Characterisation of sweet sorghum germplasm based on agro-morphological traits, molecular markers and juice related traits

By

McDonald Nundwe

BSc. Agronomy (University of Malawi)

A dissertation submitted in partial fulfilment for the academic requirements for the degree of Masters of Science in Plant Breeding

School of Agricultural, Earth and Environmental Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

January 2018

GENERAL ABSTRACT

There is rising interest for alternative energy sources because of the decline in fossil fuel production and concern over environmental pollution. Currently most biofuel is based on maize and sugar cane as raw materials. However, the use of feedstocks has triggered concerns related to food security, while sugar cane has a high-water consumption and high production requirements amongst other drawbacks. A crop which meets several requirements for biofuel (such as high biomass yield and growth rate, perennial growth, low input requirements, adaptation to the marginal areas, and tolerance to multiple stresses) is sweet sorghum. This study, therefore, aimed at characterising sweet sorghum germplasm using agro-morphological traits and molecular markers (single nucleotide polymorphisms (SNP) during the 2016-2017 summer season at two sites (Ukulinga farm and Potchefstroom). Fourteen quantitative traits were evaluated in an alpha lattice (10 x 5) design with three replications. Analysis of variance for the quantitative traits revealed high levels of genetic variability. This implies that morphological traits differed greatly with a significant G x E interaction across the two sites. Most of the accessions yielded high at Ukulinga than Potchefstroom on juice yield and %brix with a mean yield of 9 605 l/ha and 16.3%, respectively. Most of the accessions studied were early to medium maturing, as evidenced by the mean number of days to 50% flowering (74 days). Analysis of principle components showed that the first four principle components (PC) accounted for 79.12% of the total variation and that some quantitative traits were significantly positively correlated. The estimates for phenotypic coefficient of variation (PCV) were higher than those of genetic coefficient of variation (GCV) for all the traits, indicating the influence of the environment on these traits. However, GCV values for days to 50% flowering, plant height, stalk diameter and stalk yield were very close to PCV. This indicated minimal influence of the environment on the phenotypic expression of these traits. The highest broad sense heritability (H^2) of 99.2% was recorded for plant height. Juice volume had the highest expected genetic advance, expressed as a percentage of mean (GAM) of 131.2%. Days to 50% flowering were significantly and positively correlated to plant height, stalk diameter, number of leaves, stalk yield, brix, juice volume and bagasse weight, but negatively significantly correlated to panicle length, panicle width, panicle weight and 1000 grain weight. Plant height was significantly positively correlated to stalk diameter, number of leaves per plant, stalk yield, juice volume and fresh bagasse weight. Bagasse weight, brix, stalk diameter, plant height and number of leaves had a highly positive and direct contribution on juice yield. Several traits had a highly positively and indirect contribution on juice yield via these traits which had a direct contribution. This revealed primary and secondary traits with practical relevance to sweet sorghum improvement programme, because they showed direct and indirect effects on juice yield (volume), which
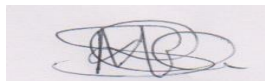
ultimately translates to sugar yield for ethanol production. Kompetitive Allele Specific Polymorphism (KASP) genotyping using 137 SNP markers revealed a considerable level of genetic diversity among the sweet sorghum accessions. Three populations were generated from the analysis. The expected heterozygosity (H_e) values ranged from 0.236 to 0.291 with a mean of 0.266. The mean of effective alleles across populations was of 1.438. The percentage of polymorphic loci ranged from 80.29% to 91.24% with a mean of 86.86%. Dissimilarity indices ranged from 0.000 to 0.583 with a mean of 0.296. The highest dissimilarity index was observed between SA 2193 and SA 2014, which implied a considerable amount of genetic diversity. Accessions were clustered into three main groups based on dissimilarity indices. The study identified SA 4490, SA 2400, SA 4495, SA 2193 and SA 4479 as superior accessions in juice yield. These accessions should be used as parents in sweet sorghum improvement programme.

DECLARATION

I, McDonald Nundwe, declare that:

1. The research reported in this dissertation, except where otherwise indicated, is my original work.
2. This thesis has not been submitted for any degree examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. All sources of information used in this literature review have been duly acknowledged.
5. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   a) their words have been re-written, but the general information attributed to them has been referenced;
   b) where their exact words have been used, then their writing has been placed in italics and inside the quotation marks and referenced.
6. This dissertation does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the referencing sections.
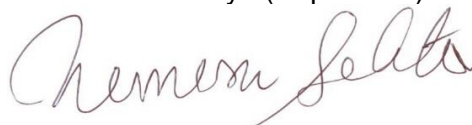
Signed

………………………………………………………………………………………………………
McDonald Nundwe

As the candidate's supervisors, we agree to submission of this dissertation:

………………………………………………………………………………………………………
Dr. Julia Sibiya (Supervisor)

………………………………………………………………………………………………………
Dr. Nemera Shargie (Co-Supervisor)

ACKNOWLEDGEMENTS

DEDICATION

To God Almighty, for the gift of life. For guiding me, and for giving me strength and encouragement to remain resilient and focused. "To Him we give praise and glory."

To my mother, Mrs. Beatrice Nundwe (nee: Mwajulo) for her love, support and advice in my upbringing. To my late father, Sydney Isaac Nundwe, for the few years that we shared, before he responded to God's call. "We loved you, but Heavens loved you the most." To my late aunt, Mrs. Georgina Chaloza (nee: Nthondo), for inspiring and challenging me to go an extra mile. To my grandmother, Mrs. Grace Nthondo for spending a significant part my life with her. To my siblings and family members, thanks for being there for me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMOVA	Analysis of molecular variance
AFLPs	Amplified fragment length polymorphisms
ANOVA	Analysis of variance
ARC-GC	Agriculture Research Council-Grain Crops
CH_4	Methane
CO_2	Carbon dioxide
CV	Coefficient of variation
DNA	Deoxyribonucleic acid
F	Fixation index
FAO	Food and Agriculture Organization
GA	Genetic advance
GAM	Genetic advance expressed as a percentage of mean
GCA	General combining ability
GCV	Genotypic coefficient of variation
GGE	Genotype main effects and genotype by environment interaction
GHG	Greenhouse gases
GEI	Genotype by environment interaction
H^2	Broad sense heritability
H_e	Expected heterozygosity within genotypes
H_o	Observed heterozygosity within genotypes
I	Information index
KASP	Kompetitive allele specific polymorphism
LSD	Least significant difference
MAB	Marker assisted breeding
MANOVA	Multivariate analysis of variance

MAS	Marker assisted selection
N_2O	Nitrous oxide
N_a	Number of alleles per locus
N_e	Number of effective alleles per locus
PC	principal component
PCA	Principle component analysis
PCR	Polymerase chain reaction
PCV	Phenotypic coefficient of variation
PI	Photoperiod insensitive
PIC	Polymorphic information content
PL	Polymorphic loci
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
SCA	Specific combining ability
SE	Standard error
SNPs	Single nucleotide polymorphisms
SPSS	Statistical package for social scientists
SSRs	Simple sequence repeats
USA	United States of America
WFP	World food programme
σ_g^2	Genotypic variance
σ_p^2	Phenotypic variance

CHAPTER 1

Introduction

Sweet sorghum and grain sorghum (*Sorghum bicolor* L. Moench) belong to the same domesticated species but the former has been selected for its high sucrose accumulation in the parenchyma of the juicy stems (Rooney et al., 2007). The stems are preferred for food grade syrup, fresh chewing, and alcohol production in countries such as Brazil, India, and Africa (Vaidyanathan et al., 1987). Under favourable conditions, a sweet sorghum crop can produce up to 13.2 metric tonnes per hectare of total sugars, which can be translated to 7682 litres of ethanol per hectare (Murray et al., 2009).

1.1 History of sweet sorghum

Documented evidence suggests that sweet sorghum was introduced into eastern Africa from Ethiopia, around 200 AD, through the local tribes, who cultivated it primarily for grain, while the sweet cane was chewed for pleasure and nutrition (Doggett, 1988). Later, the Bantu tribes migrated to the Savanna region of western and southern Africa. Here, the grain was used mostly for brewing beer. The Bantu tribes then moved the sorghum crop through their expansion from southern Cameroon region around first century AD, and the southern border of the Congo forest belt (Doggett, 1988). The current sorghums of central and southern Africa are closely related to eastern Africa, especially those of Tanzania, and are more distantly related to those of west African varieties because of the equatorial forests that acted like a barrier to their distribution (FAO, 1995).

During the first millennium BC, sorghum was shipped to India from eastern Africa as food by the chow traffic which operated for about 3000 years between east Africa (the Azanean Coast) and India through the Sebaean Lane in southern Arabia (Doggett, 1988). The Indian sorghum cultivars are closely related to those present in north eastern Africa and the coast between Cape Guardafui and Mozambique. Some literature suggests that this crop might have been distributed along the coast of southeast Asia and around China in the beginning of the Christian era (Doggett, 1988). However, it is also suggested that sorghum might have been introduced in China much earlier by the silk trade routes (FAO, 1995). Later it was introduced to the western world through Asia. Sorghum was documented in European botanical literature in 1542 and was referred to as Sorghi, this was the same name used in India (FAO,1995).

Sorghum was then carried to the Caribbean islands and other Latin American countries from west Africa through the slave trade and by navigators plying the Europe-Africa-Latin America

trade route in the early 17th century as another source for sugar production (FAO, 1995; Srinivasa et al., 2013).

1.2 Why sweet sorghum for biofuel?

There is a rising need for substitute energy sources due to the decline in fossil fuel production. Currently, the majority of biofuel is based on maize and sugar cane as raw materials. However, continued usage of these food crops has caused concerns linked to food security (Serna-Saldívar et al., 2012). The relative rise in food prices has also been attributed mainly to the use of maize for bioethanol apart from biotic and abiotic factors or changes in global consumption (WFP., 2008). Therefore, the use of major food crops such as maize cannot support the ambitious objectives of renewable fuel legislation. Environmental and economic factors have also prompted industries to embrace sweet sorghum as an energy crop.

1.2.1 Environmental benefits of sweet sorghum

Environmental factors have also contributed to the adoption of new crops devoted wholly for liquid automotive fuel to minimize the use of major farming land, irrigation water and other inputs. A devoted energy crop, preferably, must satisfy numerous requirements such as: high biomass yield and growth rate, perennial, reduced input necessities, well adapted to the local environment, easy to perfect through genetic improvement, tolerant to multiple stresses and with a good carbon sequestration rate among others (Jessup, 2009). Carbon sequestration by sweet sorghum provides a potential solution for miyigating emission of greenhouse gases (GHGs) in a standard scenario. These GHGs include methane (CH_4), carbon dioxide (CO_2) and nitrous oxide (N_2O) (Jessup, 2009).

In tropical, subtropical and arid regions of the United States, Mexico, China, India, southern Africa and other developing countries where unfavourable agronomic environments prevail, one of the most ideal crops for biofuel is sweet sorghum (*Sorghum bicolor* (L.) Moench) (Reddy et al., 2005; Zhang et al., 2010). Sorghum is a high efficient photosynthetic crop that reached a worldwide production of 56 million tonnes of grain in 2009 (FAO, 2011), just after maize, wheat, rice and barley. It is a C4 plant which is very resilient to biotic and abiotic factors as insects, drought, salinity and soil alkalinity. Further to that, this crop has one of the best rates of carbon assimilation (50 $g/m^2/day$) which in turn facilitates for a fast growth rate and a better rate of net carbon dioxide use (Prasad et al., 2007). Almodares and Hadi (2009) also reported that sorghum needs one third of the water compared to sugar cane and 80 to 90% compared to maize. Therefore, sorghum is a fairly drought tolerant crop.

Furthermore, sorghum needs almost one third of the fertilizer required by sugar cane (Kim and Day, 2011), and its growth cycle is between 3 to 5 months which allows two or three crop cycles per year instead of one normally obtained with sugar cane. Besides environmental advantages, sorghum is one of the more acquiescent crops to genetic improvement due to its genetic variability (Zhang et al., 2010). This eases plant breeding and cultivar development for adaptability across locations.

1.2.2   Economic benefits of sweet sorghum

Just like in sugar cane, ethanol production using sweet sorghum requires less energy than ethanol production based on maize, a starch rich C_4 crop (Woods, 2001). Woods (2001) also reported that maize requires the hydrolysis of starch to more easily fermentable sugars, resulting in a lower energy ratio and raised ethanol production costs. Lignin derived from sorghum biomass can be used to produce biodegradable plastics (Ashori, 2008). Lignin is a by-product of bioethanol production. Furthermore, sorghum cellulose fibre can be used to reinforce thermoplastic materials such as decomposable wood composites (Ashori, 2008).

After extraction of juice sugars, sweet sorghum bagasse can be used for manufacturing paper pulp. Belayachi and Delmas (1995) reported that the quality of the pulp obtained is like regular softwood used for paper making. The authors also reported that sorghum pulp displays a degree of adhesion higher than 80%; a low kappa number, which shows a good delignification; a high degree of polymerisation; and excellent physico-mechanical properties. This justifies the use of sweet sorghum as a major raw material for the paper industry in every region where it will be suitable to cultivate (Belayachi and Delmas, 1995). Economic policies and social pressures are used to evaluate the land-use competition for food or feedstock production for biofuel.

Any crop thrives better on fertile soil than on an impoverished soil. However, bioenergy crops like sweet sorghum can thrive on poor soils under improved crop husbandry practices. Therefore, while farmers cultivate food crops on arable land, they can use their marginal soils for biofuel production (Basavaraj et al., 2013). Sweet sorghum also requires low nitrogen fertilizer compared to maize and sugar cane. This low production cost for sweet sorghum crop compensates for the yield gap.

1.3   World production of sorghum

By 2014, America was the world leading sorghum producer, followed by Europe (FAO, 2017), as summarized in Table 1.1.

