	0.18	0.17
Brix	16.98	10.71	14.89	1.59	10.02	17.93	***	***	***	NS	14.39	0.59	0.75
Glucose content g/L	12.38	5.49	7.88	0.99	5.02	19.97	***	***	***	NS	29.34	0.47	0.64
Sucrose content g/plant	112.82	48.89	115.92	15.83	42.87	156.12	***	***	***	NS	28.82	0.52	0.75
Sugar content g/plant	134.40	59.15	131.68	16.49	44.39	176.87	***	***	***	NS	24.966	0.5	0.6

GLM: General linear model Std dev: standard deviation min: minimum max: maximum Env: environment G*E: genotype x environment Hardate: Harvest date CV: coefficient of variation H²: Heritability R²: Repeatability

*** significant at 0.001 probability level NS non significant level



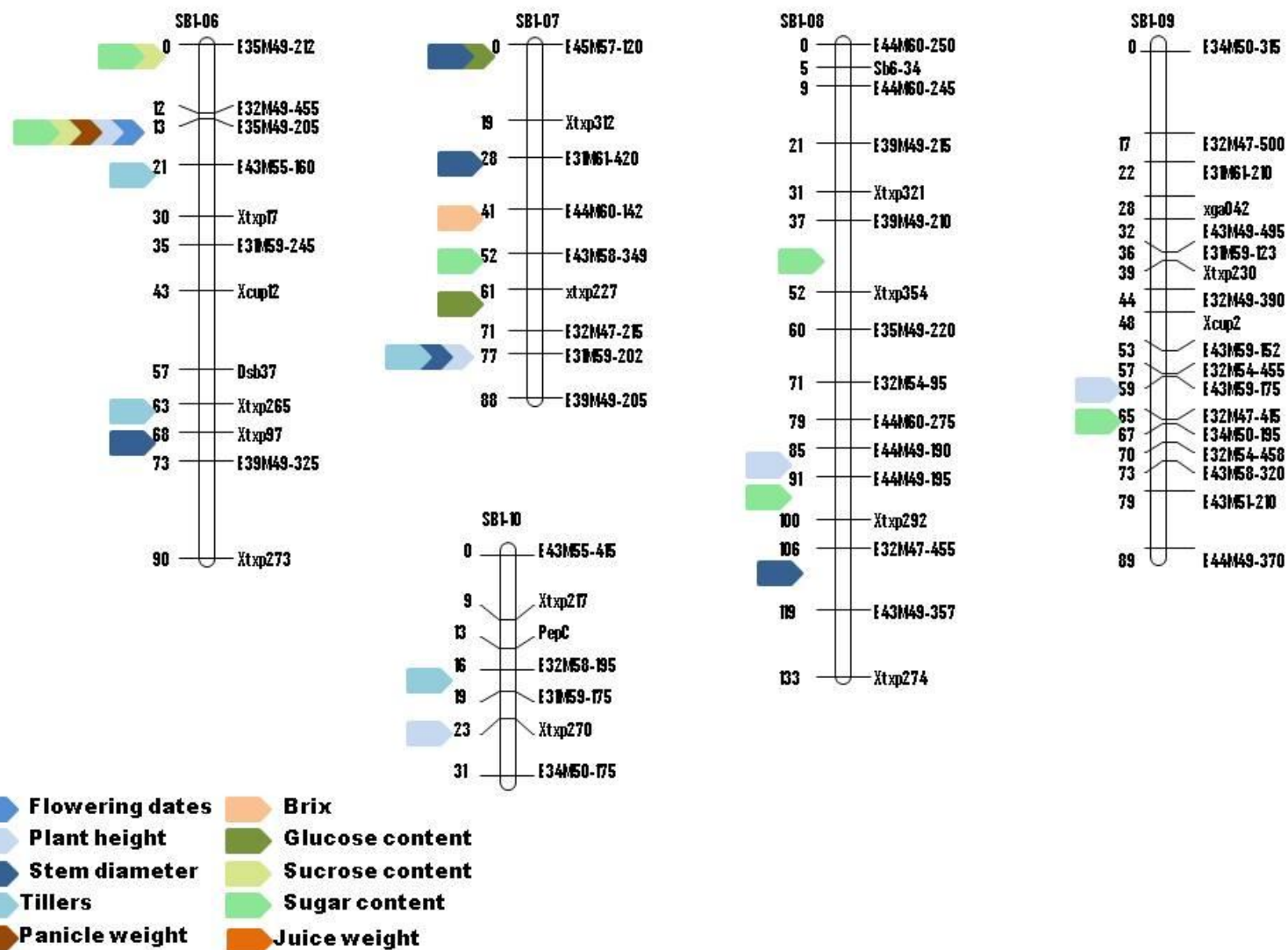


Figure 4.1. Genetic linkage map of *Sorghum bicolor* L. Moench based on 157 SSR, AFLP and EST-SSRs. The map covers 1029.08 cM with the average distance of 6.55 cM. The QTLs presented are detected by composite interval mapping and are significant either in combined environments or in each single environment or both. The details of the presented QTLs which include the position, additive effect and phenotypic variation explained are in table 4.3, supplementary results.

Table 4.2: Digenic epistatic loci associated with sugar-related traits

Traits	Chromosome	Left marker	chromosome	Left marker	Additive effect	Intercept	R ² %
Flowering dates	SBI-4a	E32M58-115	SBI-07	Xtxp227	-1.437	105.36	3
	SBI-4a	E32M58-115	SBI-08	Xtxp354	-0.023		5.7
Plant Height [cm]	SBI-01	E44M60-364	SBI-06	Xtxp17	11.197	149.44	6.3
	SBI-01	E44M60-364	SBI-09	E32M54-458	8.855		5.1
	SBI-01	Xtxp279	SBI-05	E31M61-130	8.255		5.3
	SBI-01	E44M48-225	SBI-02	Stgnhsbm36	-8.263		4.7
Stem diameter [cm]	no significant interaction found						
Number of Tillers	SBI-01	E43M48-150	SBI-02	Xtxp304	-0.113	2.21	7.5
	SBI-03	E44M60-152	SBI-06	Xtxp265	-0.098		5.7
	SBI-04a	E45M59-290	SBI-06	E43M55-160	0.09		5.5
Fresh panicle weight [g/plant]	SBI-01	Xtxp279	SBI-09	E43M49-495	4.768	59.89	5.5
	SBI-02	E44M49-145	SBI-07	xtp227	4.291		5.8
	SBI-05	E32M47-198	SBI-07	xtp227	-3.564		3.9
	SBI-06	E35M49-205	SBI-07	E45M57-120	4.493		5.5
	SBI-06	E35M49-205	SBI-08	E39M49-210	-4.985		4.7
	Estimated juice weight [g/plant]	no significant interaction found					
Brix	no significant interaction found						
Glucose content [g/plant]	SBI-01	E43M62-275	SBI-07	xtp227	-0.192	7.95	4.9
	SBI-03	E31M61-367	SBI-05	E32M50-130	-0.241		6.9
	SBI-03	E44M60-135	SBI-05	E43M53-348	-0.229		5.4
	SBI-07	E45M57-120	SBI-09	E43M51-210	0.207		4.5
	Sucrose content [g/l]	SBI-01	E39M49-160	SBI-4b	Xtxp51	-4.531	114.85
SBI-01		Xtxp279	SBI-06	E35M49-205	-6.12		6.2
SBI-06		E35M49-205	SBI-07	E44M60-142	5.441		6.8
Sugar content [g/l]	SBI-01	E39M49-160	SBI-4b	Xtxp51	-4.52	130.15	5.5
	SBI-01	Xtxp279	SBI-06	E35M49-205	-6.016		6.7
	SBI-02	Xtxp286	SBI-07	E43M58-349	-4.382		5.6

R² % Partial phenotypic variance explained in percentage

4.7 References

- Aitken, K.S., S. Hermann, K. Karno, G.D. Bonnett, L.C. McIntyre, and P.A. Jackson. 2008. Genetic control of yield related stalk traits in sugarcane. *Theor. Appl. Genet* 117:1191-1203
- Aitken, K.S., P.A. Jackson, and C.L. McIntyre. 2006. Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar x *Saccharum officinarum* population. *Theor. Appl. Genet* 112:1306-1317
- Al-Janabi, S., M. McClelland, C. Petersen, and B. Sobral. 1994. Phylogenetic analysis of organellar DNA sequences in the *Androponeae: Saccharinae*. *Theor. Appl. Genet* 88:933-644
- Aydin, S., W.L. Rooney, and F.R. Miller. 1997. Identification and characterization of the Ma5 and Ma6 maturity loci in Sorghum. In: Proceeding of the International Conference on genetic improvement of Sorghum and Pearl Millet, 22-27 September 1996, Lubbock, USA, INTSORMIL and ICRISAT: 641-642
- Berg, E.S., and B. Olaisen. 1994. Hybrid PCR sequencing-sequencing of PCR products using a universal primer. *Biotechniques* 17:896-901
- Bhatramakki, D., J. Dong, A. Chhabra, and G. Hart. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* 43:988-1002
- 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* 98:320-328
- Casady, A.J. 1965. Effect of a single height (*Dw*) gene of sorghum on grain yield, grain yield components, and test weight. *Crop Sci* 5:385-389
- Chantereau, J., G. Trouche, J. F. Rami, M. Deu, C. Barro, and L. Grivet. 2004. RFLP mapping of QTLs for photoperiod response in tropical sorghum. *Euphytica* 120:183-194
- Dean, R.E., J.A. Dahlberg, M.S. Hopkins, S.E. Mitchell, and S. Kresovich. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. National sorghum collection assessed with simple sequence repeats (SSR) markers. *Crop Sci* 39:1215-1221
- Doyle, J.J., and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15
- Dufour, P., M. Deu, L. Grivet, A. D'Hont, F. Paulet, A. Bouet, C. Lanaud, J.C. Glaszmann, and P. Hamon. 1997. Construction of a composite sorghum genome map and