Table 1.1   Global sorghum production statistics by regions by 2014

Region	Area cultivated (ha)	Yield (tons)	Yield per area (tons/ha)	Production % by region
America	7,224,581	27,402,003	3.8	16.1
Europe	390,410	1,376,253	3.5	0.9
Oceania	533,259	1,286,853	2.4	1.2
Asia	7,455,352	9,680,531	1.3	16.6
Africa	29,355,124	25,117,422	0.9	65.3

Sourced from (FAO, 2017).

In terms of biofuel, USA and Brazil are the world's major producers (Table 1.2). In 2015, the annual ethanol production in the USA was 14,806 million gallons (55.6 billion litres), accounting for 57.6% of the total global production. Brazil stood at 7,093 million gallons (26.8 million litres), accounting and 27.6% of the total world production (FAO, 2017).

Table 1.2   World fuel ethanol production by country or region (Million Gallons)

Country	2007	2008	2009	2010	2011	2012	2013	2014	2015
USA	6521	9309	10938	13298	13948	13300	13300	14300	14806
Brazil	5019.2	6472.2	6578	6921.5	5573.2	5577	6267	6190	7093
Europe	570.3	733.6	1040	1208.6	1167.6	1179	1371	1445	1387
China	486	501.9	542	541.55	554.76	555	696	635	813
Canada	211.3	237.7	291	356.63	462.3	449	523	510	436
Asia (minus China)	132	155.8	527	244.36	334.94	397	na		na
South America (minus Brazil)	74.9	79.2	83	200.22	198.66	223	na		na
Mexico & Central America	na	na	na	364.36	39.04	19	na		na
Australia	26.4	26.4	57	66.04	87.2	71	na		na
Africa	na	na	na	43.59	38.31	42	na		na
Other	82	128	247	66.04	na	na	na		na
Rest of World	315.3	389.4	914	984.61	698.15	752	727	1490	1147
World	13123	17644	20303	23311	22404	21812	22884	24570	25682

na = Not applicable

1.4 Problem statement and justification

Numerous sweet sorghum cultivars are distributed globally, hence providing a wide genetic base from which to develop specifically adapted, highly productive cultivars (Audilakshmi et al., 2010). Despite all the agronomic benefits of sweet sorghum as a bioenergy crop, the crop has received little scientific effort in the past toward the identification of sweet sorghum traits associated with bioenergy production. Traits like plant height, stem diameter, green biomass, stem sugar content, and stem juice extractability are the major contributors for the economic importance of sweet sorghum (Almodares et al., 2008).

These traits are quantitative and polygenically inherited in nature and are complex to be worked on directly in breeding procedures because of environmental noise. Therefore, to effectively work on these complex traits, there is a need to separate them into minor morphological, physiological and yield components, which could be simply analysed and assessed. Previous studies in the U.S. have suggested that much variability exists in juice quality, sugar content, and juice yield among sweet sorghum collections (Ali et al., 2008). However, information on the level of variation in growth (plant height and stem diameter), physiology and components of stem sugar (%brix, juice yield and stem fresh weight) among South African sweet sorghum germplasm is limited.

1.5 Objectives

The overall objective of this research was to characterise South African sweet sorghum germplasm for biofuel production.

1.5.1 Specific objectives

- To estimate the level of variability among 50 sweet sorghum germplasm using agro-morphological traits.
- To estimate genotypic and non-genotypic variance components and to determine correlations and path analysis for juice yield and juice related traits.
- To assess the level of genetic diversity among the sweet sorghum germplasm using single nucleotide polymorphism (SNP) markers.

1.6 Hypotheses

- There is no significant variation in terms of agro-morphological traits among sweet sorghum lines grown under two different environments.

- There are no significant differences in juice yield and brix content within the South African sweet sorghum gene pool.

- There is no genetic diversity among sweet sorghum lines grown in South Africa.

Dissertation outline

The dissertation is arranged in separate research chapters (Table 1.3), each following the format of an independent research paper. This is the format which was embraced and is used by the University of KwaZulu-Natal (UKZN). Therefore, there is unavoidable repetition of some references between chapters. The referencing style used in this dissertation is based on the Crop Science Society of America (CSSA) and follows the specific style in the Crop Science Journal.

Table 1.3   Overview of the dissertation structure

Chapter	Title
1	Introduction
2	Literature review
3	Characterisation of sweet sorghum germplasm using morpho-agronomical traits under two environments
4	Genetic and path coefficient analysis of juice yield and juice related traits in sweet sorghum
5	Assessment of genetic diversity in sweet sorghum accessions using Single nucleotide polymorphism (SNP)
6	General overview

References

Ali, M.L., J.F. Rajewski, P.S. Baenziger, K.S. Gill, K.M. Eslridge and I. Dweikat. 2008. Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Molecular Breeding 21: 497-509.

Almodares, A. and M.R. Hadi. 2009. Production of bioethanol from sweet sorghum: A review. African Journal of Agricultural Research 5: 772-780.

Almodares, A., R. Taheri and S. Adeli. 2008 Stalk yield and carbohydrate composition of sweet sorghum (*Sorghum bicolor* L. Moench) cultivars and lines at different growth stages. Journal of Malaysian Applied Biology 37: 31-36.

Ashori, A. 2008. Wood-plastic composites as promising green-composites for automotive industries. Bioresource Technology 99: 4661-4667.

Audilakshmi, S., A.K. Mall, M. Swarnalatha and N. Seetharama. 2010. Inheritance of sugar concentration in stalk (Brix), sucrose content, stalk and juice yield in sorghum. Biomass Bioenergy 34: 813-820.

Basavaraj, G., P.P. Rao, L. Achoth and C.R. Reddy. 2013. Assessing competitiveness of sweet sorghum for ethanol production: A policy analysis matrix approach. Agricultural Economics Research. Review 26: 31-40.

Belayachi, L. and M. Delmas. 1995. Sweet Sorghum: A quality raw material for the manufacturing of chemical paper pulp. Biomass Bioenergy 8: 411-417.

Doggett, H. 1988. Sorghum. Longman Scientific and Technical, London.

FAO. 1995. Sorghum and millet in human nutrition. http://www.fao.org/docrep/T0818e/T0818E00.htm#Contents. 24 August 2017.

FAO. 2017 Database of agricultural production. 2011. FAO Statistical Databases. A compilation of bioenergy sustainability initiatives. Bioenergy and Food Security. http://www.fao.org/faostat/en/#data/QC. 7 August 2017.

Jessup, R. 2009. Development and status of dedicated energy crops in the United States. In Vitro Cellular & Developmental Biology 45: 282-290.

Kim, M. and D. Day. 2011. Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. Journal of Industrial Microbiology & Biotechnology 38: 803-807.

Murray, S.C., W.L. Rooney, M.T. Hamblin, S.E. Mitchell and S. Kresovich. 2009. Sweet sorghum genetic diversity and association mapping for brix and height. The Plant Genome 2: 48-62.

Prasad, S., A. Singh, N. Jain and H.C. Joshi. 2007. Ethanol production from sweet sorghum syrup for utilization as automotive fuel in India, Energy & Fuels. Plant Science 21: 2415-2420.

World Food Program. 2008. Rising food prices: impact on the hungry. Database of Press Releases related to Africa.

Reddy, B.V., S. Ramesh, P.S. Reddy, B. Ramaiah, M. Salimath and R. Kachapur. 2005. Sweet sorghum-a potential alternate raw material for bio-ethanol and bio-energy. International Sorghum and Millets Newsletter 46: 79-86.

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

Serna-Saldívar, S.O., C. Chuck-Hernández, E. Pérez-Carrillo and E. Heredia-Olea. 2012. Sorghum as a multifunctional crop for the production of fuel ethanol: Current status and future trends. Bioethanol. InTech Open Science.

Srinivasa, P.R., C.G. C. Ganesh Kumar and B.V.S. Reddy. 2013. Sweet Sorghum: From Theory to Practice, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Vaidyanathan, S., K.E.P. Rao, M.H. Mengesha and R. Jambunathan. 1987. Total Sugar content in sorghum stalks and grains of selected cultivars from the world germplasm collection. Journal of the Science of Food and Agriculture 39: 289-295.

Woods, J. 2001. The potential for energy production using sweet sorghum in southern Africa. Energy for Sustainable Development 5: 31-38.

Zhang, C., G. Xie, S. Li, L. Ge and T. He. 2010. The productive potentials of sweet sorghum ethanol in China. Applied Energy 87: 2360-2368.

CHAPTER 2

Literature Review

2.1 Introduction

This chapter reviews literature on topics related to sweet sorghum characterisation and evaluation for biofuel production. The review focuses on taxonomy, growth stages, importance (or uses) of sweet sorghum, production constraints, genetic diversity, inheritance of traits, genotype by environment (G x E), interactions mechanisms accounting for sugar accumulation in the stalks, correlation of both agronomical and industrial traits and path analysis. The chapter then reviews morphological and molecular characterisation of sweet sorghum.

2.2 Taxonomy

Sweet sorghum [*Sorghum bicolor* (L.) Moench] is an annual monocot angiosperm in the grass species, belonging to the family Poaceae and sub family Panicoideae, (Serna-Saldivar, 2010). It belongs to the same species as grain sorghum and forage sorghum. Sorghum is mainly a semi-arid tropical plant. It is predominantly adapted to water stress due to several morphological and physiological features, which include wide-spread root system, waxy bloom on leaves which minimises water loss, ability to stop growth during moisture stress periods and C_4 photosynthetic pathway (Balole and Legwaila, 2005). It is termed as "the miracle crop," "the sugar cane of the desert" or "the camel among crops" (Sanderson, et al., 1992), due to its drought hardy characteristics. Therefore, it can survive the harsh climatic conditions of the arid environments (Ritter et al., 2007)**.**

The juice extracted from sweet sorghum stem is rich in sucrose and invert sugar which are fermented to produce ethanol (Prasad et al., 2007). Sweet sorghum has the potential to produce 43 Mg per hectare of juice annually under favourable conditions with an average of 11.8% of fermentable sugars (Kim and Day, 2011). Some sweet sorghum cultivars have been reported to produce sugar yields similar to sugar cane (Ratnavathi et al., 2010).

2.3 Growth stages

2.3.1 Germination

Germination in sorghum, like any other crop, is temperature dependent. House (1980), reported that germination occurs quickly in warm soils. When soil temperatures are 20˚C or above, the coleoptile first appears above the ground after 3 or 4 days, while with temperatures between 20˚C and 13 ˚C, it takes up to 10 days for the coleoptile to appear above the ground.

2.3.2 Vegetative stage

The tri-foliate stage befalls when the collars of three leaves are visible without dissecting the plant. Depending on the temperature, this stage may occur almost 10 days after emergence. The five-leaf stage occurs when the collars of five leaves are seen without dissecting the plant, and this occurs about three weeks after emergence. The root system develops quickly at this stage. Under optimum growing conditions, dry matter accumulates at nearly a constant rate. The development potential of plants is determined at this stage. Any biotic or abiotic stress at this stage will have significant impact on yield (House, 1980).

2.3.3 Heading and flowering

Sweet sorghum is mainly a self-pollinated species. However cross-pollination of 4–10% has been reported under specific conditions (Srinivasa et al., 2013). Sweet sorghum is a short-day plant, and blooming is accelerated by short days and long nights. However, varieties differ in their photoperiod sensitivity (Quinby and Karper, 1947). Tropical sweet sorghum varieties initiate the reproductive stage when day lengths return to 12 hours. Anthesis in sweet sorghum starts 30 to 40 days after germination. House (1985), reported that sorghum usually flowers in 55 to 70 days in warm climates, depending on the genotype. The flowers of sorghum open during the night or early morning with those at the top of the panicle opening first, causing variations in maturity. Anthesis continues in a sequential downward manner, with florets in a horizontal plane opening during the same period (Balole and Legwaila, 2005). About 6-10 days prior to floral initiation, the boot protrudes in the flag leaf sheath.

Two days after the inflorescence emerges from the boot, the flowers start to open. The flowering starts in the sessile spikelets at the tip of the inflorescence and advances toward the bottom for about four to five days. It takes 6-9 days for the whole panicle to complete flowering with maximum flowering occurring on the third or fourth day (House, 1985).

Flowering marks the end of the vegetative growth due to meristematic activity. Genetic makeup and the environment mainly influence the time required for transformation from the vegetative to reproductive phase (House, 1980). Mathur et al. (2017) reported that delayed anthesis is necessary for more biomass accumulation. Several studies have been done on QTL controlling flowering days. Murphy et al. (2014) identified loci *Ma1-Ma6* as the regions influencing photoperiod sensitivity maturity in sorghum. They also reported that *Ma1* codes for a flowering suppressor *SbPRR37,* which is circadian clock-regulated and suppresses flowering during long days. In another study, Calvino and Messing (2013) identified a collection of *miRNA169* on chromosome 1 and another on chromosome 7. These clusters presented significant synteny with the chromosomal sections having linked *bHLH* and constans-like genes from both monocots and dicots. This suggested a strong conservation among flowering and plant height-related genes and *miRNAs* which account for the linkage drag evident in drought and flowering traits.

2.4 Importance of sweet sorghum

Sweet sorghum as a multipurpose crop has the potential as a substitute biofuel raw material without much threat on food and fodder security (Srinivasarao et al., 2015). The grain can be used to produce food. Sorghum grain comprises 60 to 70% of starch, like maize. In general, the composition of sorghum is comparable to maize, with a few minor significant differences particularly in protein and fat compositions. Sorghum grain has an average 1% less fat and 1.5 to 2.0% more crude protein compared to maize (Serna-Saldivar, 2010).

Almodares and Hadi (2009) reported that sweet sorghum juice can be used for syrup, molasses, sugar and ethanol production with average fermentation efficiencies ranging from 85 to 90%. An earlier report by Huligol and Govind (2004) suggested that high efficiency in sweet sorghum is attributed to higher total reducing sugar content compared to sugar cane juice. In their study on bagasse which is the fibrous residue after juice extraction, Seetharama et al. (2002) suggested that bagasse can be used to generate power. Further, the authors reported that bagasse from sweet sorghum had a higher biological value than the bagasse from sugar cane when used as forage for animals, because it is rich in minor nutrients and minerals.

2.5   Production Constraints

2.5.1   Cultivars

Ideal cultivars for ethanol production are those which produce medium to large-sized, strong, erect stalks with high total sugars. Furthermore, cultivars must be well adapted across environments and resilient to prevailing biotic and abiotic stresses (Pfeiffer et al., 2013). However, most of the sweet sorghum cultivars, according to Srinivasa et al. (2009) are affected by genotype by environment (G × E) interactions for juice related traits. The genotypes that perform well in the rainy season do not certainly replicate their performance during post rainy season and vice versa.

2.5.2   Pests and diseases

Like grain sorghum, important sweet sorghum diseases are leaf anthracnose, red stalk rot, and maize dwarf mosaic virus and charcoal rot (Das et al., 2008). These diseases are mostly controlled through crop rotation and use of improved resistant varieties. However, most of the genotypes cultivated by farmers are not high yielding and are not well adapted to multiple stresses (Srinivasa et al., 2009).

2.5.3   Harvesting and storage

Commercial production of sweet sorghum for biofuel requires mechanical harvesting. Because the sugars start degrading after harvest, juice extraction must be done within a few days if the stalks are harvested whole. This short shelf-life of the raw juice makes storage difficult. The juice must be fermented quickly after extraction. This means that processing plants need to be located on-farm or a few kilometres away (Reddy and Reddy 2003).

2.5.4   Economic challenges

Sweet sorghum production needs a considerable amount of start-up capital which includes land preparation and purchase of seed, fertilizer, and other inputs necessary for crop production. Profitability is also dependent on transportation costs to the market. This implies that longer distances minimise the possibility of positive economic earnings (Pfeiffer et al., 2013). Therefore, use of superior cultivars could help offset higher harvest and transportation expenses.

2.6   Current status on sweet sorghum research and production in South Africa

In a report by Johnson and Matsika (2006), the authors suggested that sub-Saharan Africa has the highest ability for bioethanol production because it has vast areas of underutilized land, sub-tropical climate and low productivity levels. However, low annual production level of 5000 litres per hectare of ethanol are realized (Reddy et al., 2007).

Sweet sorghum was identified as a potential biofuel crop for bioethanol production due to the high fermentable sugars in its stalks and its ability to thrive in marginal environments in South Africa (Tsuchihashi and Goto, 2008). In an earlier report by Prasad et al. (2007), the authors reported that adoption of improved sweet sorghum cultivars in South Africa as feedstock could increase the size of the current sorghum market by twofold, if yields of improved varieties are superior over the present average. However, Musango and Brent (2011) lamented that little effort has been directed towards sweet sorghum improvement programme for commercial production in South Africa. Makanda et al. (2009) also reported that South African sweet sorghum accessions have never been characterised regarding their suitability for biofuel production. The authors further reported that there were not many breeding programmes that underline stem sugar content, especially in sorghum cultivars. Chinyama (2016) also reported that majority of the germplasm found in South Africa are landraces, particularly grown by small-scale farmers for domestic consumption.

2.7   The genetic diversity amongst sweet sorghum germplasm

Genetic diversity can be defined as any difference in nucleotides, genes, chromosomes or whole genomes of individuals (Wang et al., 2009), and this is usually expressed phenotypically. Genetic diversity can be estimated among different accessions or organisms within same species (intraspecific), among species (interspecific) and between genus and families (Mittal and Dubey, 2009). Accession origins are major sources of genetic diversity, and accessions of the same source or breeding background have the lowest genetic distance (Lekgari and Dweikat, 2014).

Determination of genetic diversity among sweet sorghum accessions is a gateway to the study of evolutionary forces that determine their existence. This also helps to identify collections of similar genotypes for the purposes of conservation and proper use of genetic resources, in addition to the protection of property rights. Phenotypic estimates are used to show the extent of genetic connection and consequently the resemblance in phenotypic traits may show genetic similarity of genotypes (Cox et al., 1985). Genetic diversity study of a crop species involves cultivation of sub-samples and the use of phenotypic markers.

14

Phenotypic and genetic diversity are significant in genetic conservation, assessment and utilization of genetic resources, and the study of breeding germplasm for determining uniqueness and genetic composition (Sergio and Gianni, 2005). Multivariate analysis of variance (MANOVA) has been reported to be suitable for the assessment of population divergence while studying quantitative and qualitative traits (Sneath and Sokal, 1973). Genetic distance is used in approximating the diversity in plant populations. Euclidian distance measure is the most frequently used statistics for assessing genetic distance between two groups or organisms using statistical measures based on phenotypic data.

The phenotypic variations are characteristically due to genetic complexity from multiple interacting loci; with allelic effects which are sensitive to the environmental conditions each individual experiences (Lynch and Walsh, 1998). Significant traits in sweet sorghum such as height, length of internode, number of tillers, panicle length, peduncle exertion and yield of grain per area are quantitative in nature. For qualitative traits, a large extent of genetic diversity exists among sweet and grain sorghum populations regarding glume colour, panicle compactness, panicle shape, grain colour and grain cover (Kisua et al., 2015).

Many studies have been done on genetic diversity of sweet sorghum. Doggett (1970) reported that the highest genetic and phenotypic diversity in both wild and cultivated accessions of sorghum are found in central Africa. Kisua et al. (2015) reported that sweet sorghum accessions are genetically closely related to each other but highly diverged from grain sorghum. In a previous study, Lekgari and Dweikat (2014) studied 142 sweet sorghum germplasm of diverse origin with molecular and morphological markers and produced five distinct cluster groups. The authors however argued that although the clusters were not identical they complemented each other. In another previous study, Murray et al. (2009) genotyped a diverse panel of 125 sweet sorghum using 47 simple sequence repeats (SSRs) and 322 single nucleotide polymorphisms (SNPs). Three major groups of sweet sorghum were identified, each with several subgroupings. This study suggests existence of wide genetic diversity for sweet sorghum from which suitable parents can be selected for breeding programme.

2.8 Inheritance of traits (qualitative and quantitative traits)

To facilitate the breeding for cultivars with desirable traits, it is necessary to gather information on the genetics of such traits on how they are inherited including the gene action controlling such traits. Understanding the relative effects of gene action would be necessary in deciding the breeding strategies and suitable selection criteria to develop sweet sorghum cultivars with improved sucrose levels and juice yields. Audilakshmi et al. (2010) reported that heterosis for

plant height, green stalk yield, and commercial stalk sugar is significant. The authors further reported that inheritance of brix and stalk weight in sugar stalk is under both additive and non-additive gene action. This suggests that both general combining ability (GCA) and specific combining ability (SCA) were significant. However, the inheritance of the traits was mainly controlled by non-additive gene action. Similar findings were reported by Yang et al. (2009) and Shiringani et al. (2010) who independently indicated that sugar composition is a quantitative trait controlled by multiple gene effects with environmental interaction. Felderhoff et al. (2012) also suggested that brix was an additive trait. Therefore, breeding for high sugar concentration in sweet sorghum hybrids entails both parents to have high brix values. In order to identify the genomic regions linked to sugar content in sweet sorghum, Yun-long et al. (2006) crossed a high sugar content inbred line, early Foger with another inbred line, N32B. Analysis of 207 segregating populations identified two QTL, which explained total phenotypic variation for %brix ranging from 22.2 to 25%. In a later report Murray et al. (2008) studied a population generated from sweet sorghum cultivar Rio and grain sorghum cultivar BTx623. The QTL, which controlled yield and stem sugar content and QTL that controlled grain yield did not have pleiotropic effects on each other. This led to identification of some QTL for sugar components on SBI-01, SBI-02, SBI-03, SBI-05, SBI-06, SBI-07, SBI-10.

2.9   Genotype by environment (G x E) interaction

Mather and Caligari (1976) defined genotype by environment interaction (G x E) as the differential response of genotypes under changes in the environmental conditions. It also refers to situations where the combined effects of genotype and environment are significantly greater or significantly lower, than the predicted sum of the separate effects.

The genotype and environment interaction influences the selection procedures of genotypes, including sweet sorghum. Studies on adaptability and stability provide information about the performance of each genotype under different environmental conditions (Rono et al., 2016). The phenotypic performance of each genotype is influenced by abiotic and biotic factors; some genotypes may perform well in one environment but fail in another environment (Fentie et al., 2013). These factors include rainfall, temperature, soil fertility, light, pests, diseases and crop husbandry practices that differ across sites and significantly influence yield potential of crop varieties. Environmental factors can also mean different planting dates. These factors make it hard to establish the superiority of cultivar across varying environments (Aslam et al., 1993).

The G x E interaction must be assessed for yields, which are cane, biomass, juice yield and sugar content (Rono et al., 2016). The interaction affects crop performance, complicates breeding, testing and selection for superior genotypes. The G x E interaction changes the

rankings of genotypes in various environments; thus, an increase in G x E interaction weakens the association between genotypic and phenotypic qualities making it difficult to discriminate superior genotype across locations (Yan and Kang, 2002).

Performance of cultivars for quantitative traits such as yield and other characters, which influence yield, sometimes differs with environments. Consequently, to develop a cultivar with high yielding potential and which is stable across sites, attention should be given to the importance of stability performance for the genotypes under varying environments and their interactions (Ghazy et al., 2012).

According to a report by Audilakshmi et al. (2010), G x E interaction influence is manifested for plant height, stalk weight, reducing sugars and percent of total sugar. Rono et al. (2016) reported significant impact of the interaction on days to anthesis and juice yield. The authors also reported a positive correlation between temperature and brix values. On plant height, however, previous studies have reported lower environment effect on the height variation among sorghum genotypes. This suggests that the trait is controlled by a few genes, with a minor environmental influence. In a genetic study, Quinby and Karper (1954), reported four loci that influence sorghum height. These are *Dw1, Dw2, Dw3* and *Dw4*. In another genetic linkage mapping studies, Salas et al. (2009) showed that sorghum height is influenced by only a few major QTL that have high heritability. These earlier reports however contradict with a recent report by Abubakar and Bubuche (2013) on biomass, where it was reported that genotype by environment interaction was highly significant on plant height. This suggests that selection for this trait cannot be done across locations.

Therefore, several studies have focused on height as secondary trait to identify the genes correlated to biomass yield (Salas et al., 2009). In a study to determine genotype by sowing date interactions for sugar yield components and to identify high yielding and stable or specifically adapted genotypes for cultivation across the year, Reddy et al. (2014) used genotype plus genotype by environment (GGE) biplot analysis to evaluate five sweet sorghum genotypes for three seasons. The authors reported that date of sowing was the main source of variation accounting to 58% to 94% of the genotype plus sowing date plus genotype by sowing date (G+S+GS) for sugar yield and related traits. This suggests that sowing dates must be considered in an attempt to realise higher yields.

2.10 Sugar accumulation in the stalk

Sweet sorghum juice contains 12–22% (w/v) sugar. Sugar composition in the stem juice is mostly estimated using a refractometer. A refractometer value of 1% brix indicates 1 gram of soluble solids in 100 ml juice (Laopaiboon et al., 2007)

Sweet sorghum accumulates sucrose in the stem. This feature is rare among plants, making sweet sorghum a suitable source of bioethanol (Calvino et al., 2008). Contradicting findings, however, have been reported on sucrose accumulation in sweet sorghum (Burks et al., 2015). Sweet sorghum, like sugar cane, accumulates sucrose in stem parenchyma cells, but phloem filling and the timing of sucrose accumulation both differs between the two species. In sugar cane, sucrose transfer in mature internodes is largely symplasmic, with sugars loaded into the phloem via plasmodesmata (Tarpley and Vietor, 2007). In sweet sorghum, sucrose transfer is symplasmic in growing internodes but apoplastic in mature internodes. Lingle (1987) reported that sweet sorghum sugar accumulation increases after panicle emergence, once internode elongation has stopped. Other studies suggest that the timing of sucrose accumulation is similar in sweet sorghum and sugar cane, beginning while stem elongation is still occurring (Hoffmann-Thoma et al., 1996; Gutjahr et al., 2013). However, studies on enzymatic control and carbon transport indicate that the mechanism of accumulation for sugar cane and sweet sorghum is different, as reported by Lingle (1987) and Tarpley and Vietor (2007).

Gutjahr et al. (2013) suggested that these contradictory findings might arise from the use of photoperiod-sensitive (PS) versus photoperiod-insensitive (PI) sorghum cultivars. The authors reported that in PI genotypes or PS genotypes grown under short days, internode maturation is generally coincident with flowering, whereas in PS genotypes grown under long days, flowering is delayed relative to internode maturation, and sugar levels at anthesis are substantially higher.

Sucrose content normally depends on its metabolism, transport, and storage. The significant genes in sucrose metabolism are *sucrose phosphate synthase* (SPSS) and *sucrose synthase* (*SUS*); their products catalyze rate-limiting steps in this metabolic pathway (Ludewig and Flugge, 2013). Invertase (INV) degrades sucrose to glucose and fructose, thus determining whether sugar molecules are stored as sucrose or starch. These factors synergistically influence stem sucrose content (Qazi et al., 2012).

To determine the optimum harvest stage for realizing high sugar yield, it is necessary to study the dynamic flux of component stalk sugars like glucose, fructose and sucrose in the juice at post-flowering and with respect to various phenological stages. This is important for sweet

sorghum commercialization and value chain sustenance. Kumar et al. (2010), reported that sucrose (a component which accounts for about 70% of major fermentable sugar) was high at physiological maturity, but highest at post-physiological maturity, with no trade-off in terms of brix (%) and juice yield. However, the authors reported that there was significant genotype x stage interaction for juice yield, %brix, sucrose, fructose levels, pH and glucose content. This implies that harvest date is critical for quality yield and yield components. Similar findings were also reported by Zegada-Lizarazu and Monti (2012), that sucrose concentrations in sweet sorghum increase greatly after flowering.