- comparison with sugarcane, a related complex polyploid. *Theor. Appl. Genet* 94:409–418
- Feltus, F.A., S.E. Hart, K.F. Schertz, A.M. Casa, S. Kresovich, S. Abraham, P.E. Klein, P.J. Brown, and A.H. Paterson. 2006. Alignment of genetic maps and QTLs between inter- and intra-specific sorghum populations. *Theor. Appl. Genet* 112:1295-1305
- Graham, D., and K.J. Lessmann. 1966. Effect of height on yield and yield components of two isogenic lines of *Sorghum vulgare* Pers. *Crop Sci* 6:372–374
- Grivet, L., A. D'Hont, D. Roques, P. Feldmann, C. Lanaud, and J. C. Glaszmann. 1996. RFLP mapping in cultivated sugarcane (*Saccharum spp.*): Genome organization in a highly polyploid and aneuploid interspecific hybrid. *Genetics* 142:987–1000
- Guimarães, C.T., G.R. Sills, and B.W.S. Sobral. 1997. Comparative mapping of *Andropogoneae*: *Saccharum* L. (sugarcane) and its relation to sorghum and maize. *Proc. Natl. Acad. Sci* 94:14261–14266
- Hasan, M., F. Seyis, A.G. Badani, J. Pons-Kühnemann, W. Friedt, W. Lühs, and R. J. Snowdon. 2006. Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. *Genet. Resour. Crop Evol* 53:793-802
- Hasan, M., W. Friedt, J. Pons-Kühnemann, N.M. Freitag, K. Link, and R.J. Snowdon. 2008. Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *Napus*). *Theor. Appl. Genet* 116:1035-1049
- Holland, J.B. 2001. Epistasis and plant breeding. *Plant Breed. Rev* 21:27-92
- Jordan, D.R., R.E. Casu, P. Besse, B.C. Carroll, N. Berding, and C.L. McIntyre. 2004. Markers associated with stalk number and suckering in sugarcane collocate with tillering and rhizomatousness QTLs. *Genome* 47:988-993
- Kim, J-S., P.E. Klein, R.R. Klein, H.J. Price, J.E. Mullet, and D.M. Stelly. 2005. Chromosome identification and nomenclature of *Sorghum bicolor*. *Genetics* 169: 1169-1173
- Kong, L., J. Dong, and G.E. Hart. 2000. Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple sequence repeats (SSRs). *Theor. Appl. Genet* 101:438-448
- Laurie, D.A. 1997. Comparative genetics of flowering time. *Plant Mol. Biol* 35:167–177
- Lingle, S.E. 1997. Seasonal internodes development and sugar metabolism in sugarcane. *Crop Sci* 37:1222–1227
- Mace, E.S., J-F. Rami, S. Bouchet, P.E. Klein, R.R. Klein, A. Kilian, P. Wenzl, L. Xia, K. Halloran, and D.R. Jordan. 2009. A consensus genetic map of sorghum that intergrades

- multiple component maps and high-throughput diversity array technology (DarT) markers. *BMC Plant Biol.* 9.
- Milligan, S.B., K.A. Gravois, K.P. Vischoff, and F.A. Martin. 1990. Crop effects on genetic relationships among sugarcane. *Crop Sci* 30:927-931
- Ming, R., S.-C. Liu, Y.-R. Lin, J. da Silva, W. Wilson, et al. 1998. Detailed alignment of *Saccharum* and *Sorghum* chromosomes: Comparative organization of closely-related diploid and polyploid genomes. *Genetics* 150:1663–1682
- Ming, R., S.C. Liu, P.H. Moore, J.E. Irvine, and A.H. Paterson. 2001. QTL analysis in a complex autopolyploid: Genetic control of sugar content in sugarcane. *Genome Res* 11:2075–2084
- Ming, R., Y.-W. Wang, X. Draye, P.H. Moore, J.E. Irvine, and A.H. Paterson. 2002. Molecular dissection of complex traits in autopolyploids: mapping QTLs affecting sugar yield and related traits in sugarcane. *Theor. Appl. Genet* 105:332-345
- Murray, S.C., A. Sharma, W.L. Rooney, P.E. Klein, J.E. Mullet, S.E. Mitchel, and S. Kresovich. 2008. Genetic improvement of sorghum as biofuel feedstock I: QTL for stem sugar and grain non-structural carbohydrates. *Crop Sci* 48:2165-2179
- Natoli, A., C. Gorni, F. Chegani, P. Ajmone Marsan, C. Colombi, C. Lorenzoni, and A. Marocco. 2002. Identification of QTLs associated with sweet sorghum quality. *Maydica* 47:311-322
- Parth, D.K., D.R. Jordan, E.A.B. Aitken, E.S. Mace, P. Jun-ai, C.L. McIntyre, and I.D. Godwin. 2008. QTL analysis of ergot resistance in sorghum. *Theor. Appl. Genet* 117:369-382
- Paterson, A.H., Y.-R. Lin, Z. Li, K.F. Schertz, J.F. Doebley, S.R.M. Pinson, S.-C. Liu, J.W. Stansel, and J.E. Irvine. 1995. *Science* 296:1714-1718
- Paterson, A.H., M. Freeling, and T. Sasaki. 2005. Grains of knowledge: Genomics of model cereals. *Genome res* 15:1643-1650
- Pereira, M.G., and M. Lee. 1995. Identification of genomic regions affecting plant height in sorghum and maize. *Theor. Appl. Genet* 90:380-388
- Purcell, S., and P.C. Sham. 2004. Epistasis in quantitative trait locus linkage analysis: Interaction or main effects? *Behav. Genet* 34:143-152
- Quinby, J.R. 1967. The maturity genes of sorghum. In: *Advances in agronomy XIX*. Norman, A.G. ed., Academic Press Inc. New York :267-305
- Quinby, J.R., and R.E. Karper. 1954. Inheritance of height in sorghum. *Agron. J* 46:211–216

- Rae, A.M., N.R. Street, K.M. Robinson, N. Harris, and G. Taylor. 2009. Five QTL hotspots for yield in short rotation coppice bioenergy poplar: The poplar Biomass Loci. *BMC Plant Biol* 9:23
- Rami, J.F., P. Dufour, G. Trouche, G. Fliedel, C. Mestres, F. Davrieux, P. Blancard, and P. Hamon. 1998. Quantitative trait loci for grain quality, productivity, morphological, and agronomical traits in sorghum (*Sorghum bicolor* L. Moench). *Theor. Appl. Genet* 97:605-616
- Ritter, K.B., D.R. Jordan, S.C. Chapman, I.D. Godwin, E.S. Mace, and C.L. McIntyre. 2008. Identification of QTL for sugar-related traits in sweet x grain sorghum (*Sorghum bicolor* L. Moench) recombinant inbred population. *Mol. Breed* 22:367-384
- Routman, E.J., and J.M. Cheverud. 1997. Gene effects on a quantitative trait: two locus epistatic effects measured at microsatellite markers and at estimated QTL. *Evolution* 51:1654-1662
- Rygulla, W., R.J. Snowdon, C. Eynck, B. Koopmann, A. van Tiedemann, W. Lühs, and W. Friedt. 2007. Broadening the genetic basis of *Verticillium longisporum* resistance in *Brassica napus* by interspecific hybridisation. *Phytopathology* 97:1391-1396
- Schloss, S.J., S.E. Mitchell, G.M. White, R. Kukatla, J.E. Bowers, A.H. Paterson, and S. Kresovich. 2002. Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. *Theor. Appl. Genet* 105:912-920
- Srinivas, G., K. Satish, R. Madhusudhana, and N. Seetharama. 2009. Exploration and mapping of microsatellite markers from subtracted drought stress ESTs in *Sorghum bicolor* (L.) Moench. *Theor. Appl. Genet* 118:703-717
- Srinivas, G., K. Satish, R. Madhusudhana, R.N. Reddy, S.M. Mohan, and N. Seetharama. 2009. Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellites markers in sorghum. *Theor. Appl. Genet* 118:1439-1454
- Srinivas, G., K. Satish, S.M. Mohali, R.N. Reddy, R. Madhusudhana, D. Balakrishna, B.B. Venkatesh, C.J. Howarth, and N. Seetharama. 2008. Development of genic-microsatellite markers for sorghum staygreen QTL using comparative genomic approach with rice. *Theor. Appl. Genet* 117:703-717
- Taramino, G., R. Tarchini, S. Ferrario, M. Lee., and M.E. Pe. 1997. Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. *Theor. Appl. Genet* 95:66-72

- Ungerer, M.C., S.S. Halldorsdottir, J.L. Modliszekwski, T.F.C. Mackay, and M.D. Purugganan. 2002. Quantitative trait loci for inflorescence development in *Arabidopsis thaliana*. *Genetics* 160:1133-1152
- Uptmoor, R., W.G. Wenzel, W. Friedt, G. Donalson, K. Ayisi, and F. Ordon. 2003. Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions from southern Africa by RAPDs, AFLPs and SSRs. *Theor. Appl. Genet* 106:1316-1325
- Utz, H.F., and A.E. Melchinger. 1994. Comparison of different approaches to interval mapping of quantitative trait loci. In: *Biometrics in plant breeding: Applications of molecular markers*. van Ooijen, J.W. and J. Jansen (ed), wageningen:195-204
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, and M. Kuiper. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 21:4407-4414
- Whitlock, M.C., P.C. Phillips, B.G. Moore, and S.J. Tonsor. 1995. Multiple fitness peaks and epistasis. *Ann. Rev. Ecol. Syst.* 26:601-629
- Xu, X.Y., G.H. Bai, B.F. Carver, and G.E. Shaner. 2005. Mapping of QTLs prolonging the latent period of *Puccinia triticina* infection in wheat. *Theor. Appl. Genet* 110:244-251
- Yano, M., Y. Harushima, Y. Nagamura, N. Kurata, Y. Minobe, and T. Sasaki. 1997. Identification of quantitative trait loci controlling heading date in rice using a high-density linkage map. *Theor. Appl. Genet* 95:1025–1032
- Yun-long, B., Y. Seiji, I. Maiko, and C. Hong-Wei. 2006. QTLs for sugar content of stalk in sweet sorghum (*Sorghum bicolor* L. Moench). *Agric. Sci. In China* 5:736-744