2.11 Correlation of traits and path coefficient analysis

Information on correlation (direction and magnitude) of traits is very helpful in breeding programme for further understanding of the genetic mechanism of stalk yield-related traits and sugar concentration of stalk juice (Kisua et al., 2015). This is critical in determining whether the traits are genetically controlled and heritable for transmission to the desired genotypes. In sweet sorghum, therefore, it is important to understand correlation of both agronomical traits (such as plant height, stalk diameter, stalk yield, panicle mass) and industrial traits such as juice extraction (%), reducing sugar content, stalk fibre ratio, sucrose percentage in a stalk, total soluble solute percentage (brix), just to mention a few (Mahajan et al., 2011).

Many traits such as green stalk yield, stalk sugar content, stalk juice extractability and grain yield have been shown to be major contributors to its economic superiority (Bala et al., 1996). However, these traits are quantitatively and polygenically inherited in nature and very difficult to manipulate directly in breeding procedures. Therefore, to effectively improve these complex traits, they need to be separated into smaller morphological, physiological and genetic components, which are easily analysed and evaluated. Correlations between the traits are important for successful selections (both direct and indirect) in breeding programmes. Significantly positive correlations imply that the changes of two variables are in the same direction, while negative correlations indicate their inverse associations with each other. Positive correlation of traits implies that simultaneous selection and improvement of traits can be done.

Several researchers have reported similar findings on association of traits in sweet sorghum. In a genetic diversity study, Kisua et al. (2015) evaluated sweet and grain sorghum accessions and reported a strong positive association between number of leaves and plant height. An earlier study by Audilakshmi et al. (2010) also reported significant positive correlations between plant height and stalk yield and between stem diameter and juiciness. In another correlation study, Shukla et al. (2017) reported that sweet sorghums do not have to be tall to

accumulate sugar. However, the authors identified a positive correlation between height and sugar accumulation.

Makanda et al. (2009) reported a negative and highly significant correlation between grain yield and stem biomass. This suggests existence of antagonistic genetic mechanisms between these two traits. In a report by Lombardi et al. (2015), days to anthesis and plant height presented positive phenotypic correlations with total stalk yield per hectare. This indicates that taller and late flowering plants are associated with higher stalk yield. An earlier study by Reddy et al. (2005) reported a negative association between panicle mass and %brix.

Although correlation information is key in plant breeding, individual analysis is limited to a pair of traits. Furthermore, simple correlation estimates may not signify the actual association between two traits, since there may be interference by a third trait or group of traits that might distort the correlation estimates. Therefore, a path analysis is employed which reveals direct and indirect variables which are significantly correlated to the dependent variable (Wright, 1921). This allows separating of correlation coefficient into direct and indirect effects of several traits towards a dependent variable and thus helps in evaluating the cause-effect association as well as effective selection (Dewey and Lu, 1959). Further, Dewey and Lu (1959) also reported that path analysis was critical for better understanding of correlations among traits, which is a pathway for knowledge on specificity of the genetic material being studied.

2.12 Sweet sorghum characterisation

Characterisation of germplasm is essential for the sustainable maintenance and continuous use of crop genetic resources (Sergio and Gianni, 2005). This involves, mainly, identifying heritable traits, leading to classification that facilitates enhanced utilization of germplasm and successful breeding programme (Upadhyaya et al., 2008). Morphological and molecular characterisation is recommended in sorghum accessions collected from different regions to guarantee a further comprehensive and informative characterisation. Results of molecular studies are regarded as complementary to agronomic and morphological characterisation (Karp et al., 1997).

2.12.1 Morphological characterisation

According to Franco et al. (2001), morphological markers are phenotypic characters such as flower colour, seed colour and shape, growth habits, pigmentation, texture, maturity, yield, and pest and disease resistance. Plant selections are done based on morphological characteristics that can be discriminated and that are co-inherited with the favourable trait. Morphological

traits displayed by sweet sorghum landraces determine selection and preference by the farmers. Majority of the farmers keep sorghum landraces that are early maturing, high yielding, tolerant to drought, pests and diseases, and with grain colour producing appealing products (Muii et al., 2013). Morphological markers have, in the past, been extremely useful to plant breeders. Sergio and Gianni (2005) argued that although these methods remain effective, morphological or pedigree evaluations are influenced by the environment and management practices. They are also limited in many ways and are influenced by the developmental stage of the plant. Semagn et al. (2006) suggested that because morphological markers are highly influenced by the environmental conditions, it is necessary to use of molecular marker data to supplement or complement their clustering.

2.12.2 Molecular characterisation

Semagn et al. (2006) defined molecular markers as polymorphisms which are found naturally in populations that show neutral sites of variation at DNA sequence level. They are biological features that are determined by allelic forms of genes or genetic loci and can be passed on to successive generations. Molecular markers are also referred to as DNA markers. DNA markers are defined as a fragment of DNA showing mutations/variations, which can be used to detect polymorphism between different genotypes or alleles of a gene for a sequence of DNA in a population or gene pool (Southern, 1975).

The use of molecular markers allows plant breeders and geneticists to locate and understand the basics of several gene interactions which determine complex traits (Haussmann et al., 2000). Jiangfeng et al. (2014) also stated that molecular markers can be used as experimental probes or tags to trace an organism, a tissue, a cell, a nucleus, a chromosome or a gene. Several molecular markers such as random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphism (AFLP), restriction fragment length polymorphisms (RFLPs) and microsatellites have previously been applied on genetic diversity studies for crops.

Mehmood et al. (2008) assessed the genetic diversity and phylogenetic relationship among diverse sorghum varieties by Genome DNA fingerprinting using RAPDs, and observed a significant level of genetic diversity among the sorghum varieties. Using AFLPs and SSRs, Menz et al. (2004) examined the diversity in elite sorghum R- and B-lines compared to exotic and converted germplasm, and the classification of germplasm by the two different markers. Cluster analysis failed to reveal clear differentiation between the R- and B-lines, whilst AFLP markers produced clusters in better agreement to pedigree information due to their distribution and coverage of the genome. Kimberley et al. (2007) examined the genetic relatedness between sweet sorghum and grain sorghum, using AFLPs. A cluster analysis discriminated

older sweet sorghum accessions (collected in mid 1800s) from those developed and released during the early to mid-1900s. Sweet sorghum lines were largely distinguished from grain sorghum.

2.12.3   Microsatellites

Microsatellites are the third classes of molecular markers. These are DNA sequences of mono-, di-, tri-, tetra- and pentanucleotide units repeated in tandem, which are widely dispersed in the genome (Tautz and Renz, 1984; Powell et al., 1996). Microsatellites are extensively distributed throughout the genome, they are co-dominant, highly polymorphic and transferable between species. These attributes provide the foundation for their successful application in a wide range of fundamental and applicable fields (Chistiakov et al., 2006). They also have a significant degree of allelic variation, making them valuable as genetic markers (Xu, 2010). The major mutation mechanism in microsatellite tracts is slipped-strand mispairing (Levinson and Gutman, 1987). When slipped-strand mispairing arises within a microsatellite array during DNA synthesis, it can result in the gain or loss of one or more repeat units depending on whether the newly synthesized DNA chain or the template chain loops out. Microsatellites include simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) (Gupta et al., 2001).

2.12.3.1   Simple sequence repeats (SSRs)

The SSRs are more abundant in non-coding regions than in coding regions. They are regarded as the most efficient markers although their usage is limited due to long laborious steps involved to develop them (Rakoczy-trojanowask and Bolibok, 2004). Using SSRs, Folkertsma et al. (2005) studied the genetic diversity of guinea-race sorghum accessions and found most of the variation was among accessions from semi-arid and Sahelian Africa and the least among accessions from South Asia, thus supporting earlier reports on the spread of the guinea-race sorghums across Africa and south Asia. In another study, Ngugi and Onyango (2012) genotyped sorghum germplasm grown in Kenya using SSR markers to estimate the level of genetic diversity present and reported a wide range of variation in quality index from 0.005 to 0.39. The average polymorphic information content (PIC) value observed was 0.6241 signifying a high level of diversity. The gene diversity index ranged between 0.2419 and 0.9313 with a mean of 0.6627 per locus. Using ten SSR markers, Chinyama (2016) studied twenty-four sweet sorghum germplasm obtained from the International Crops Research for Semi-Arid Tropics (ICRISAT) and African Centre for Crop Improvement (ACCI) for genetic diversity. In this study, SSR markers were able to discriminate ACCI and ICRISAT accessions into separate cluster groups. PIC values ranged from 0.32 to 0.86, with an overall mean of

0.62, which implies high level of genetic variability. These results suggest reliability of SSR markers in genetic diversity studies.

2.12.3.2 Single nucleotide polymorphisms (SNPs)

A SNP is an individual nucleotide base variation between two DNA sequences. SNPs can be categorized according to nucleotide substitution as either transitions (C/T or G/A) or transversions (C/G, A/T, C/A or T/G). For instance, sequenced DNA fragments from two different individuals, AAGC**C**TA to AAGC**T**TA, comprise a single nucleotide difference (Edwards et al., 2007). Gupta et al. (2001) reported that SNPs are the most abundant molecular markers which are highly polymorphic and are closely distributed across genome. Jehan and Lakhanpaul (2006) stated that the abundance of SNPs in nature offsets their limitation of being biallelic and less informative. In a genetic diversity and association mapping study of sweet sorghum genotypes using SNP markers, Murray et al. (2009) observed three main genetic clusters of sweet sorghums. Based on observed phenotypes and known background, sweet sorghum accessions were clustered into three groups as historical and modern syrup, modern sugar/energy types, and amber types. Using information on population structure and relatedness, marker-trait association mapping was done for height and stem sugar (brix) traits. Three significant associations for height were also identified.

Using SNP markers to study complex traits such flowering time in maize, Buckler et al. (2009) reported that the genetic architecture of anthesis time is controlled by small additive QTL rather than a single large effect QTL. In wheat genomic study by Akhunov et al. (2009), gene-based SNP markers were established which confer resistance to leaf rust and powdery mildew diseases. In a study by Yu et al. (2011) on QTL analysis in rice for yield and three-yield-component traits, number of tillers per plant, number of grains per panicle, and grain weight compared a SNP-based map to that of a previous RFLP/SSR-based QTL map generated using the same mapping population. Using the ultra-high-density SNP map, the authors reported that this map was more powerful and had high resolution unlike the RFLP/SSR map. As suggested by Mammadov et al. (2012), beside the power and the resolution, maps based on high-density SNP markers are also highly appropriate for fine mapping and cloning of QTL.

2.13 Summary of literature review

This literature review has established that sweet sorghum is a potential crop for biofuel production due to the high level of fermentable sugars in its stalks and its ability to thrive in marginal environments. The review has further established that:

- There is a considerable level of genetic diversity on central and southern African sweet sorghum accessions.
- Inheritance of brix and stalk weight in sugar stalk is subject to both additive and non-additive gene action.
- Sweet sorghum is affected by G x E interactions of sugar yield and related traits.
- SNP markers have previously been used before in genetic diversity study analysis of both grain and sweet sorghum. They have gained popularity because of their high frequency in the genome and their biallelic nature.
- Little effort has been directed towards characterisation of sweet sorghum accessions for biofuel production in South Africa.
- Most of the cultivated sweet sorghum cultivars in south Africa are landraces and cultivated by smallholder farmers which have not been characterised for biofuel production.

References

Abubakar, L. and T.S. Bubuche. 2013. Genotype x environment interaction on biomass production in sorghum (*Sorghum bicolor* L. Moench) in North-Western Nigerian. African Journal of Agricultural Research 8: 4460-4465.

Akhunov, E., C. Nicolet and J. Dvorak. 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. Theoretical and Applied Genetics 119: 507-517.

Almodares, A. and M.R. Hadi. 2009. Production of bioethanol from sweet sorghum: A review. African Journal of Agricultural Research 5: 772-780.

Aslam, M., A.N. Khan, M.S. Mirza, Mi I.K and U. Ullah. 1993. Evaluation of different soybean genotypes for stability in yield performance across diverse environments. Pakistan Journal of Agricultural Resource 14: 297-303.

Audilakshmi, S., A.K. Mall, M. Swarnalatha and N. Seetharama. 2010. Inheritance of sugar concentration in stalk (Brix), sucrose content, stalk and juice yield in sorghum. Biomass Bioenergy 34: 813-820.

Balole, T.V. and G.M. Legwaila. 2005. *Sorghum bicolor* (L.) Moench.) PROTA Wageningen, Netherlands.

Buckler, E.S., J.B. Holland and P.J. Bradbury. 2009. The genetic architecture of maize flowering time. Science 325: 714-718.

Burks, P.S., C.M. Kaiser, E.M. Hawkins and B. P.J. Brown. 2015. Genomewide Association for Sugar Yield in Sweet Sorghum. Crop Science 55: 2138-2148.

Calvino, M. and J. Messing. 2013. Discovery of microRNA169 gene copies in genomes of flowering plants through positional information. Genome Biology Evolution 5: 402-417.

Calvino, M., R. Bruggmann and J. Messing. 2008. Screen of genes linked to high sugar content in stems by comparative genomics. Rice 1: 166-176.

Chinyama, S. 2016. Assessment of Sweet Sorghum Lines for Genetic Diversity Using Quantitative Traits and SSR Markers. MSc dissertation, University of KwaZulu-Natal, South Africa.

Chistiakov, D.A., B. Hellemans and F.A.M. Volckaert. 2006. Microsatellites and their genomic distribution, evolution, function and applications. A review with special reference to fish genetics Aquaculture 255: 1-29.

Cox, T.S., Y.T. Kiang, M.B. Gorman and D.M. Rodgers. 1985. Relationship between coefficient of parentage and genetic similarity indices in the soybean. Crop Science 25: 529-532.

Das, I.K., B. Fakrudin and D.K. Arora. 2008. RAPD cluster analysis and chlorate sensitivity of some Indian isolates of *Macrophomina phaseolina* from sorghum and their relationships with pathogenicity. Microbiological Research 163: 215-224.

Dewey, D.R. and K.H. Lu. 1959. A correlation and path coefficient analysis of components of crested wheatgrass seed production. Agronomy Journal 51: 515-518.

Doggett, H. 1970. Sorghum (Tropical Agriculture Series). Longmans, Green & Company., Limited, London.

Edwards, D., J.W. Forster, D. Chagne and J. Batley. 2007. What is SNPs? Association Mapping in Plants: 41-52.

Felderhoff, T.J., S.C. Murray, P.E. Klein, A. Sharma, M.T. Hamblin, S. Kresovich, et al. 2012 QTLs for energy-related traits in a sweet × grain sorghum mapping population. Crop Science 52:20-40.

Fentie, M., A. Assefa and K. Beleke. 2013. AMMI analysis of yield performance and stability of finger millet genotypes across different environments. World Journal of Agricultural Sciences 9:231-237.

Franco, J., J. Crossa, J. Ribaut, J. Bertran, M. Warburton and M. Khairallah. 2001. A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. Theoretical and Applied Genetics 103: 944-952.

Folkertsma, R., H. Frederick, W. Rattunde, S. Chandra, G. Raju and C. Hash. 2005. The pattern of genetic diversity of Guinearace *Sorghum bicolor* (L.) Moench landraces as revealed with SSR markers. Theory and Applied Genetics 111: 399-409.

Ghazy, M.F., M.S. Shadia and N. Magda. 2012. Stability analysis and genotype x environment interactions for forage sorghum hybrids (*Sorghum bicolor*, L. Moench). Agricultural Research 38: 142-153.

Gupta, P.K., J.K. Roy and M. Prasad. 2001. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Current Science 80: 524-535.

Gutjahr, S., A. Clément-Vidal, A. Soutiras, N. Sonderegger, S. Braconnier, M. Dingkuhn and D. Luquet. 2013. Grain, sugar and biomass accumulation in photoperiod-sensitive sorghums. II. Biochemical processes at internode level and interaction with phenology. Functional Plant Biology 40: 355-368.

Haussmann B.I.G., Obilana A.B., Ayiecho P.O., Blum A., Schipprack W. and G. H.H. Gieger. 2000. Yield and yield stability of four population types of grain sorghum in a semi-arid area of Kenya. Crop Science 40: 319-329.

Hoffmann-Thoma, G., K. Hinkel, P. Nicolay and J. Willenbrink. 1996. Sucrose accumulation in sweet sorghum stem internodes in relation to growth. Plant Physiology 97: 277-284.

House, L.R. 1980. A guide to sorghum breeding. International Crops Research Institute for the Semi-Arid Tropics ICRISAT, Andhra Pradesh 502 324, India

House, L.R. 1985. A Guide to Sorghum Breeding second edition. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Andhra Pradesh 502 324, India

Huligol, R.V., and M. Govind. 2004. A trial with sweet sorghum. Alternative uses of sorghum and pearl millet in Asia: proceedings of the Expert Meeting India. ICRISAT.

Jehan, T. and S. Lakhanpaul. 2006. Single nucleotide polymorphism (SNP)-Methods and applications in plant genetics: A review Indian Journal of Biotechnology 5: 435-459

Jiangfeng, H.E., Z. Xiaoqing, A. Laroche, L. Zhen-Xiang, L. Hongkui and L. Ziqin. 2014. Genotyping by sequence (GBS), an ultimate marker assisted selection (MAS) tool to accelerate plant breeding. Plant Science 5: 1-8.

Johnson, F.X. and E. Matsika. 2006. Bio-energy trade and regional development: the case of bio-ethanol in southern Africa. Energy for Sustainable Development 10: 42-53.

Karp, A., S. Kresovich, K.V. Bhatt, W.G. Ayad and T. Hodgkin. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies IPGR technical bulletins International Plant Genetics Resources Institute, Rome, Italy.

Kim, M. and D. Day. 2011. Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. Journal of Industrial Microbiology & Biotechnology 38: 803-807.

Kimberley, B., K.B. Ritter, C.L. McIntyre, I.D. Godwin, D.R. Jordan and S.C. Chapman. 2007. An assessment of the genetic relationship between sweet and grain sorghums, within *Sorghum bicolor* ssp. *bicolor* (L.) Moench, using AFLP markers   Euphytica 157: 161-176.

Kisua, J., K. Mwikamba, M. Makobe and A. Muigai. 2015. Genetic diversity of sweet and grain sorghum using phenotypic markers. International Journal of Biosciences 6: 34-46.

Kumar C.G., Fatima A., Rao P.S., Reddy B.V.S., Rathore A., Rao R. N., Khalid S., Kumar A.A and Kamal A. 2010. Characterisation of Improved Sweet Sorghum Genotypes for Biochemical Parameters, Sugar Yield and Its Attributes at Different Phenological Stages. International Crops Research Institute for the Semi-Arid Tropics 12: 322-328.

Laopaiboon, L., P. Thanonkeo, P. Jaisil and P. Laopiboon. 2007. Ethanol production from sweet sorghum juice in batch and fed-batch fermentations by *Saccharomyces cerevisiae*. World Journal of Microbiology and Biotechnology 23: 1497-1501.

Lekgari, A. and D. I. Dweikat. 2014. Assessment of genetic variability of 142 sweet sorghum germplasm of diverse origin with molecular and morphological markers. open Journal of Ecology 4: 371-393.

Levinson, G. and G.A. Gutman. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Molecular Biology and Evolution 4 203-221.

Lingle, S.E. 1987. Sucrose metabolism in the primary culm of sweet sorghum during development. Crop Science 27: 1214-1219.

Lombardi G.M.R., Nunes J.A.R., Parrella R.A.C., Teixeira D.H.L., Bruzi A.T., Durães N.N.L., Fagundes T.G. 2015. Path analysis of agro-industrial traits in sweet sorghum. Genetics and Molecular Research 14:16392-16402.

Ludewig, F. and U.I. Flugge. 2013. Role of metabolite transporters in source-sink carbon allocation. Front Plant Science 4: 231.

Lynch, M. and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates.

Mahajan, R., P. Wadikar, S. Pole and M. Dhuppe. 2011. Variability, correlation and path analysis studies in sorghum. Research Journal of Agricultural Sciences 2: 101-103.

Makanda, I., P. Tongoona and J. Derera. 2009. Quantification of genotypic variability for stem sugar accumulation and associated traits in new sweet sorghum varieties. African Crop Science Conference Proceedings 9: 391-398.

Mammadov, J., R. Aggarwal, R. Buyyarapu and S. Kumpatla. 2012. SNP Markers and Their Impact on Plant Breeding. International Journal of Plant Genomics 2012: 388-413.

Mather, K. and P.D.S. Caligari. 1976. Genotype x environment interactions IV. The effect of the background genotype. Heredity 36: 41-48.

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

Mehmood, S., A. Bashir, A. Ahmad, Z.K. Akram, N. Jabeen and M. Gulfraz. 2008   Molecular Characterisation of Regional Sorghum bicolor varieties from Pakistan. Pakistan Journal of Botany 40: 2015-2021.

Menz, M., R. Klein, N. Unruh, W. Rooney, P. Klein and J. Mullet. 2004. Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Science 44: 1236-1244.

Mittal, N. and A.K. Dubey. 2009. Microsatellite markers - A new practice of DNA based markers in molecular genetics. Pharmacognosy Reviews 3 235-246.

Muii, C.W., R.M. Muasya and D.T. Kirubi. 2013. Baseline survey on factors affecting sorghum production and use in eastern Kenya. African Journal of Food, Agriculture, Nutrition and Development 13 7339-7353.

Murphy, R.L., D.T. Morishige, J.A. Brady, W.L. Rooney, S. Yang, P.E. Klein and J.E. Mullet. 2014. Ghd7 (Ma6) represses Sorghum flowering in long days: Ghd7 alleles enhance biomass accumulation and grain production. Plant Genome 7: 1-10.

Murray, S.C., W.L. Rooney, M.T. Hamblin, S.E. Mitchell and S. Kresovich. 2009 Sweet sorghum genetic diversity and association mapping for brix and height The Plant Genome 2: 48-62.

Murray, S.C., A. Sharma, W.L. Rooney, P.E. Klein, J.E. Mullet, S.E. Mitchell and S. Kresovich. 2008. Genetic improvement of sorghum as a biofuel feedstock I: QTL for stem sugar and grain nonstructural carbohdrates. Crop Science 48: 2165-2179.

Musango, J.K. and A.C. Brent. 2011. A conceptual framework for energy technology sustainability assessment. Energy for Sustainable Development 15: 84-91.

Ngugi, K. and C.M. Onyango. 2012. Analysis of the molecular diversity of Kenyan sorghum germplasm using microsatellites. Crop Science Biotechnology 15 189-194.

Pfeiffer, T., M. Montross and M. Barrett. 2013. Sweet Sorghum for Biofuel. University of Kentucky Agricultural Communications. U.S.A. p. 1-3.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and A. Rafalski. 1996. The Comparison of RFLP, RAPD, AFLP and SSR (Microsatellite) Markers for Germplasm Analysis. Molecular Breeding 3: 225-238.

Prasad, S., A. Singh, N. Jain and H.C. Joshi. 2007. Ethanol production from sweet sorghum syrup for utilization as automotive fuel in India, Energy & Fuels. Plant Science 21: 2415-2420.

Qazi, H.A., S. Paranjpe and S. Bhargava. 2012. Stem sugar accumulation in sweet sorghum-activity and expression of sucrose metabolizing enzymes and sucrose transporters. Plant Physiology 169: 605-613.

Quinby, J.R. and R.E. Karper. 1947. The effect of short photoperiod on sorghum varieties and first-generation hybrids. Agricultural Research 75: 295-300.

Quinby, J.R. and R.E. Karper. 1954. Inheritance of height in sorghum. Agronomy 46: 211-216.

Rakoczy-trojanowask, M. and H. Bolibok. 2004 Characteristics and a comparison of three classes of microsatellite-based markers and their application in plants. Cellular and Molecular Biology Letters: 221-238.

Ratnavathi, C.V., K. Suresh, B.V.S. Kumar, M. Pallavi, V.V. Komala and S. N. 2010 Study on genotypic variation for ethanol production from sweet sorghum juice. Biomass Bioenergy 34: 947-952.

Reddy, B.V.S. and Reddy, S.P. 2003. Sweet sorghum: characteristics and potential. International Sorghum and Millets Newsletter p. 26-28.

Reddy, B.V., S. Ramesh, P.S. Reddy, B. Ramaiah, M. Salimath and R. Kachapur. 2005. Sweet Sorghum-a Potential Alternate Raw Material for Bio-Ethanol and Bio-Energy. International Sorghum and Millets Newsletter 46: 79-86.

Reddy, B.V., A.A. Kumar and S. Ramesh. 2007. Sweet sorghum: A water saving bioenergy crop. International Crops Research Institute for the Semi-Arid Tropics. Patancheru-502 324, Andhra Pradesh, India.

Reddy, P.S., B.V.S. Reddy and P.S. Rao. 2014. Genotype by sowing date interaction effects on sugar yield components in sweet sorghum (*Sorghum bicolor* L. Moench). Jourbal of Breeding and Genetics 46: 241-255.

Ritter, K.B., M.C. Lynne, I.D. Godwin, D.R. Jordan and S.C. Chapman. 2007 An assessment of the genetic relationship between sweet and grain sorghums, within *Sorghum bicolor* ssp. *bicolor* (L.) Moench, using AFLP markers. Euphytica 157: 161-176

Rono J.K., Cheruiyot E.K, Othira J.O, Njuguna V.W., Macharia J.K, Owuoche J., et al. 2016. Adaptability and Stability Study of Selected Sweet Sorghum Genotypes for Ethanol Production under Different Environments Using AMMI Analysis and GGE Biplots. The Scientific World Journal 2016: 1-14.

Salas Fernandez, M.G., P.W. Becraft, Y. Yin and T. Lubberstedt. 2009. From dwarves to giants? Plant height manipulation for biomass yield. Trends Plant Science 14: 454-461.

Sanderson, M.A., R.M. Jones, J. Ward and R. Wolfe. 1992. Silage sorghum performance trial at Stephenville. Forage Research in Texas. Texas Agricultural Experimental Station, Stephenville, USA.

Seetharama, N., D.B. Rao, C.V. Ratnavathi, S.M. Parwez, B. Mathew, K. Singh, et al. 2002. Sweet sorghum-an ancillary sugar crop. Indian Farming 36: 7-8.

Semagn, K., A. Bjornstad and M.N. Ndjiondjop. 2006. An overview of molecular marker methods for plants. African Journal of Biotechnology 5: 2540-2568.

Sergio, L., and B. Gianni. 2005. Molecular markers-based analysis for crop germplasm preservation. FAO Meeting on the role of biotechnology for the characterisation and conservation of crops, forestry, animal and fishery genetic resources, Turin, Italy.

Serna-Saldivar, S. 2010. Cereal Grains: Properties, Processing, and Nutritional Attributes. CRC Press. United States of America.

Shiringani, A. and M. Frisch. 2010. Genetic mapping of QTLs for sugar-related traits in a RIL population of *Sorghum bicolor* L. Moench Theoretical and Applied Genetics 121: 323–336.

Shukla, S., T.J. Felderhoff, A. Saballos and W. Vermerris. 2017. The relationship between plant height and sugar accumulation in the stems of sweet sorghum (*Sorghum bicolor* (L.) Moench). Field Crops Research 203: 181-191.

Sneath, P.H.A. and R.R. Sokal. 1973. Numerical Taxonomy Freeman, London.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology 98: 503-517.

Srinivasa R.P., Rao S.S., Seetharama N., Umakath A.V., Reddy S.P, Reddy, B.V.S and Gowda, C.L.L. 2009. Sweet sorghum for biofuel and strategies for its improvement. International Crops Research Institute for the Semi-Arid Tropics., Andhra Pradesh, India: p. 80.