Chapter 5

Identification of QTL for fibre-related traits in grain x sweet sorghum

Amukelani Shiringani and Wolfgang Friedt

Abstract: Compared to temperate grasses, as well as maize, sorghum received less attention in terms of cell wall components improvement. With the aid of high throughput genomics technology, a significant knowledge and understanding of cell wall components quantity, structure, properties and the precise interrelationship will be accumulated. The objectives of the study were to identify 1) additive QTLs, 2) additive x additive interactions, and QTL x environment interaction 3) additive QTLs with multiple effects and co-localization in fibre – related traits. Fibre components such as neutral detergent fibres (NDF), acidic detergent fibre (ADF), acidic detergent lignin (ADL), cellulose, hemicellulose, fresh leaf mass, stripped stalk mass, dry stalk mass, fresh biomass, and dry biomass were analysed from a population of 213 grain x sweet sorghum recombinant inbred lines. The population was planted in 4 environments. The phenotypic performance was analysed using general linear model of SAS[®] 9.0 software. The additive QTLs, additive x additive interactions, QTL x environment interactions were analysed using composite interval mapping (CIM) of PLABQTL. The results showed that genotype and environment were highly significantly different among each other, and a significant genotype x environmental interaction was observed. All source of variance were not significant in dry stalk mass. CIM detected few additive QTLs associated with fibre related traits distributed across the ten chromosomes explaining the phenotypic variation that ranged from 6.1 to 24.7%. Abundant co-localization of these QTLs was observed across all chromosomes, and the highest cluster was observed on chromosome 6. Some QTLs showed multiple effects, however, the allele for each trait was favouring the parent with the increasing effect. QTL x environment interaction was observed in QTLs showing multiple effects. Additive x additive interaction was observed in seven out of ten traits indicating the importance of epistatic analysis. The phenotypic variation explained by the digenic interaction was lower compared to the individual QTL; however, there was no clear rule that controlled the additive effect.

5.1 Introduction

Cellulose is the most abundant polymer on earth and it is the core of the plant cell wall (Lerouxel et al., 2006; Taylor, 2008). In primary cell wall, cellulose is important for load-

bearing network and determining the orientation of cell expansion, while in secondary cell wall it provides mechanical properties that allow the cell to stand upright and facilitate the xylem vessels to function properly (Taylor, 2008). Cellulose is more abundant in secondary cell wall having 40% dry mass compared to in primary cell wall that has 15-30% dry mass (Sticklen, 2008).

Recently, high throughput genomic technologies made a significant contribution in understanding the molecular details involved in cellulose biosynthesis (Dhugga, 2007; Lerouxel et al., 2006). Genome sequencing revealed that all higher plants have multiple cellulose synthase catalytic subunit (CESA) proteins encoded by the CESA genes, which share a conserved structure. Plant genomes contain multiple CESA genes as part of the gene family, which may assemble to form a functional complex. Cellulose is made by a complex enzyme system (Dhugga, 2007; Hazen et al., 2002; Lerouxel et al., 2006) that is integrated into the plasma membrane after being assembled in the Golgi bodies.

Unlike cellulose that is a branched polymer and contains anhydrous glucose, hemicellulose is an unbranched polymer containing glucose, xylose, mannose, arabinose, and galactose. These sugar components resulted in classification of hemicelluloses into xylogans that contains β -1,4-glucan backbone, mannans that contain β -1,4-linked mannose (glucose and mannose), glucuronoarabinoxylans that contain β -1,4-linked xylan backbone, and mixed linkage glucans that contain linkages of both β -1,4-glucan and β -1,3-glucan. Cosgrove (2005) reported that xyloglucan and arabinoxylan are two of the most abundant hemicelluloses.

Structural similarities between β -1, 4-glucan chains of cellulose and the backbones of various β -linked hemicellulosic polysaccharides suggests that Cellulose Synthase like (CSL) genes might be involved in the biosynthesis of hemicelluloses (Lerouxel et al., 2006). The CSL are a family of genes that have a sequence similar to the CESA genes and appear to be present in all plant genomes. Hemicelluloses are synthesized in the Golgi, package into vesicles and then exported to the cell wall through exocytosis.

Lignin is an amorphous polymer consisting of three aromatic alcohols (monolignols): *p*-coumaryl, coniferyl and sinapyl alcohols (Buranov and Mazza, 2008). The respective aromatic constituents of these alcohols in the polymer are *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) moieties (a part or half a molecule), which are transported from the cytosol to

the apoplast. During lignification process, these monolignols produce a complex three-dimensional amorphous lignin polymer via linkages which lack the regular and ordered units such as those in cellulose or protein. Dixon et al. (2001) reported that the mechanisms that determine the relative proportions of these linkage types in a particular lignin polymer are currently not known.

Lignin content in grass cell wall ranges from 6-12% of the dry weight (Jung and Castler, 2006), and together with lignin composition varies among major groups of higher plants and among species (Whetten et al., 1998). For example, wood lignin mainly contain guaiacyl and syringyl units, whereas those of herbaceous plants contain all three units in significant amounts with different ratios (Billa et al., 1997). The ratio of S/G in grasses increases as the plant ages (Grabber, 2004) and H occurs in grasses in a very low amount (Grabber, 2004). Lignin does not exist in plant tissue as an independent polymer but it is bonded with other polymers such as cellulose and hemicelluloses forming lignin-carbohydrates complexes (LCC). The LCC from herbaceous crops are structurally different from those in woods and contain ferulic bridges between lignin and carbohydrates (arabinoxylans) via ester-linked ferulic acids. Thus, these LCC are often referred to as lignin/phenolics-carbohydrates complexes. Ferulic acid is attached to lignin with ether bonds and to carbohydrates with ester bond.

Unlike in woody plants, lignin biosynthesis is not well studied in grasses; the structure and properties of lignin, and the precise interrelationships with other cell wall components are not well understood (Buranov and Mazza, 2008). Recently however, efforts are being made to understand lignin biosynthesis in grasses because of the growing interest in bioethanol production from renewable biomass. Among grasses, maize received the most attention in genetic and genomic studies related to cell wall lignification and degradability (Barrierè et al., 2003), but with the complete sequence of rice and sorghum genome a better understanding of pathways and genes involved will be accumulated.

There is a growing interest of enhancing plants for conversion to biofuel to replace petroleum while reducing greenhouse gases. Among the potential plants such as maize, sugarcane and switchgrass, sorghum has attractive characteristics for development as a model biomass crop (Sarath et al., 2008). A successful packaging of desired genes while deleterious genes are excluded has been done in many crops through conventional plant breeding and molecular

plant breeding. Improvements in biomass quality in sorghum is dependent in genetic variability of the species, the heritability of the traits, intensity of selection, and the ability of the plant breeder to understand the genetic architecture controlling these traits. The objectives of this chapter were to 1) identify chromosomal regions linked to fibre-related traits, 2) to identify epistatic regions associated with these traits and 3) identify stable quantitative trait loci.

5.2 Material and methods

5.2.1 Phenotyping analysis

The population of Recombinant inbred lines (RIL) derived from a cross between grain sorghum (M71) and sweet sorghum (SS79) discussed in chapter two was used for this study. Details about experimental trials and data collection have been discussed in chapter three. The data was analyzed using SAS[®] 9.0 software. Descriptive statistics was analysed using the LSmeans command and the analysis of the variance was done using the TYPE III method of the general linear model.

5.2.2 Genotyping

DNA extraction was performed according to Doyle and Doyle (1990). A total of 200 SSR primer pairs reported by Bhatramakki et al. (2000), Dean et al. (1999), Kong et al. (2000), Schloss et al. (2002), and Taramino et al. (1997), and 20 Est-SSRs reported by Srinivas et al., (2008; 2009) were used to screen parents of the mapping population. The primers were fluorescently labelled with the M13 tailings. The reaction protocol and thermal cycler touch-down protocol were carried in the same way as in chapter 4. A total of 256 AFLP primer combinations were also included. The AFLP amplification was carried out following Vos et al. (1995) and using the Gibco AFLP core reagents. The sequence of the primers carried the amplification steps were as follows: +0 EcoR1, +0 Mse1, +1 EcoR1, and +1 Mse1 were 5'-GAC TGC GTA CCA ATT C-3', 5'-GAT GAG TCC TGA GTA A-3', 5'-GAC TGC GTA CCA ATT CA-3', and 5'-GAT GAG TCC TGA GTA AC-3 respectively as reported by Uptmoor et al. (2003). The samples were denatured at 94 °c for 5 min and then 2 µl samples was loaded in 8% polyacrylamide gel. Amplification products were separated and visualized using LI-COR model 4200 (MWG Biotech, Ebersberg, Germany) automated DNA sequencing system. This system can detect amplification products that were labelled with two

different dyes: IRD-700 nm and IRD-800 NM. The construction of the linkage maps and analysis of the QTL is similar to what is reported in chapter 4.

5.3 Results

5.3.1 Phenotyping results

The average performance of parents and recombinant inbred lines across locations are presented in table 5.1. Quantification of fibres using van Soest et al. (1991) methods and NIRS showed that the male parent M71 has higher cellulose, hemicellulose, NDF, ADF and ADL content than the female parent SS79. Although cell wall components were not quantified in these parents before, this was expected because sweet sorghum genotypes have been selected for easy-to-peel in Limpopo. In terms of biomass components and/or total biomass, SS79 weighed more than M71. The average performance of recombinant inbred lines fell within the range of the parents but transgressive segregation was also observed.