Srinivasarao, P., C.G. Kumar, R.S. Prakasham, A.U. Rao, B.V.S. Reddy. 2015. Sweet sorghum: Breeding and bioproducts. Industrial crops. New York.

Tarpley, L. and D.M. Vietor. 2007. Compartmentation of sucrose during radial transfer in mature sorghum culm. Plant Biology 7: 1-10.

Tautz, D. and M. Renz. 1984 Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Research 12: 4127-4138.

Tsuchihashi, N. and Y. Goto. 2008. Year-round cultivation of sweet sorghum [*Sorghum bicolor* (L.) Moench] through a combination of seed and ratoon cropping in Indonesian Savanna. Plant Production Science 11: 377-384.

Upadhyaya, H., C. Gowda and D. Sastry. 2008. Plant genetic resources management: collection, characterisation, conservation and utilization. Journal of Semi-Arid Tropics Agricultural Research 6: 1-16.

Wang, M.L., C. Zhu, N.A. Barkley, Z. Chen, J.E. Erpelding, S.C. Murray, M.R. Tuinstra, T. Tesso, G.A. Pederson and J. Yu. 2009. Genetic diversity and population structure

analysis of accessions in the US historic sweet sorghum collection. Theoretical and Applied Genetics 120: 13-23.

Wright, S. 1921. Correlation and causation. Journal of Agricultural Research 20: 557-587.

Yan, W. and M.S. Kang. 2002. GGE Biplot Analysis: A Graphical Tool for Breeders. CRC Press.

Yang, M., L. Liu, L. Li, B. Wang, J. Chang and L. G. 2009. Correlation analysis between sucrose synthase expression and sucrose accumulation in sweet sorghum [*Sorghum bicolor* (L.) Moench]. Acta, Agronomic Sinca 35: 185-189.

Xu, Y. 2010. Molecular Plant Breeding. International Maize and Wheat Improvement Center (CIMMYT), Mexico.

Yu, H., W. Xie and J. Wang. 2011. Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. PLoS ONE 6.

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). Agricultural Science China 5: 736-744.

Zegada-Lizarazu, W. and A. Monti. 2012. Are we ready to cultivate sweet sorghum as a bioenergy feedstock? A review on field management practices. Biomass Bioenergy 40: 1-12.

CHAPTER 3

Characterisation of sweet sorghum germplasm based on morpho-agronomic traits and industrial traits under two environments.

Abstract

Sweet sorghum has the potential to become a major bioenergy source because of high sugar accumulation in its juicy stalk, and its low production input requirements. Exploring genetic diversity of sweet sorghum is significant to generate reliable information for development and improvement of the crop as an energy source. This study aimed at characterising sweet sorghum germplasm using agro-morphological traits under two environments (Potchefstroom and Ukulinga). Fifty sweet sorghum germplasm accessions from Agriculture Research Council-Grain Crops (ARC-GC) were used in the study. The experiment was laid out in a 10 x 5 alpha lattice design with three replicates in both environments. Wide genetic diversity was observed on almost all traits studied. Under combined analysis of variance, the studied sweet sorghum accessions were classified as early to medium maturing. The number of days to flowering ranged from 59 to 100 days. Juice yield ranged from 437 l/ha to 22000 l/ha, while %brix ranged from 3.1 to 19.2%. Genotype by environment (G x E) interaction was significant for almost all the traits under study. The principle component analysis produced four components which were effective and accounted for 79.12% of the variance. The principle biplot showed some quantitative traits that were positively and significantly correlated. The study identified SA4490, SA2400, SA4495, SA2193 and SA4479 as superior accessions in juice yield. These accessions should be used as parents in sweet sorghum improvement programme.

Key words: biofuel, genetic diversity, principle component analysis, sweet sorghum

3.1 Introduction

There is an increasing need for substitute energy sources because of the decline in fossil fuel production and environmental concern. The most common biofuel raw materials are maize and sugar cane. However, the use of these feedstocks poses a threat on food security with the rise in human population (Serna-Saldívar et al., 2012). The relative rise in food prices has also been attributed primarily to the use of maize for bioethanol apart from other factors like drought or changes in global consumption (Brinkman et al., 2009).

Therefore, the use of food crops such as maize cannot support the objectives of renewable fuel legislation. An alternative bioenergy crop such as sweet sorghum is thus needed. Sweet sorghum has several advantages such as high biomass yield and growth rate, minimum input requirements and is fairly tolerant to multiple stresses among others (Jessup, 2009).

There are many types of sweet sorghum germplasm distributed globally, providing a diverse genetic base from which to develop regionally specific, highly productive cultivars (Audilakshmi et al., 2010). Despite all the agronomic advantages sweet sorghum has as a bioenergy crop, it received little scientific research effort toward the identification of traits relevant to biofuel production. Therefore, there is a need to characterise existing sweet sorghum germplasm for both agronomical and industrial traits.

Characterisation of germplasm is essential for the sustainable maintenance and increased use of crop genetic resources (Sergio and Gianni, 2005). This involves identifying heritable traits, leading to classification that facilitates enhanced utilisation of germplasm and successful breeding programme (Upadhyaya et al., 2008). Morphological and molecular characterisation is recommended in accessions collected from different regions to ensure a more comprehensive and informative characterisation (Atokple, 2003).

Plant selections are done based on morphological characteristics that can be discriminated and that are co-inherited with the desired trait. Morphological traits displayed by sweet sorghum landraces determine selection and preference by the farmers and breeders (Franco et al., 2001). According to Ali et al. (2008) and Murray et al. (2009), the most useful morphological traits for selection in sweet sorghum are plant height, stalk diameter, juice yield and brix.

As suggested by Lynch and Walsh (1998), the phenotypic variations in sweet sorghum, just like any other crop, are due to genetic complexity from multiple interacting loci, with allelic effects that are influenced by the environment. Many important traits in sorghum, such as

height are quantitative or polygenic, showing continuous variability. According to Semagn et al. (2010), a quantitative trait is a measurable feature, attribute or characteristic that depends on the cumulative action of many genes and their interaction with the environment which can differ among individuals over a given range to produce a continuous distribution of phenotypes. These phenotypic variations account for the genetic diversity.

Determination of genetic diversity by phenotypic measurements in sorghum is a gateway to the study of evolutionary forces that control their existence. Phenotypic estimates are used to reveal the extent of genetic relatedness and therefore the similarity in phenotype characteristics may show genetic similarity of genotypes (Kisua et al., 2015). In this study, sweet sorghum accessions were characterised for genetic diversity and relatedness using phenotypic traits for identification of potential parents for future sweet stem sorghum breeding programme.

3.2 Hypothesis

The hypothsis being tested for this study is that there is no significant variation in the agro-morphological traits among sweet sorghum lines grown under two different environments.

3.3 Materials and methods

Fifty sweet sorghum germplasm (Table 3.1) were used in this study. These were South African accessions and were obtained from the Agricultural Research Council-Grain Crops (ARC-GC).

Table 3.1   List of sweet sorghum genotypes used in the study

Entry No.	Accession Designation	Entry No.	Accession Designation
1	SA 0096	26	SA 4477
2	SA 0163	27	SA 4479
3	SA 0240	28	SA 4481
4	SA 0311	29	SA 4482
5	SA 0312	30	SA 4483
6	SA 0317	31	SA 4484
7	SA 0831a	32	SA 4485
8	SA 1238	33	SA 4486
9	SA 1242	34	SA 4488
10	SA 1330	35	SA 4489
11	SA 1382	36	SA 4490
12	SA 160	37	SA 4491
13	SA 1904	38	SA 4494
14	SA 2012	39	SA 4495

Entry No.	Accession Designation	Entry No.	Accession Designation
15	SA 2013	40	SA 4501
16	SA 2014	41	SA 4503
17	SA 2034	42	SA 4506
18	SA 2036	43	SA 4510
19	SA 2116	44	SA 4520
20	SA 2193	45	SA 4523
21	SA 2218	46	SA 4524
22	SA 2249	47	SA 4525
23	SA 2330a	48	SA 4526
24	SA 2400	49	SA 4528
25	SA 4476	50	SA 4534

3.3.1   Experimental sites

The field trials were conducted during the 2016/2017 summer season in Potchefstroom (ARC-GC Research Farm), North West Province and at Ukulinga Research farm in Pietermaritzburg, KwaZulu-Natal Province. Potchefstroom is located at 26°74'S latitude; 27°8'E longitude and altitude of 1344 m above sea level. It experiences an average minimum and maximum temperature of 9.61°C and 25.48°C, respectively. It is characterised by loamy clay soils and receives average annual rainfall of 618.9 mm. Ukulinga Research farm is located at latitude of 29°67'S, longitude of 30°41'E and 806 m above sea level. It is characterised by clay soils. It experiences a mean annual temperature of 17.9°C and receives a mean annual rainfall of 738 mm.

3.3.2   Field trial design and management

Experiments were laid out in a 10 x 5 alpha lattice design as explained by Cochran and Cox (1992) with three replications. The experimental materials were planted in a single row of 5 m long with inter-row spacing of 0.75 m and intra-row spacing of 0.2 m at both sites. Planting was done on 7th December 2017 and 14th December 2017 for Potchefstroom and Ukulinga, respectively. The experiments were conducted under rain fed conditions with supplementary irrigation when needed. A compound fertilizer (3:2:1) NPK was applied as at a rate of 100 kg/ha at Potchefstroom while at Ukulinga a compound fertiliser of 2:3:4 (NPK) was applied at a rate of 250 kg/ha. A mixture of pre-emergence herbicides Dual Gold and Basagran® was applied during planting to control weeds. This was augmented by hand weeding to keep the trial fields weed free, throughout the growing season. All standard agronomic practices were carried out as required.

3.4 Data collection

Quantitative characters were measured based on sorghum descriptors (IBPGR, 1993). Five randomly selected plants were tagged on each plot, and the data were collected. Date to flowering data, however, was collected from the whole plot.

Days to 50% flowering: This was recorded as the duration of days from planting, up to the time when 50% of the plants within a plot had started shedding pollen.

Plant height: This was measured as the distance from the base of the plant to the tip of the panicle of the main stem. A measuring tape was used to measure the height. This was expressed in centimetres (cm).

Stalk diameter (thickness): This was done to determine the external thickness of the plant. It was measured using a Vernier caliper, and it was expressed in millimetres (mm). It was done during harvesting.

Number of leaves per plant: This was done by counting the total number of leaves from the base of the plant up to the last leaf.

Lodging score: This was done by determining the percentage at which plants on the plot had fallen to the ground. A scale of 1 to 5 was used, where 1 = no lodging while 5 = all plants on the plot lodged.

Pest score (sweet sorghum stem borer): Sweet sorghum stem borer scoring was done according to Davis and Williams (1992) scale rating. It was done using a scale of 0 to 9, where 0 = no visible damage and 9 = whole plot was destroyed by the pests.

Stalk yield: This was determined by weighing the total mass of the harvested stalks (without leaves and panicles) before crushing. It was expressed in kilograms and converted to t/ha. In each plot, a 1 m row was harvested in the middle. Harvesting was done at hard dough stage for each entry.

The stalks were manually harvested, and transported to the laboratory for juice extraction and further data collection. Juice extraction was done using a hydraulic three roller press.

Brix: This was measured by chopping five representative stalks and dropping the juice on the hand-held refractometer (a laboratory device for measurement of sugar content in sweet sorghum stalk). Chopping was done on the 4th internode from the tip. Three drops were collected on the refractometer. The readings were expressed in percentage (%).

Bagasse fresh weight: This was done by weighing the bagasse soon after crushing and extracting juice. It was expressed in kilograms and then converted to tonnes per hectare (t/ha).

Juice volume: Juice volume was measured using a measuring cylinder. It was expressed in millilitres (ml) and converted to litres per hectare (l/ha).

Panicle length: This was measured from the base of panicle to the tip of the panicle, where the first branch starts, and it was done at maturity.

Panicle width: This was measured as width of panicle in natural position on the widest part, and it was expressed in cm.

Panicle weight: This was done by weighing dry panicles per plot before threshing, then it was converted to tonnes per hectare (t/ha).

1000 grain weight: This was taken by weighing 1000 grains at 12% moisture content, using a digital scale.

3.5 Statistical analysis

3.5.1 Analysis of variance (ANOVA)

Analysis of variance for the quantitative data was performed using GenStat statistical package, 18th edition, using the unbalanced treatment structural design. This was aimed at determining the genetic diversity of the measured quantitative traits. Descriptive statistics (mean value, coefficient of variation (%CV), least significant difference at 5% level (LSD 0.05) were used to compare levels of variation for morphological and agronomic traits both within and between two environments in this field experiment.

3.5.2 Principal component analysis

The quantitative standardised means for the two locations (Potchefstroom and Ukulinga) were subjected to principal component analysis (PCA) in a Multivariate analysis to assess the proportional contribution of each trait to the entire genetic variation, as suggested by Reddy et al. (2009).

3.6 Results

3.6.1 Weather data

The mean seasonal temperature (Table 3.2) for the two sites varied. Ukulinga recorded higher mean minimum and maximum temperatures (18°C and 33°C respectively) than Potchefstroom.