Analysis of the variance using general linear model showed that genotypes and environments were highly significant ($P < 0.001$) in all the traits investigated except for dry stalk mass. There was a significant interaction between genotype and environment ($P < 0.001$) in most traits except in NDF and dry stalk mass. Although harvest dates were expected to be the significant source of variation in most traits, it was significant only in ADF, ADL and cellulose. The coefficient of variation ranged from 6.14 to 34.37% where ADF had the lowest and fresh leaf mass had the highest.

Higher broad sense heritability were observed in ADF ($H^2=0.84$), cellulose ($H^2=0.82$), stripped stalk mass ($H^2=0.79$), and NDF ($H^2=0.70$) and the lowest broad sense heritability were observed in dry stalk mass ($H^2=0.005$), dry biomass ($H^2=0.13$), and hemicellulose ($H^2=0.24$). Traits that showed higher heritability also showed higher repeatability. Low heritability and repeatability can be the results of sampling error coupled with environmental effects and that can lessen success in enhancing these traits.

5.3.2 QTL analysis

Figure 5.1 shows the significant QTL detected by CIM and supplementary table 5.3 provides the full information of the additive QTLs detected. Environments 01, 02, 03, and 04

throughout the chapter represent the field experiment in Gross Gerau 2007, Gross Gerau 2008, Rauischholzhausen 2007, and Rauischholzhausen 2008, respectively.

5.3.2.1 Acidic detergent fibres

A total of seventeen additive QTLs were detected on all chromosomes associated with acidic detergent fibres. Out of these QTLs, eight showed a significant influence to the phenotype either by using the overall mean or in each single environment. These QTLs had a LOD score range of 3.14 and 7.99. The QTL with a LOD score of 7.99 on SBI-03 position 86 cM contributed the highest phenotypic variation of 17.8% followed by the QTL with LOD score 6.91 on SBI-06 position 58 cM, which contributed 15.6%. Higher contribution of additive effects among the detected QTL was found on SBI-06 left flanked by E35M49-205 of 1.141 g/dry weight and SBI-07 left flanked by E31M59-202 of 1.21 g/dry weight. The positive signs indicate the allele contribution by the M71 parent. In general, the additive effects of all the QTLs were having either negative or positive effect, indicating the contribution of both parents in decreasing and increasing acidic detergent fibres.

5.3.2.2 Neutral detergent fibre

A total of seventeen additive QTLs were detected associated with NDF distributed on all chromosomes. These QTLs were detected at LOD score ≥ 4.0 and significant QTLs having LOD score between 2.5 and 4.0 were reported as suggestive. Six QTLs and two suggestive QTLs were significant using either overall mean or single environments. The phenotypic variation explained by the detected QTLs ranged from 7.5 to 18.8% where the lowest had LOD score 2.98 on SB1-01 flanked by Xtxp279 and the highest had LOD score 8.18 on SBI-07 flanked by E31M59-202. QTLs showing high additive effects were detected on chromosome 02, 06, and 07 valued at 1.21, 1.79 and 1.23 g/dry weight respectively. These effects showed the contribution of the M71 parents in increasing neutral detergent fibre. In all detected QTLs however, both negative and positive additive effects were observed showing the contribution of both the parental alleles in this trait.

5.3.2.3 Acidic detergent lignin

A total of five out of sixteen detected additive QTLs were significant in influencing the phenotype across the environments or using the overall data set. The significant QTLs were detected on SBI-4b, SBI-06, SBI-07, SBI-08, and SBI-09. The QTL detected in SBI-09 with the left flanking marker Xtxp230 had LOD score of 2.79 and therefore considered as

suggestive QTL. The phenotypic variation explained by all detected QTLs ranged from 6.6 to 18.9% where the lowest had a LOD score of 2.79 on SBI-09 at position 36 cM and the highest had a LOD score 4.28 on SBI-09 at position 130 cM. The additive effects of the detected QTLs were variable showing contribution by both parents.

5.3.2.4 Cellulose content

A total of nineteen additive QTLs were detected at LOD score ≥ 4.0 distributed on nine chromosomes except on chromosome 1. The biggest cluster was observed on chromosome 6, where 5 QTLs were detected by the overall environmental mean and significant influenced the phenotype either in combined or single environment. Twelve out of nineteen QTLs significantly influenced the phenotype. The phenotypic variation explained by these QTLs ranged from 7.4 to 17.6%. A total of 67% of the detected QTLs showed a positive additive effect which showed the contribution of the male parent in increasing cellulose quantity.

5.3.2.5 Hemicellulose content

At LOD score ≥ 3.0 , nine QTLs associated with hemicelluloses content were detected. These QTLs were distributed across eight chromosomes except on chromosome 1 and 9. Six out of nine detected QTL significantly influenced the phenotype in three environments (1, 2 and 3) and also with the overall mean. The QTL that had LOD score of 10.99 on SBI-06 at position 28 cM explained the highest phenotypic variation of 23.6%. This QTL showed the additive effects of 0.448 g/dry weight. Overall, the detected QTL explained the phenotypic variation that ranged from 7.4 to 23.6%. Only two out of nine detected QTLs showed the allele contribution by the female parent, which signified the allele contribution by the male parent.

5.3.2.6 Fresh leaf mass

A total of six QTLs associated with fresh leaf mass were detected on SBI-02, SBI-03, SBI-04a, SBI-04b, SBI-06, and SBI-08. Out of these six QTLs, four were significantly influenced the phenotype in overall environments, and in environment 1, 2 and 3. The phenotypic variation explained by these QTL was 54.6%. Only one QTL showed positive additive effects while the other five showed negative additive effects.

5.3.2.7 Stripped stalk mass

A total of fifteen additive QTLs associated with stripped stalk mass were detected at LOD score > 4.0 , but significant QTLs at LOD < 4.0 were reported as suggestive QTLs. These QTLs

were distributed on nine chromosomes except on chromosome eight. The biggest clusters were observed on SBI-01 and SBI-06 with three QTLs each, where SBI-06 consistent of only significant QTLs. Ten significant QTLs were detected using the overall environmental means and each environment. The phenotypic variation explained by these QTLs ranged from 7.9 to 17.4%. The additive effects of these detected QTLs showed that both parents contributed alleles for decreasing or increasing stripped stalk weight.

5.3.2.8 Dry stalk mass

A total of seven out of eleven detected additive QTLs significantly influenced the dry stalk mass in all the three environments using the overall environmental means of dry stalk mass. These QTLs were distributed on six chromosomes. The phenotypic variation explained by all detected QTLs ranged from 7.0 to 16.7%. The highest phenotypic variation was explained by QTL on SBI-06 with LOD score of 7.46 position 14 cM followed by the QTL on SBI-07 with LOD score of 5.72 at position 82 cM. These two QTLs significantly influenced the dry stalk mass using the overall environmental means and other three environments. Two out of eleven QTLs showed positive additive effects, showing that most of the alleles were contributed by SS79, parent with high stalk mass.

5.3.2.9 Fresh biomass

A total of thirteen additive QTLs were detected where ten significantly influenced the fresh biomass across all environments. These significant QTLs were distributed on all chromosomes, where chromosome 1 was found with three QTLs, the highest number detected. The phenotypic variation explained by these QTLs ranged from 6.1 to 24.7%. The additive effects of the detected QTLs ranged from -0.04 to -52.56 g/plant favouring SS79, the high biomass parent and 2.84 to 57.50 g/plants favouring the M71, the low biomass parent.

5.3.2.10 Dry biomass

Thirteen QTLs and four suggestive QTLs associated with dry biomass were detected in 8 chromosomes. The biggest clusters were detected on chromosome 1 and chromosome 2, where each consisted of 4 QTLs. Six out of seventeen QTLs were found significant in combined environments or environment 1 or 2. None of the QTLs significantly influenced dry biomass in at least 2 environments or in environment 3 and 4. The phenotypic variation explained by these QTLs ranged from 7.0 to 12.5%. The additive effects of these QTLs showed that both parents contributed to the dry biomass.

5.3.3 QTL x Environmental interaction and epistasis

QTL x environment interaction is an important component for quantitative traits. QTLs detected in one environment but not in another might indicate QTL x environmental interaction (Veldboom and Lee, 1996). A total of twenty-seven QTLs showing a significant QTL x environmental interaction were detected in most traits except in hemicelluloses, NDF, and dry biomass (table 5.3, supplementary results). Dry biomass was quantified in 2 environments, and probably that is why no QTL x environmental interaction was detected. A total of seven QTLs showing significant QTL x environmental interactions were detected in stripped stalk mass, five were detected in dry stalk mass and fresh biomass, four were detected in fresh leaf weight, two in cellulose and ADF, and one in ADL. QTLs that showed pleiotropic effects were the ones that showed significant QTL x environmental interaction. For example, the QTL on SBI-06 at position 14 cM showed significant QTL x environment interaction in most traits.

Epistatic effects is important for studying complex traits and genetic model for QTL mapping, assuming no epistasis could lead to a biased estimation of QTL parameters (Cao et al., 2001). A total of seventeen significant digenic additive x additive interaction pairs were detected in cell wall components studied and related agronomic traits (table 5.2). Most of the additive x additive interactions detected were involving SBI-01, showing the importance of this chromosome in digenic pairs. A total of four digenic epistatic QTLs were detected associated with hemicellulose, and fresh biomass, three digenic epistatic QTLs for dry stalk mass, two digenic epistatic QTLs for stripped stalk mass and cellulose, one digenic epistatic QTL for ADL and NDF. There were no significant additive x additive interaction detected for ADF, dry biomass, and fresh leaf mass. Cao et al., (2001) reported that detecting epistatic of pair wise QTLs for a trait depends on the developmental stage of the plants. They concluded that interaction between QTLs and background or modifying loci might be the prevalent form of epistasis affecting the behaviour of quantitative traits.

The total phenotypic variation explained by interactions for ADL was 3.3% from one digenic interaction, NDF was 6.7% from one digenic interaction, stripped stalk mass was 10.6% from two digenic interactions, cellulose was 12.4% from two digenic interactions, dry stalk mass was 14% from three digenic interactions, hemicelluloses was 30.3% from 4 digenic interactions and 16.7% from 4 digenic interactions. As observed in sugar related traits

(chapter 4), the total phenotypic variations explained by the interactions were lower than that of corresponding main effects. The importance of additive x additive interaction effects in total genetic effects may be trait-dependent.

5.3.4 Co-localization and pleiotropic effect

Composite interval mapping detected more QTLs that co-localized across the ten chromosomes (supplementary table 5.3, and Figure 5.1). Most traits were detected more than once in one chromosome. The highest co-localization was observed on chromosome 06, where the QTLs of all the traits were detected summing to twenty-five QTLs. The lowest was found on chromosome 10, where only one QTL per trait was observed summing to seven QTLs. Most of the chromosomes harboured QTLs ranging from 14-18 in total.

QTLs showing multiple effects were detected on different chromosome (Supplementary table 5.3, and figure 5.1). The biggest clusters with significant QTLs were observed on SBI-06 at position 14, SBI-07 at position 77 and SBI-03 at position 83. The QTL on SBI-06 was detected in most of the traits except in hemicelluloses and fresh leaf mass. The phenotypic variation explained by this QTL ranged from 8.4 to 17.4%, where the lowest was seen in cellulose and the highest in stripped stalk mass. The QTL on SBI-07 was detected in most traits except for dry biomass, stripped stalk mass and fresh leaf mass. The phenotypic variation explained by this QTL ranged from 9.0 to 18.9%.

In the two clusters observed in SBI-06 and SBI-07, the alleles for the cell wall components (cellulose, hemicelluloses, ADL, ADF, NDF,) were contributed by M71 parent, while the alleles for agronomic mass (fresh and dry biomass, stripped stalk mass and dry stalk mass) were contributed by the SS79 parents. This indicated that both parents contributed in increasing fibre content and agronomic mass, considering that M71 has high fibre content and SS79 has high biomass. ADL was also observed in this region, which is not desired by the breeder. Lignin does not appear in the cell wall alone, it is cross-linked with cellulose and hemicelluloses, thereby prohibiting cellulose fibre from swelling, which reduces surface area access to the enzyme. Cellulases irreversibly adsorb to lignin, thus preventing their action on the cellulose. This necessitates the use of more enzymes to saturate these non-productive sites adsorption sites on the surface of biomass and lead to high enzyme cost. However, declaring QTLs as pleiotropic is difficult since completely linked QTLs can behave in that way. This

can be validated by conducting a large number of crosses that may or may not produce recombinants.

5.4 Discussion

The increasing demand of food and the rising concerns about climate change and energy security pushes *Sorghum bicolor* to the top of global agendas. The main attraction is its ability to provide renewable energy products, industrial commodities, as well as food and animal feed. Carpita and McCann (2008) reported that sorghum, as well as maize show the potential to provide abundant and sustainable resources of lignocellulosic biomass for the production of biofuel and can also serve as genetic models for the improvement of C₄ bioenergy grasses. Both crops are annual grasses and give the farmers flexibility, carry C₄ photosynthesis, and maize has historical depth of genetic knowledge and sorghum has been recently sequenced.

In this study, we quantified and detected QTLs that are linked with cell wall content and biomass content. There is limited literature on the study at the moment. However, a recent work by Murray et al. (2008) quantified and detected QTLs that are linked with stem fibre content, leaf fibre content and agronomical traits associated. As reported in our study, they also found co-localization of QTLs associated with fibre-related traits. Unlike in our study, their QTLs were distributed across nine linkage groups where no QTL was detected in linkage group two. In our study, using additive x additive model resulted in detection of more QTLs having high additive effect. The additive effect is important for breeding because it can positively aid selection. Maize has received lot of attention for fibre content research in many parts of the world. Before the biofuel storm, maize cell wall components were studied for digestibility in ruminants (Moore and Hatfield, 1994, Cardinal et al., 2003). In maize, Cardinal et al. (2003) reported clustering and overlapping of QTLs for cell wall components such as lignin, NDF and ADF in the stalk and Barrierè et al. (2008) reported co-localization of the QTLs associated with lignin content, lignin monomeric composition and cell digestibility. In our study, co-localization and overlapping of genes were observed. This suggests that the genes of cell wall components may be linked in the genome.

The role of epistasis in crop improvement is captivating because there is clear evidence that genes interact. Although epistasis is difficult to estimate, understanding the amount and type of epistasis present can enhance the reliability of predictions and the design of plant breeding

programs (Lamkey and Lee, 1993). It is possible to detect QTLs with additive and epistatic effects, and QTL x environmental interactions using PLABQTL. In our study, epistasis was detected in seven out of ten traits studied. The inability to detect epistasis cannot be taken as evidence for the absence of epistasis in those traits, because of the cancelling of epistatic effects among loci, genotype x environmental interactions and sample size. And estimates using the analysis of the variance approach can generally yield non significant QTLs (Lamkey and Lee, 1993), a method used in our study. The genetic basis of heterosis has been ascribed to overdominance at a single locus, or dominance complementation, and multilocus epistasis. The results of additive x additive interactions in our study showed the contribution of both parents, and the additive effect per trait was not consistent, it differed among pairs. However, it is desired to fix additive x additive effect that increases lignocellulosic biomass for marker-assisted selection.

The main goal of plant breeders in bioenergy project is to develop cultivars with high biomass yield, cellulose and hemicelluloses content, which will be stable across years and environments. However, the plants interact with environment from germination to maturity, and this interaction differs from one genotype to another and this phenomenon complicates plant breeding. The identification of non environment specific QTLs or QTLs with minor QTL x environmental effect should be useful for marker-assisted selection. In our study, stability of the QTL across environments was analysed. It was observed that some QTLs are stable across environments and others are influenced largely or less by the environment. QTL x environmental interaction was mostly observed in agronomic traits than in cell wall components and in QTLs that showed high pleiotropic effects. In pleiotropic effects, the interaction was trait dependant.

5.5 Conclusion

The ultimate goal to identify QTL associated with fibre-related traits in this study is to enhance sorghum as the potential bioenergy source. Pronounced phenotypic variation was observed among the genotypes and significant QTLs were detected among traits. Many QTLs associated with these traits co-localized and clustered mostly on chromosome 1, 3, 6, and 7. Digenic epistatic interaction was found for most of the traits and also some QTLs that showed multiple effects. Additive QTLs were detected across environments and significant QTL x environmental interactions were observed, too.

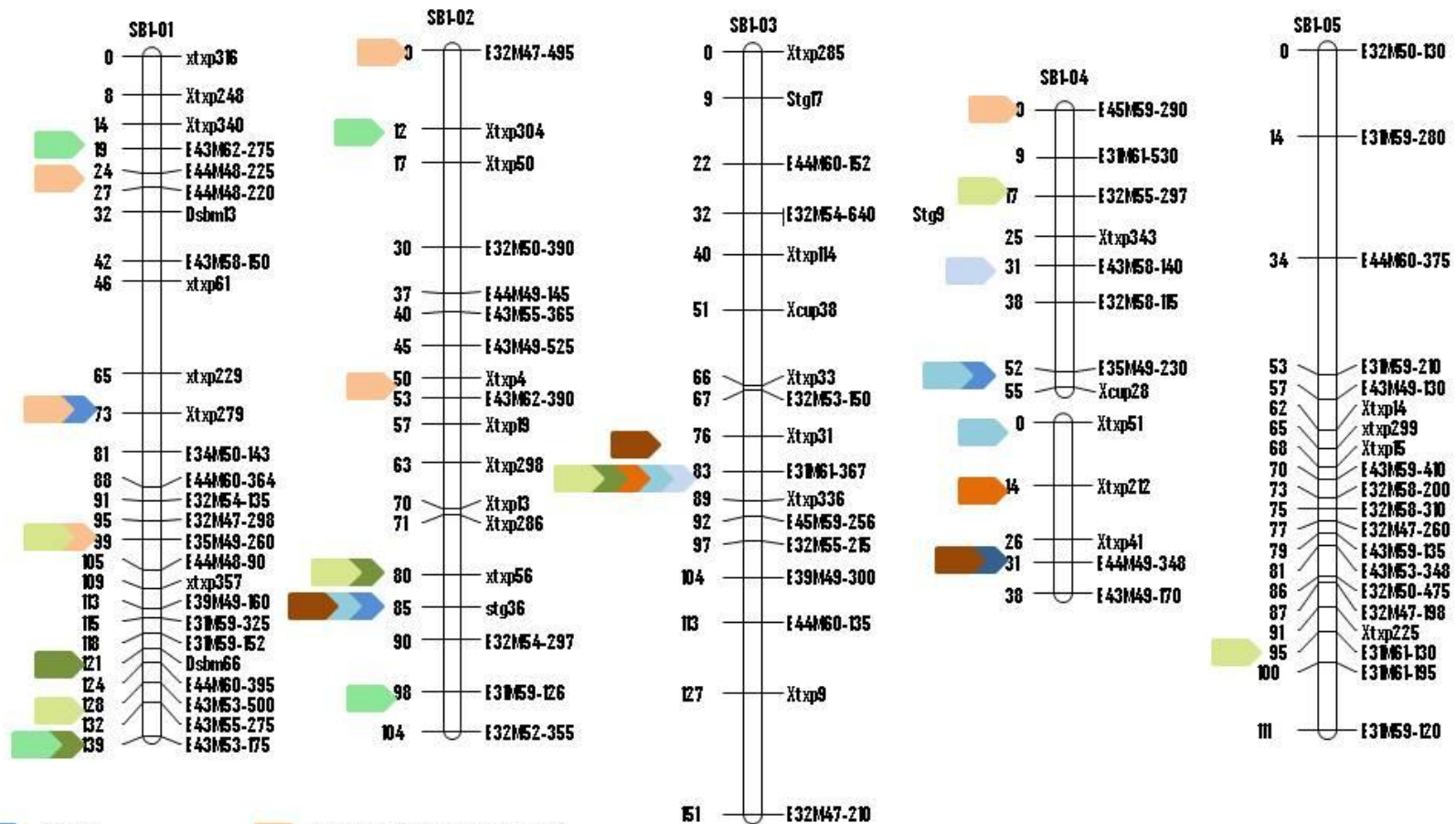
5.6 Acknowledgement

The authors would like to gratefully acknowledge the valuable technical assistance of Nelly Weiss, Swetlana Renner, and Adriana Ochoa Fandiño for fibre content analysis, as well as the financial support of the Deutscher Akademische Austausch Dienst (DAAD) to Amukelani Lacrechia Shiringani.