Table 3.2 Data for mean temperatures and rainfall for Potchefstroom and Ukulinga during 2016-2017 growing season

Month	Potchefstroom			Ukulinga		
	T_n	T_x	Rain (mm)	T_n	T_x	Rain (mm)
Oct-16	12	30	55	7	35	65
Nov-16	15	30	95	12	37	76
Dec-16	17	33	20	14	38	33
Jan-17	16	28	29	17	27	70
Feb-17	17	27	27	18	27	94
Mar-17	15	28	34	12	37	32
Apr-17	10	25	46	14	36	37
May-17	5	23	11	7	31	57
Jun-17	2	22	0	6	28	1
Total			317	_	_	465
Mean	12	27	_	12	33	_

T_n = Minimum temperature (°C), T_x = Maximum temperature (°C)

3.6.2 Analysis of quantitative traits at Potchefstroom.

The analysis of variance for almost all quantitative traits at Potchefstroom indicated a highly significant difference ($p < 0.001$) (Table 3.3). This indicates presence of a wide genetic variation. A wide variation was observed in number of days to 50% flowering. The range was 57 to 101 days. The observed grand mean for number of days to 50% flowering was 81. SA 0311 flowered early (57 days) in Potchefstroom while SA 4481 flowered late (101 days). Wide genetic diversity existed in plant height as well. Plant height ranged from 124.4 cm to 328.3 cm. SA 4482 and SA 4495 were respectively the shortest and tallest accessions. The mean height observed for all the accessions was 215.8 cm. The lowest stalk diameter (thickness) was observed in SA 0311 with a value of 10.0 mm. SA 4490 had the highest stalk diameter of 21.0 mm; the grand mean, however was 17.0 mm. The number of leaves per plant ranged from 7.0 for SA 0240 to 15.0 for SA 2193, with a mean of 11.0. Non-significant differences were observed on lodging score in Potchefstroom environment. The lowest lodging score was

1.0 for almost all the genotypes under study, while the highest score of 4.0 was observed for SA 4494.

Pest damage also varied amongst the genotypes. The lowest score of 2.0 was observed in SA 0096 and SA 0311, whilst the highest stem borer score observed was 6.0 in SA 1238, SA 4489 and SA 4523. A wide range of genetic diversity was also manifested in stalk yield where the yield ranged from 8.4 t/ha for SA 0240 to 57.9 t/ha for SA 2400. The grand mean for stalk yield was 31.6 t/ha. The mean for panicle length was 26.2 cm, with a minimum length of 13.4 cm for SA 2193 and a maximum length of 36.2 cm for SA 2116. The mean for panicle width was 6.7 cm while the minimum panicle width was observed in SA 2193 with a value of 4.4 cm. SA 2116 had a maximum width with a value of 9.3 cm. Non-significant differences were observed in weight of panicles where the mean was 3.6 t/ha. SA 4481 had the lowest panicle weight of 1.2 t/ha while the highest panicle weight was observed in SA 4479 (8.7 t/ha). Variation was also observed in 1000 grain weight. SA 4491 had the least weight of 9.3 g while SA 4525 had the highest weight of 27.4 g for 1000 grains. The mean for 1000 grain weight was 16.1 g.

Brix readings for sugar content varied greatly amongst accessions under study. The lowest reading was observed in SA 2034 with a value of 3.1%. SA 1330 had the highest sugar content with a value of 19.2%. The mean sugar content was 10.4%. Genetic diversity was also observed in juice volume with a minimum of 350 l/ha to 20 178 l/ha for SA 0240 and SA 4490, respectively. The mean for juice volume was 9394.0 l/ha. Bagasse fresh weight ranged from 4.3 t/ha for SA 0240 to 30.7 t/ha for SA 2400, with an overall mean of 17.7 t/ha.

Table 3.3   Means values for 14 quantitative traits measured at Potchefstroom, 2017

Accession	DFL	PHT	SD	LPP	LS	PS	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
SA0096	98.0	228.4	18.1	12.5	1.0	2.0	36.4	20.6	6.0	3.8	9.8	14.7	9213.0	20.2
SA0163	73.3	156.1	14.7	10.1	1.0	5.3	17.9	29.5	7.5	4.5	23.4	4.5	3692.0	12.4
SA0240	60.7	125.3	11.3	7.2	1.0	5.0	8.4	25.0	6.4	2.8	21.1	4.9	350.0	4.3
SA0311	56.7	144.1	10.4	7.8	1.0	2.0	12.1	22.0	5.2	2.0	19.7	13.6	4210.0	7.2
SA0312	83.0	201.1	20.3	11.4	1.0	2.7	40.4	23.0	6.8	5.0	11.1	11.7	11956.0	25.5
SA0317	75.3	212.7	18.7	11.9	1.0	2.3	39.1	18.0	5.0	4.5	24.8	9.1	11750.0	22.1
SA0831a	72.3	185.1	16.0	9.8	1.0	3.0	35.8	24.4	6.1	4.3	14.6	6.7	10434.0	13.6
SA1238	63.7	163.8	12.1	9.1	1.0	6.3	9.5	34.2	7.9	3.4	20.2	6.6	1503.0	6.2
SA1242	84.7	202.5	15.8	11.7	1.0	3.0	27.6	22.7	5.5	2.2	9.8	11.2	8367.0	15.1
SA1330	91.0	281.3	20.3	12.6	1.0	2.7	38.2	18.8	5.5	1.5	15.9	19.2	12865.0	21.7
SA1382	81.7	269.1	14.2	11.1	1.0	2.7	27.7	34.0	8.0	4.6	13.6	13.6	8686.0	15.7
SA160	88.0	232.8	16.9	10.9	1.0	2.0	30.0	31.9	8.1	4.1	12.5	15.5	9259.0	18.1
SA1904	64.0	174.3	12.3	8.2	1.0	3.0	14.5	25.4	6.2	3.2	20.2	10.9	5462.0	8.1
SA2012	81.7	217.7	16.0	11.7	1.0	3.3	29.1	24.0	6.6	3.1	12.2	11.4	10934.0	15.1
SA2013	98.7	169.3	13.8	12.3	1.0	2.3	22.9	21.9	5.4	4.0	9.4	13.9	6044.0	11.4
SA2014	66.0	234.8	14.4	9.1	1.0	4.3	22.6	31.7	7.6	5.8	24.2	5.0	4958.0	14.0
SA2034	70.7	148.5	10.8	8.7	1.0	3.3	13.3	27.3	7.7	2.4	24.0	3.1	2482.0	8.3
SA2036	72.3	178.4	16.1	9.9	1.0	2.3	22.3	22.2	5.6	2.3	19.4	17.4	5733.0	13.0
SA2116	85.7	240.6	15.7	10.2	1.0	1.7	23.8	36.2	9.3	1.7	17.6	10.1	2300.0	15.3
SA2193	100.0	239.6	19.0	14.9	1.0	3.3	44.2	13.4	4.4	2.9	16.3	10.6	16133.0	24.8
SA2218	97.3	170.3	17.9	12.4	1.0	2.3	36.2	22.4	5.6	4.0	9.8	14.0	11000.0	22.2
SA2249	86.3	212.8	14.6	10.5	1.0	3.3	24.6	22.5	5.9	2.3	13.2	13.2	8271.0	13.1
SA2330a	88.7	215.2	15.6	11.8	1.0	2.3	31.0	22.6	5.7	2.5	12.4	13.4	9000.0	15.3
SA2400	80.7	189.7	18.2	11.3	1.0	2.7	57.9	25.0	7.3	3.5	12.7	13.1	18022.0	30.7
SA4476	82.3	223.2	18.4	10.1	1.0	4.3	26.6	28.7	7.1	3.0	14.3	13.8	9287.0	16.4
SA4477	97.7	201.4	19.8	12.9	1.0	2.0	51.9	24.0	6.4	3.1	11.7	11.7	12778.0	26.1
SA4479	70.0	242.0	18.6	11.6	1.0	4.7	46.3	30.8	8.3	8.7	20.5	7.2	15193.0	23.7

Accession	DFL	PHT	SD	LPP	LS	PS	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
SA4481	101.0	256.5	19.1	13.7	1.0	2.3	37.9	15.2	4.5	1.2	12.8	9.7	12444.0	20.9
SA4482	72.7	124.4	13.3	8.0	1.0	2.0	8.5	25.1	5.6	2.4	19.8	13.5	1098.0	4.9
SA4483	82.3	203.4	19.7	10.7	1.0	3.0	34.5	22.8	5.7	4.1	12.0	10.9	11982.0	20.5
SA4484	76.7	208.9	17.7	10.3	1.0	3.3	30.4	29.8	6.9	3.9	14.4	9.8	10883.0	14.3
SA4485	86.3	236.5	19.9	11.7	1.0	2.0	40.6	25.9	6.3	3.6	14.6	11.9	14318.0	24.3
SA4486	78.3	253.3	16.0	10.6	1.0	4.3	30.9	30.7	7.6	3.4	14.7	7.3	10887.0	16.0
SA4488	70.3	207.7	16.5	10.5	1.0	3.7	30.3	24.7	6.3	4.0	13.9	13.1	10102.0	14.6
SA4489	75.0	191.1	17.7	9.4	1.0	5.7	30.4	27.3	7.3	4.3	19.3	8.0	11167.0	15.8
SA4490	90.3	284.1	21.4	14.1	1.0	3.3	54.8	31.2	8.0	2.9	14.5	10.6	20178.0	28.0
SA4491	90.0	207.4	15.7	12.0	1.0	2.3	30.1	20.9	5.3	3.6	9.3	11.0	10222.0	15.1
SA4494	68.7	291.6	19.1	9.4	4.0	6.0	37.8	30.4	8.7	3.6	20.6	7.1	11667.0	21.5
SA4495	93.7	328.3	19.7	14.2	1.0	4.0	50.7	33.4	8.4	2.9	16.8	11.1	16778.0	26.4
SA4501	89.7	244.7	20.3	12.5	1.0	2.7	40.2	26.6	7.2	3.4	12.6	10.2	13111.0	20.5
SA4503	69.7	272.1	13.3	10.7	1.0	3.0	17.7	33.9	7.9	6.3	26.3	7.2	1804.0	13.7
SA4506	79.3	248.6	18.7	11.5	1.0	2.7	38.5	25.3	6.4	2.9	16.3	13.4	12333.0	22.1
SA4510	98.7	181.5	16.7	12.5	1.0	2.3	32.3	21.7	5.5	3.4	9.4	14.5	8491.0	14.6
SA4520	92.0	223.2	17.2	12.2	1.0	2.0	33.9	25.5	6.6	3.5	12.6	14.4	13156.0	17.2
SA 4523	69.3	145.4	14.1	9.7	1.0	6.3	15.3	23.0	6.7	2.1	23.1	3.8	2440.0	8.0
SA 4524	72.0	265.0	20.8	12.4	1.0	4.0	51.9	32.0	8.4	5.1	19.6	9.3	12178.0	30.0
SA 4525	70.7	246.6	14.4	10.7	1.0	3.0	21.1	35.6	8.7	6.5	27.4	7.3	2489.0	13.8
SA 4526	82.7	253.0	19.3	11.9	1.0	2.3	45.8	27.4	7.1	4.1	16.9	17.1	11311.0	23.5
SA 4528	77.7	215.9	16.1	10.7	1.0	3.3	34.6	26.4	6.9	4.4	13.5	11.1	10400.0	14.2
SA 4534	84.3	242.9	19.1	11.9	1.0	2.0	42.8	31.9	8.1	3.9	12.2	13.5	10378.0	23.9
Mean	80.8	215.8	16.7	11.1	1.1	3.2	31.6	26.2	6.7	3.9	16.1	10.9	9394.0	17.3
L.S.D $_{0.05}$	3.40	7.33	2.36	1.47	0.00	1.38	4.07	2.81	0.57	0.27	2.34	12.34	2381.00	11.46
%C.V.	2.59	2.09	8.69	8.20	0.00	26.62	7.93	6.61	5.25	4.23	8.94	1.10	15.62	1.62

DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS= Pest damage score (%), SY = Stalk yield (t/ha), PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix (%), JV = Juice volume, BFW = Bagasse fresh weight (t/ha)

Table 3.4   Mean squares for six quantitative traits measured at Potchefstroom, 2017

Source of Variation	DF	DFL	PHT	SD	LPP	LS	PS
Rep	2	23.13	45.5	45.13	3.96	0.00	0.38
Rep.IB	12	77.52***	1307.38***	5.04*	2.24**	484136959942327***	3.50***
Germplasm	49	364.11***	5575.77***	23.21***	7.96***	0.50***	3.85***
Residual	86	4.38	20.38	2.11	0.82	0.50	0.73
Total	149	128.82	1951.31	9.86	3.33	0.18	1.97

DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS = Pest damage score (%), IB = Incomplete bloc

Table 3.5   Mean squares for 8 quantitative traits measured at Potchefstroom, 2017

Source of Variation	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
Rep	17.12	9.60	0.77	0.07	1.70	8.37	12065170	10.81
Rep.IB	88.29***	42.67***	2.27***	0.07***	6.97***	16.18***	16314655***	23.75***
Germplasm	434.67***	71.09***	3.79***	0.03ns	69.99***	35.81***	61730551.00***	121.62***
Residual	6.29	2.99	0.13	0.03	2.08	1.81	2152477.00	3.91
Total	153.91	28.67	1.51	0.03	24.80	14.24	23018900.00	44.31

SY = Stalk yield (t/ha), PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix, JV = Juice volume, BFW = Bagasse fresh weight (t/ha)

3.6.3 Analysis of quantitative traits for Ukulinga

Analysis of variance for Ukulinga revealed significant differences for almost all quantitative traits under study (Table 3.6). The mean number of days to 50% flowering was 66. SA 1382 and SA 4481 were early and late flowering accessions with 56 and 98 days, respectively. Plant height ranged from 157.4 cm for SA 4482 to 295.6 cm for SA 4494. The grand mean was 339.4 cm. Stalk diameter varied from 10.0 mm for SA 0311 and SA 0240 to 30.0 mm for SA 4523, with an overall mean of 16.0 mm. There was also a significant difference in number of leaves per plant. SA 4484 had 7.0 leaves per plant while SA 2193 had 14.0 leaves per plant, with an overall mean of 11.0 leaves per plant. The mean lodging score was 1.0 for almost all accessions. The maximum value for lodging score was 3.0 observed in SA 4494.The lowest score for pest infestation observed was 1.0 in most accessions, whilst the highest score observed was 2.0 in a few accessions. The overall mean for pest score was 1.0.