Table 5.1 Descriptive statistics and analysis of variance

Traits	Parents		RECOMBINANT INBRED LINES				Anova with GLM						
	SS79	M71	Mean	std dev	min	max	Genotype	En v	G*E	Hardate	CV	H ²	R ²
ADF g/dry weight	29.72	36.11	33.67	2.37	28.27	40.65	***	***	***	***	6.14	0.84	0.92
NDF g/dry weight	5.67	7.50	52.01	3.19	44.93	60.81	***	***	***	NS	10.64	0.7	0.76
ADL g/dry weight	46.58	57.99	7.23	0.64	5.73	8.86	***	***	***	***	10.35	0.67	0.84
Cellulose g/dry weight	24.03	29.23	26.44	2.03	22.03	33.12	***	***	***	***	6.89	0.82	0.91
Hemicellulose g/dry weight	16.53	19.90	18.34	1.01	16.11	21.93	**	***	NS	NS	27.32	0.24	0.24
Fresh leaf mass g/plant	122.48	59.85	106.96	22.61	60.42	192.59	***	***	***	NS	34.37	0.51	0.61
Stripped stalk mass g/plant	642.54	191.57	375.86	140.02	118.04	754.9	***	***	***	NS	33.84	0.79	0.9
Dry stalk mass g/plant	73.65	35.28	53.9	21.48	0	123.26	NS	NS	NS	NS	-	0.005	0.004
Fresh biomass g/plant	1102.74	472.72	570.76	130.1	277.46	944.09	***	***	***	NS	33.49	0.34	0.52
Dry biomass g/plant	177.94	81.02	92.34	22.82	25.59	158.67	***	***	***	NS	28.29	0.13	0.36

GLM: General linear model ADF: Acidic detergent fibre NDF: Neutral detergent fibre ADL: Acidic detergent lignin Std dev: standard deviation min: minimum max: maximum Env: environment G*E: genotype x environment Hardate: Harvest date CV: coefficient of variation H²: Heritability R²: Repeatability *** significant at 0.001 probability level NS non significant level



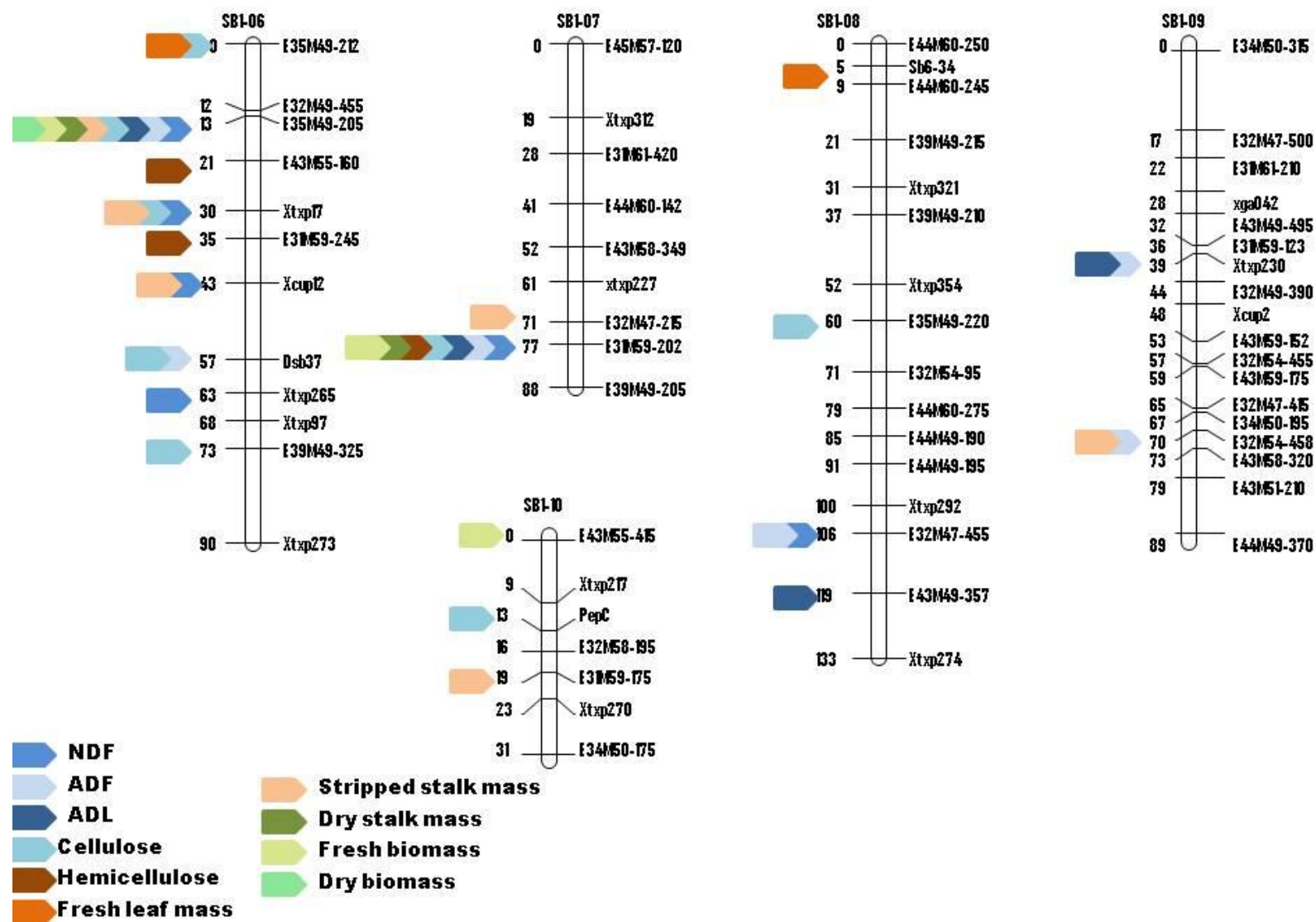


Figure 5.1. Genetic linkage map of *Sorghum bicolor* L. Moench based on 157 SSR, AFLP and EST-SSRs. The map covers 1029.08 cM with the average distance of 6.55 cM. The QTLs presented are detected by composite interval mapping and are significant either in combined environments or in each single environment or both. The details of the presented QTLs which include the position, additive effect and phenotypic variation explained are in table 5.3, supplementary results.

Table 5.2. Digenic epistasis associated with fibre-related traits

Traits	Chromosome	Left marker	chromosome	Left marker	Additive effect	Mean	R ² %
ADF g/dry weight	none						
NDF	SBI-01	E43M53-500	SBI-03	E31M61-367	-0.766		6.7
ADL g/dry weight	SBI-02	E32M47-495	SBI-07	E31M59-202	-0.122	7.264	3.3
Cellulose g/dry weight	SBI-01	E35M49-260	SBI-05	E44M60-375	0.495	26.586	7.4
	SBI-01	Xtxp279	SBI-06	Dsb37	-0.42		5.0
Hemicellulose	SBI-02	Stgnhsbm36	SBI-05	E31M59-210	-0.274	18.457	7.6
g/dry weight	SBI-4b	E44M49-348	SBI-05	E31M59-210	-0.233		5.4
	SBI-4b	E44M49-348	SBI-06	E43M55-160	-0.332		11.3
	SBI-07	E31M59-202	SBI-08	E39M49-215	0.24		6.0
Fresh leaf mass g/plant	none						
Stripped stalk mass g/plant	SBI-01	E35M49-260	SBI-09	E43M59-152	31.277	362.290	5.1
	SBI-01	Xtxp279	SBI-06	Xtxp17	41.147		5.5
Dry stalk mass g/plant	SBI-01	E43M53-175	SBI-08	Xtxp292	-4.356	51.112	3.4
	SBI-02	Xtxp56	SBI-07	E31M59-202	-4.517		3.8
	SBI-4b	E44M49-348	SBI-06	E43M55-160	5.378		6.8
Fresh biomass g/plant	SBI-01	E43M53-175	SBI-09	E31M59-123	29.143	559.48	4.3
	SBI-01	E35M49-260	SBI-02	Xtxp50	-36.24		4.3
	SBI-01	E35M49-260	SBI-09	E32M47-415	29.719		4.4
	SBI-01	Xtxp248	SBI-10	E43M55-415	-26.96		3.7
Dry Biomass g/plant	none						

ADF: Acidic detergent fibre NDF: Neutral detergent fibre ADL: Acidic detergent lignin R²%; phenotypic variation explained in percentage

5.7 References

- Barrierè, Y., C. Gulliet, D. Goffner, and M. Pichon. 2003. Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. A review. *Anim. Res* 52:193-228
- Barrierè, Y., J. Thomas, and D. Denoue. 2008. QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838 X F286. *Plant Sci* 175:585-595
- Billa, E., D. P. Koullas, B. Monties, and E. G. Koukios. 1997. Structure and composition of sweet sorghum stalk components. *Ind. Crops Prod* 6:297-302
- Bhatramakki, D., J. Dong, A. Chhabra, and G. Hart. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* 43:988-1002
- Buranov, A.U., and G. Mazza. 2008. Lignin in straw of herbaceous crops. *Ind. Crops Prod* 28:237-259
- Cao, G.Q., J. Zhu, C.X. He., Y.M. Gao, and P. Wu. 2001. QTL analysis for epistatic effects and QTL x environment interaction effects on final height of rice (*Oryza sativa* L.). *Acta. Genet. Sin* 28:135-143
- Cardinal, A.J., M. Lee, and K.J. Moore. 2003. Genetic mapping and analysis of quantitative trait loci affecting fibre and lignin content in maize. *Theor. Appl. Genet* 106:866-874
- Carpita, N.C., and M.C. McCann. 2008. Maize and sorghum: genetic resources for bioenergy grasses. *Trends Plant Sci* 13:415-420
- Cosgrove, D.J. 2005. Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol* 6:850-861
- Dean, R.E., J.A. Dahlberg, M.S. Hopkins, S.E. Mutchell, and S. Kresovich. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. National sorghum collection assessed with simple sequence repeats (SSR) markers. *Crop Sci* 39:1215-1221
- Dixon, R.A., F. Chen, D. Guo, and K. Parvathi. 2001. The biosynthesis of monolignols: a "metabolic grid", or independent pathways to guaiacyl and syringyl units? *Phytochem* 57:1069-1084
- Dhugga, K.S. 2007. Maize biomass yield and composition of biofuel. *Crop Sci* 47:2211-2227
- Doyle, J.J., and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15
- Grabber, J.H., J. Ralph, C. Lapierre, and Y. Barrierè. 2004. Genetic and molecular basis of grass cell-wall degradability. I. Lignin-cell wall matrix interactions. *C.R. Biologies* 327:455-465

- Hazen, S.P., J.S. Scott-Craig, and J.D. Walton. 2002. Cellulose synthase-like genes of rice, *Plant Physiol* 128:336-340
- Jung, H.G., and M.D. Castler. 2006. Maize stem tissues: Cell wall concentration and composition during development. *Crop Sci* 46:1793-1800
- Kong, L., J. Dong, and G.E. Hart. 2000. Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple sequence repeats (SSRs). *Theor. Appl. Genet* 101:438-448
- Lamkey, K.R., and M. Lee. 1993. Quantitative genetics, molecular markers, and plant improvement. In B.C. Imrie and J.B. Hacker (ed.) *focused plant improvement: Towards responsible and sustainable agriculture*. Proc.10th Australian plant breeding conference., Gold Coast, 18 -23 April 1993. Organising committee, Australian convention Travel Service: Canberra:104-115
- Lerouxel, O., D.M. Cavalier, A.H. Liepman, and K. Keegstra. 2006. Biosynthesis of plant cell wall polysaccharides-a complex process. *Curr. Op. in plant biol* 6:621-630
- Lin, Y., and S. Tanaka. 2006. Ethanol fermentation from biomass resources: current state and prospects. *Appl. Microbiol. Biotechnol* 69:627-642
- Moore, K.J., and R.D. Hatfield. 1994. Carbohydrates and forage quality. In G.C. Fahey et al. (ed). *Forage quality, evaluation, and utilization*. ASA, CSSA, and SSSA, Madison, WI
- Murray, S.C., W.L. Rooney, S.E. Mitchel, A. Sharma, P.E. Klein, J.E. Mullet, and S. Kresovich. 2008. Genetic improvement of sorghum as biofuel feedstock. II. QTL for stem and leaf non-structural carbohydrates. *Crop Sci* 48:2181-2193
- Sarath, G., R.B. Mitchell, S.E. Sattler, D. Funnell, J.F. Pedersen, R.A. Graybosch, and K.P. Vogel. 2008. Opportunities and roadblocks in utilization forages and small grains for liquid fuels. *J. Ind. Microbiol. Biotechnol* 35:343-354
- Schloss, S.J., S.E. Mitchell, G.M. White, R. Kukatla, J.E. Bowers, A.H. Paterson, and S. Kresovich. 2002. Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. *Theor. Appl. Genet* 105:912-920
- Srinivas, G., K. Satish, R. Madhusudhana, and N. Seetharama. 2009. Exploration and mapping of microsatellite markers from subtracted drought stress ESTs in *Sorghum bicolor* (L.) Moench. *Theor. Appl. Genet* 118:703-717
- Srinivas, G., K. Satish, S.M. Mohali, R.N. Reddy, R. Madhusudhana, D. Balakrishna, B.B. Venkatesh, C.J. Howarth, and N. Seetharama. 2008. Development of genic-microsatellite markers for sorghum staygreen QTL using comparative genomic approach with rice. *Theor. Appl. Genet* 117:703-717

- Sticklen, M.B. 2008. Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. *Nature Rev. Genet* 9:433-443
- Taramino, G., R. Tarchini, S. Ferrario, M. Lee, and M.E. Pe. 1997. Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. *Theor. Appl. Genet* 95:66-72
- Taylor, N.G. 2008. Cellulose biosynthesis and deposition in higher plants. *New phytologist* 178:239-252
- Uptmoor, R., W.G. Wenzel, W. Friedt, G. Donalson, K. Ayisi, and F. Ordon. 2003. Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions from southern Africa by RAPDs, AFLPs and SSRs. *Theor. Appl. Genet* 106:1316-1325
- Van Soest, P.J., J.B. Robertson, and B.A. Lewis. 1991. Methods for dietary fibre, neutral detergent fibre, and nonstarch polysaccharides in relation to animal nutrition. *J. Dairy Sci* 74:583-3597
- Veldboom, L.R., and M. Lee. 1996. Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: II. Plant height and flowering. *Crop Sci* 36:1320-1327
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, and M. Kuiper. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 21:4407-4414
- Whetten, R.S., J.J. Mackay, and R.R. Sederoff. 1998. Recent advances in understanding lignin biosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol* 49:585-609

Summary

The increasing demand of food and the rising concerns about climate change and energy security pushes *Sorghum bicolor* to the top of global agendas. The main attraction is its ability to provide food and animal feed as well as renewable energy products, and industrial commodities. The species comprises grain and sweet sorghum types. Sweet sorghum has major advantages compared to other sugar and biomass crops. Firstly, sorghum is a species showing extensive genetic variation, including drought and heat tolerant genotypes, which enable the usage of marginal land that is not suitable for cultivating other crops. Secondly, competition between the use of land for food or for energy is less because the grains can be used for food or feed while stems can be used for biofuel production. In the current study, field trials and related genetic analysis were carried out to identify major chromosomal regions that are linked to biofuel-related traits. An experimental population of 213 RILs from a cross between grain sorghum (M71) and sweet sorghum (SS79) was planted in 5 environments in Germany in the vegetation periods 2007 and 2008. This population segregates for alleles controlling biomass yield, juice content, sugar content, fibre content, and many other traits. The approach to identify these chromosomal regions was (1) to phenotype the RIL population for the mentioned traits, (2) to construct a genetic map from the RILs, and (3) to correlate the phenotypic data and genetic map informations.

The phenotypic data was analysed for individual location and environments (years*location) using mixed model from SAS[®] 9.1 version. Genotypes were treated as fixed effects while all other components were treated as random effects. Regarding, sugar-related traits, we analysed brix, glucose content, sucrose content, and sugar content, measured stem diameter, stem juice weight, plant height and fresh panicle weight, and counted the days to anthesis and the number of tillers per plant. We partitioned the variance components as genetic, location, year, genetic x location, genetic x year, location x year, genetic x location x year, and harvest date. The results showed that although location contributed highly to the phenotype in terms of percentage, it was not a significant source of variance. The interaction between genotypes and environment was observed in most traits except for stem juice weight, fresh panicle weight and glucose content. Genetic variance contributed small but significant to the phenotypic variance. And thus the heritability estimates were moderate to high in all traits except stem juice weight ($H^2=0.18$). Pearson correlation estimates showed that these traits were significantly correlated with each other. Flowering dates correlated with all the traits while

number of tillers only positively correlated with stem juice weight, and negatively correlated with sugar content and flowering dates.

In fibre-related traits, we quantified fibre content traits (acidic detergent fibre, acidic detergent lignin, neutral detergent fibre, cellulose, hemicelluloses), and measured related agronomic traits (fresh leaf mass, stripped stalk mass, dry stalk mass, fresh biomass and dry biomass). We partitioned our variance components as genetic, environment, harvest date and genotype x environment. Genetic variance was significant in all traits except for dry biomass while environment was not significant for all traits. There was a significant genotype x environment interaction in all traits except hemicelluloses. Dry stalk mass, dry biomass and hemicelluloses had low heritability estimates of ≤ 0.24 . Pearson correlation estimates showed that fibre content traits significantly and positively correlated with each other as well as agronomic traits, while there were negative correlations among agronomic and fibre content traits.

A total of 213 RILs were used to construct the genetic linkage map using 157 markers, i.e. simple sequence repeats (SSR), amplified fragment length polymorphism (AFLP) and expressed sequence tag-simple sequence repeats (EST-SSR). The linkage map was compared to other sorghum published maps and linkage groups were assigned according to the recent sorghum nomenclature. The phenotypic data of each trait was averaged to obtain the combined environmental means to detect QTLs. Composite interval mapping (CIM) of the PLABQTL software was used for analysis. The model AA was used to estimate additive x additive QTL interaction, and the command 'environ' was used to estimate significant QTL in each environment and to detect QTL x environmental interaction.

CIM detected a few additive QTLs per trait across the 10 sorghum chromosomes. The additive QTLs co-localized and showed clusters on some chromosome, where the most prominent clusters were observed on chromosomes 1, 2, 6, and 7. QTLs showing multiple effects were observed on all chromosomes except chromosome 5 and 10 and most of these pleiotropic QTLs showed the contribution of the positive allele in each trait. For example, a QTL on SBI-06, position 14, showed the allele contribution of M71 (high cellulose and compact panicle parent) for cellulose and fresh panicle weight, along with the allele contribution of SS79 (high sucrose and biomass parent) for sucrose and fresh biomass. However, QTLs that showed pleiotropic effect also exhibited QTL x environment interaction. Most of the QTLs detected for agronomic biomass (i.e. stripped stalk mass, fresh leaf mass

and fresh biomass) interacted significantly with the environment. Additive x additive interaction was detected in 14 out of 20 traits studied totalling 41 digenic pairs. This indicates the abundance of digenic epistasis across the genome. However, failure to detect QTL x QTL interaction can never be interpreted as being absent but can be attributed to QTL x environment interaction, sampling error or accuracy of statistic methods.

Among other characteristics, stem sugar components and lignocellulosic biomass are the important characteristics of *Sorghum bicolor* to determine it as one of the major bioenergy crops. There is limited literature on sugar and fibre related traits in sorghum at the moment, and this study among other few, serves as one of the base investigations on the topic. The phenotypic variation of the lines, the moderate to high heritability of the desired traits, the significant correlation and pleiotropic QTLs to the positive direction identified, the additive stable QTLs and their interactions detected suggest that it will be possible to develop improved lines and hybrids by breeding and to ultimately enhance the sorghum plant as a bioenergy crop.

Zusammenfassung

Die steigende Nachfrage nach Nahrungsmitteln und regenerativen Energieträgern sowie die Sorge um den fortschreitenden Klimawandel machen Sorghum zu einer hochinteressanten landwirtschaftlichen Pflanze. Seine hauptsächlichen Vorzüge liegen in seiner gleichzeitigen Verwendbarkeit als Nahrungs- und Futtermittel, als Energieträger und als Rohstofflieferant für industrielle Erzeugnisse. Die Art Sorghum umfaßt sowohl Korn- als auch Zuckertypen, wobei sich Zuckersorghum sich gegenüber anderen Zucker- und Biomassekulturen durch einige wesentliche Eigenschaften auszeichnet: Erstens zeigt Sorghum eine enorme genetische Variabilität, die sogar dürre- und hitzetolerante Genotypen umfaßt. Zweitens verringert die gleichzeitige Nutzung des Sorghums für Nahrungs- und Futtermittelgewinnung aus dem Korn und Biokraftstoff aus den Stängeln den Wettbewerb um die Landnutzung durch verschiedene Pflanzen für unterschiedliche Zwecke. In der vorliegenden Arbeit wurden Feldversuche und genetische Analysen mit dem Ziel durchgeführt, verantwortliche chromosomale Regionen zu identifizieren, die mit Eigenschaften gekoppelt sind, welche die Eignung zur Gewinnung von Biokraftstoff beeinflussen. Eine Versuchspopulation von 213 RILs (rekombinante Inzuchtlinien) aus einer Kreuzung zwischen KornSORGHUM (M71) und Zuckersorghum (SS79) wurde in fünf Umwelten in Deutschland während der Vegetationsperioden 2007 und 2008 gepflanzt. Diese Population spaltet in Allele, die den Biomasseertrag, den Saftgehalt, den Zuckergehalt, den Fasergehalt, sowie viele weitere Eigenschaften kontrollieren. Der Ansatz, diese chromosomalen Regionen zu identifizieren, umfaßte (1) die Phänotypisierung der RIL-Population für die genannten Merkmale, (2) die Erstellung einer genetischen Karte der RIL-Linien und (3) die Korrelation der phänotypischen Daten mit den genotypischen Daten.

Die phänotypischen Daten wurden für einzelne Standorte und Umwelten mit SAS[®] Version 9.1 im gemischten Model (*mixed model*) analysiert. Die Genotypen wurden dabei als fixe Effekte betrachtet, während die übrigen Komponenten als zufällige Effekte behandelt wurden. In Hinblick auf zuckerrelevante Eigenschaften analysierten wir den Brix, den Sucrose- und Glucosegehalt, den Zuckergehalt, den Stängeldurchmesser, die Stängelsaftmenge, die Pflanzhöhe, das Frischgewicht der Rispen, die Zeitspanne bis zur Blüte und die Anzahl der Triebe pro Pflanze. Die Varianz-Komponenten unterteilten wir in Genotyp, Standort, Jahr, Genotyp x Standort, Standort x Jahr, Genotyp x Standort x Jahr und Erntedaten. Die Ergebnisse zeigen, daß der Standort trotz seines großen Einflusses auf den Phänotyp keine signifikante Quelle für die Varianz ist. Mit Ausnahme von Stängelsaftmenge, Frischgewicht

der Rispe und Glucosegehalt wurden Wechselwirkungen zwischen Genotyp und Umwelt für alle Merkmale beobachtet. Die genetische Varianz beeinflusst die phänotypische Varianz zwar wenig, jedoch signifikant. Die geschätzten Heritabilitäten waren daher mittel bis hoch in allen Merkmalen mit Ausnahme der Stängelsaftmenge ($H^2=0.18$). Pearsons Korrelationsschätzungen zeigen, daß diese Merkmale auch untereinander signifikant korreliert sind. So korrelieren die Blütedaten mit allen Merkmalen. Die Anzahl der Stängel pro Pflanze ist dagegen zwar positiv mit der Stängelsaftmenge korreliert, aber negativ mit dem Zuckergehalt und Blütedaten.

Bezüglich der Fasermerkmale wurden Fasergehalte, wie Säure-detergenfaser, Neutral-detergenzfaser, Säure-detergenzignin, Zellulose und Hemizellulose und pflanzenbauliche Merkmale, wie Blattfrischmasse, Geschälte-Stängel-Masse, Stängeltrockenmasse, Gesamtfrischbiomasse, Gesamttrockenbiomasse bestimmt. Die Varianz-Komponenten unterteilten wir in Genotyp, Umwelt, Erntezeitpunkt, Genotyp x Umwelt. Die genetische Varianz ist signifikant in allen Merkmalen mit Ausnahme der Gesamttrockenbiomasse. Die Umwelt ist hingegen in keinem der Merkmale signifikant. Die Genotyp-Umwelt-Wechselwirkung ist für alle Merkmale außer Hemizellulose signifikant. Stängeltrockenmasse, Gesamttrockenbiomasse und Hemizellulose weisen niedrige Heritabilitätsschätzungen von ≤ 0.24 auf. Pearsons Korrelationsschätzungen zeigen, daß sowohl jeweils Fasergehaltmerkmale als auch pflanzenbauliche Merkmale signifikant und positiv untereinander korreliert sind. Zwischen diesen Merkmalsgruppen werden allerdings negative Korrelationen beobachtet.

Eine genetische Kopplungskarte wurde aus den genotypischen Daten von 213 RILs und 157 Markern, wie *simple sequence repeats* (SSR), *amplified fragment length polymorphism* (AFLP) and *expressed sequence tag-simple sequence repeats* (EST-SSR) erstellt. Diese Karte wurde mit anderen publizierten Sorghum-Karten verglichen und die Kopplungsgruppen entsprechend der aktuellen Sorghum-Nomenklatur bezeichnet. Die phänotypischen Daten wurden gemittelt, um mittels der kombinierten Umweltdaten QTLs zu detektieren. Die QTL-Analyse wurde mit der Software PLABQTL unter Nutzung der *Composite interval mapping* – Methode (CIM) durchgeführt. Additive x additive QTL-Wechselwirkungen wurden mit dem Model AA geschätzt. Der Befehl “environ“ wurde angewandt, um signifikante QTLs in jeder Umwelt zu ermitteln und QTL-Umwelt-Wechselwirkungen zu ermitteln.

Es konnten einige additive QTLs pro Merkmal in den 10 Sorghumchromosomen detektiert werden. Die additiven QTL ko-lokalisieren und bilden auf einigen Chromosomen Cluster. Die bedeutendsten Cluster finden sich auf den Chromosomen 1,2,6 und 7. Pleiotrope QTLs konnten auf allen Chromosomen außer Chromosom 5 und 10 beobachtet werden. Die meisten von ihnen beeinflussen das jeweils positive Allel der betroffenen Merkmale. Zum Beispiel zeigt ein QTL auf SBI-06, Position 14, den allelischen Einfluß des Elters M71 (Hoch-Zellulose und kompakte Rispe) auf Zellulose und Rispenfrischmasse, sowie den allelischen Einfluß des Elters SS79 (Hoch-Sucrose, Hoch-Biomasse) auf Sucrose und Gesamtfrischbiomasse. QTLs mit pleiotropen Effekten zeigen darüberhinaus QTL-Umwelt-Wechselwirkungen. Die meisten ermittelten QTL für pflanzenbauliche Merkmale (z.B. Geschälte-Stängel-Masse, Blattfrischmasse, Gesamtfrischbiomasse) zeigen signifikante Wechselwirkungen mit der Umwelt. In 14 von 20 untersuchten Merkmalen wurden Additive x Additive Wechselwirkungen beobachtet, insgesamt sind es 41 digenische Paare. Dies verdeutlicht die Verbreitung digenischer Epistasie über das Genom. Dessenungeachtet darf eine nicht nachgewiesene QTL-QTL-Wechselwirkung nicht ignoriert werden. Sie kann auch der Genauigkeit oder Fehlerhaftigkeit der statistischen Methoden und Probennahme, oder aber einer möglichen QTL-Umwelt-Wechselwirkung zugeschrieben werden.

Neben anderen Charakteristika sind Stängelzuckerkomponenten und Lignozellulose-Biomasse diejenigen Merkmale von *Sorghum bicolor*, die zur seiner Bedeutung als Bioenergieträger beitragen. Gegenwärtig ist nur wenig zu Zucker- und Fasermerkmalen in Sorghum publiziert. Die vorliegende Arbeit dient neben anderen wenigen der Grundlagenforschung auf diesem Gebiet. Die phänotypische Varianz der Linien, die mittlere bis hohe Heritabilität der gewünschten Merkmale, die signifikante Korrelation, sowie pleiotrope QTL bezüglich der positiven Allele zusammen mit stabilen QTLs und ihren Wechselwirkungen lassen erwarten, daß die Entwicklung verbesserter Linien und Hybriden durch Züchtung möglich ist, wodurch Sorghum als Bioenergiepflanze gestärkt wird.