A wide genetic diversity was also revealed in stalk yield per hectare. The yield ranged from 11.8 t/ha for SA 2034 to 75.5 t/ha for SA 4495 with an overall mean of 34.3 t/ha. Variation was also observed on panicle length and panicle width. The shortest panicle was 13.6 cm for SA 2193, while SA 2116 had the longest panicle of 36.5 cm. The grand mean for the panicle length was 25.3 cm. The mean for panicle width was 6.5 cm, and the minimum width was 3.9 cm for SA 1330 while the maximum width was 9.5 cm for SA 2249. Panicle weight ranged from 0.3 t/ha for SA 4483 to 7.4 t/ha for SA 4485 with a grand mean of 3.5 t/ha. Wide variation was exhibited in 1000 grain weight as well, where the minimum weight was 12.3 g for SA 4483 and the maximum was 31.1 g for SA 0163 with an overall mean of 20.7 g.

Significant differences were observed for the industrial traits such as juice brix, juice volume and fresh bagasse weight. The minimum brix value of 6.2 was observed in accession SA 0240 and a maximum 22.8 for SA 1330, with a grand mean of 16.3. The trend was extended to juice volume where the grand mean was 9605.0 l/ha with a minimum volume of 524 l/ha for SA 0240 and a maximum of 27 222 l/ha for SA 4495. Fresh bagasse weight ranged from 5.1 t/ha for SA 0240 to 35 t/ha for SA 4495. The overall mean was 16.8 t/ha.

Table 3.6   Means for 14 quantitative traits measured at Ukulinga, 2017

Accession	DFL	PHT	SD	LPP	LS	PS	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
SA 0096	67.7	234.3	17.3	11.0	1.0	1.3	38.7	19.5	6.2	3.9	16.4	16.5	12567.0	16.9
SA 0163	57.3	160.9	15.3	8.7	1.0	1.0	20.3	26.9	4.1	1.1	31.1	14.8	2409.0	9.5
SA 0240	56.7	163.5	11.1	8.8	1.0	1.0	14.6	24.2	4.6	0.8	23.4	6.2	524.0	5.1
SA 0311	73.0	188.4	10.5	8.8	1.0	1.0	21.3	20.2	4.4	2.1	20.0	20.5	4267.0	12.8
SA 0312	62.0	243.4	17.3	12.0	1.0	1.3	45.4	19.0	5.1	1.0	17.3	20.4	12667.0	20.8
SA 0317	67.3	228.3	12.3	10.2	1.0	1.0	26.7	16.1	4.0	0.5	24.1	19.4	4778.0	12.5
SA 0831a	68.7	204.6	16.9	10.9	1.0	1.0	34.7	22.8	9.0	5.5	19.7	16.4	10489.0	15.2
SA 1238	62.3	187.1	12.7	8.9	1.0	1.0	16.2	36.4	5.1	6.0	27.2	14.9	1342.0	9.4
SA 1242	64.0	235.3	16.6	10.7	1.0	1.3	35.0	22.5	8.1	5.3	22.6	17.0	11033.0	15.9
SA 1330	71.3	280.8	17.0	11.1	1.0	1.7	41.0	16.9	3.9	4.1	15.0	22.8	12667.0	22.2
SA 1382	55.7	244.5	13.4	10.6	1.0	1.0	24.8	33.0	7.7	3.5	18.4	14.8	8111.0	13.2
SA 160	63.7	242.2	13.3	11.7	1.3	1.0	27.3	25.9	8.3	3.4	16.7	17.7	7800.0	13.1
SA 1904	57.3	228.7	12.7	8.2	1.0	1.0	21.3	24.0	6.6	4.0	21.0	19.5	5080.0	10.6
SA 2012	62.7	226.9	17.0	11.3	1.0	1.7	34.7	22.9	8.9	6.7	22.2	15.2	11476.0	17.4
SA 2013	67.3	186.2	18.7	9.6	1.0	1.3	36.9	21.7	6.6	6.4	17.6	12.9	8900.0	15.6
SA 2014	67.0	279.2	15.3	10.8	1.0	1.7	33.0	33.2	5.6	3.4	25.0	13.4	11333.0	18.5
SA 2034	69.3	166.8	14.1	8.7	1.0	1.0	11.8	32.0	6.2	2.0	26.2	15.1	991.0	6.2
SA 2036	66.0	264.9	16.9	9.7	1.0	1.3	34.7	22.0	6.1	3.1	21.8	19.8	11222.0	15.4
SA 2116	69.3	287.7	13.1	9.3	1.0	1.0	28.6	36.5	6.3	1.1	14.7	16.0	3262.0	22.8
SA 2193	71.3	257.7	17.1	13.6	1.0	1.3	41.8	13.6	4.5	0.7	23.9	18.2	9778.0	25.0
SA 2218	65.0	197.0	16.0	11.4	1.0	1.0	27.7	21.3	7.5	5.9	16.4	16.5	7156.0	13.6
SA 2249	65.3	229.3	16.2	10.7	1.0	1.0	35.0	22.7	9.5	5.6	22.3	15.5	10000.0	16.9
SA 2330a	65.3	227.9	16.1	11.2	1.0	1.3	34.4	33.8	8.6	3.0	21.2	14.9	10844.0	18.0
SA 2400	72.3	267.7	18.3	11.6	1.0	1.3	38.5	21.3	8.1	1.2	20.3	19.4	11822.0	22.0
SA 4476	67.7	249.3	16.9	9.8	1.0	1.3	36.5	29.1	8.5	4.3	22.4	14.3	12178.0	17.5
SA 4477	72.0	242.5	19.5	10.4	1.0	1.3	33.9	24.1	6.6	6.0	17.5	18.3	13222.0	16.7
SA 4479	59.0	247.4	17.7	10.9	1.0	1.3	44.7	26.5	9.0	4.4	23.8	17.7	10222.0	18.8

Accession	DFL	PHT	SD	LPP	LS	PS	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
SA 4481	98.0	257.6	19.6	13.3	1.0	1.7	59.7	14.9	5.9	2.7	19.9	17.7	18972.0	34.1
SA 4482	66.0	157.4	13.7	7.0	1.0	1.0	17.8	25.7	4.6	0.4	21.2	17.8	2756.0	8.2
SA 4483	58.7	267.4	15.7	11.0	1.0	1.0	36.4	19.2	5.1	0.3	12.3	20.5	8444.0	17.3
SA 4484	65.7	246.7	20.3	11.0	1.0	2.0	46.1	26.0	7.0	5.0	21.8	17.7	16778.0	22.2
SA 4485	62.3	245.6	18.2	10.0	1.0	1.3	46.9	23.2	9.3	7.4	19.7	15.9	16533.0	19.3
SA 4486	73.7	265.5	15.6	10.1	1.0	2.3	32.8	28.4	5.5	6.0	25.3	9.9	9911.0	16.0
SA 4488	66.7	232.0	17.1	10.0	1.0	2.0	34.0	24.1	6.0	5.2	19.4	15.5	11667.0	14.8
SA 4489	63.3	244.4	17.9	9.3	1.0	1.0	33.5	26.9	6.7	1.5	21.8	13.9	10600.0	15.2
SA 4490	69.7	287.4	18.3	12.5	1.0	1.0	44.7	29.7	7.7	2.4	21.6	19.2	13822.0	21.6
SA 4491	68.3	240.3	14.8	10.7	2.0	1.3	25.4	18.9	6.1	3.5	15.2	14.7	9667.0	13.5
SA 4494	66.7	295.6	17.1	9.3	2.7	1.0	46.6	28.4	8.0	3.2	23.2	17.5	13067.0	24.9
SA 4495	69.7	286.8	21.9	12.6	1.3	1.7	75.5	34.7	6.1	1.5	22.6	18.3	27222.0	35.7
SA 4501	68.3	256.4	16.2	11.3	1.0	1.0	33.7	26.4	7.4	4.0	18.1	18.2	9067.0	15.7
SA 4503	65.7	265.7	14.1	10.7	1.0	1.0	33.3	31.9	6.1	2.1	22.3	6.8	2738.0	13.9
SA 4506	67.7	264.7	18.5	10.1	1.0	1.7	44.5	23.8	5.9	3.3	24.3	17.1	14000.0	19.9
SA 4510	62.0	191.2	17.0	11.6	1.0	1.0	35.1	21.0	7.4	6.0	16.3	15.0	10644.0	14.0
SA 4520	63.3	260.9	16.3	10.3	1.0	1.3	35.9	25.1	7.8	5.4	18.5	17.5	11556.0	16.9
SA 4523	66.0	174.7	26.9	10.0	1.0	1.7	25.6	29.8	4.5	3.6	23.3	15.1	2827.0	10.6
SA 4524	67.3	295.2	18.8	11.5	1.0	1.3	38.0	28.3	6.3	4.0	24.8	14.4	13720.0	19.3
SA 4525	64.7	268.0	12.4	10.5	1.0	1.0	24.6	30.8	5.7	3.3	25.0	7.2	1353.0	10.8
SA 4526	63.0	287.1	15.9	12.3	1.0	1.0	40.8	27.8	4.1	0.4	15.3	20.8	9311.0	22.1
SA 4528	69.3	257.0	16.6	11.5	1.0	1.3	34.1	25.4	6.3	4.2	19.2	14.2	13189.0	16.0
SA 4534	68.7	252.4	12.6	10.1	1.0	1.0	28.4	26.4	8.3	4.9	17.3	19.8	6276.0	14.8
Mean	66.4	239.4	16.3	10.6	1.1	1.3	34.3	25.3	6.5	3.5	20.7	16.3	9605.0	16.8
L.S.D $_{0.05}$	2.84	8.47	5.09	1.34	0.46	0.62	2.04	5.72	1.34	0.68	2.31	2.44	2206.00	2.90
%C.V.	2.63	2.18	19.25	7.81	26.77	30.10	3.66	13.94	12.60	11.97	6.86	9.24	14.15	10.64

DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS= Pest damage score (%), SY = Stalk yield (t/ha), PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix (%), JV = Juice volume, BFW = Bagasse fresh weight (t/ha)

Table 3.7   Mean squares for seven quantitative traits measured at Ukulinga, 2017

Source of Variation	DF	DFL	PHT	SD	LPP	LS	PS	SY
Rep	2	0.25	40.39	16.80	2.60	0.11	0.09	0.17
Incomplete block	12	33.27***	1627.29***	28.99***	0.97ns	0.11ns	0.43***	140.07***
Germplasm	49	109.14***	3790.89***	21.50***	4.89***	0.22***	0.28***	336.08***
Residual		3.06	27.23	9.84	0.68	0.08	0.15	1.57
Total	149	40.34	1393.98	15.31	2.11	0.13	0.21	122.72

DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS = Pest infestation score (%), SY = Stalk yield (t/ha**),** IB = Incomplete bloc

Table 3.8   Mean squares for seven quantitatve traits measured at Ukulinga, 2017

Source of Variation	PAL	PAW	PWT	GWT	BR	JV	BFW
Rep	0.93	3.62	0.02	3.63	1.78	6391529.00	16.33
Rep.IB	70.30***	7.86***	5.53***	11.73***	17.71***	36810316.00***	35.35***
Genotype	73.89***	5.73***	9.97***	39.97***	32.21***	70264837.00***	95.36***
Residual	12.44	0.68	0.18	2.02	2.25	1847286.00	3.18
Total	37.15	2.96	3.83	15.30	13.34	27223829.00	36.26

PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix, JV = Juice volume, BFW = Bagasse fresh weight (t/ha)

3.6.4   Combined analysis of quantitative traits

The ANOVA showed significant differences (p<0.001) for combined sites (Table 3.9). The number of days to 50% flowering ranged from 59 for SA 0240 to 100 for SA 4481 with a grand mean of 74. SA 4482 had the shortest plants with a height of 140.9 cm while SA 4495 had the tallest plants with a height of 328.3 cm. The overall mean for plant height was 215.8 cm. The minimum stalk diameter was 10.0 mm for SA 0311 while SA 4495 had a maximum stalk diameter of 21.0 mm. The grand mean was 17.0 mm. Number of leaves per plant ranged from 8 for SA 0311 to 14 for SA 4481. The minimum lodging score was 1.0 for almost all accessions except for SA 4494 which had the highest score of 3.0. The minimum stem borer infestation score was 2.0 while the maximum was 6.0, with an overall mean of 3.0.

A wide range of genetic diversity was revealed in stalk yield. Yields ranged from 11.5 t/ha for SA 0240 to 63.1 t/ha for SA 4495. The grand mean for stalk yield was 32.9 t/ha. The mean for panicle length was 25.7 cm, with a minimum length of 13.5 cm for SA 2193 and a maximum length of 36.4 cm for SA 2116. The mean for panicle width was 6.6 cm. The minimum panicle width was observed in SA 2193 with a value of 4.4 cm. SA 4479 had a maximum width of 8.7 cm. Significant differences were also observed in the weight of panicles. The mean was 3.7 t/ha. SA 2116 had the lowest panicle weight of 1.4 t/ha while SA 4479 had the highest value of 6.9 t/ha. Variation was also observed in 1000 grain weight. SA 4483 had the least weight of 12.2 g while SA 0163 had the highest weight of 27.2 g for 1000 grains. The mean for 1000 grain weight was 18.4 g.

Wide genetic variations were also observed for the industrial traits including juice brix, juice volume, bagasse weight and juice extraction ratio. The minimum brix value was 3.1% for SA 0240 and the maximum was 19.2% for SA 1330. The trend was extended to juice volume where the grand mean was 9500.0 l/ha, with a minimum volume of 437 l/ha for SA 0240 and a maximum of 22 000.0 l/ha for SA 4495. Fresh bagasse weight ranged from 4.7 t/ha for SA 0240 to 31.1 t/ha for SA 4495, with an overall mean of 17.0 t/ha.

Table 3.9   Means for 14 quantitative traits under combined analysis, 2017

Accession	DFL	PHT	SD	LPP	LS	PS	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
SA 0096	98.0	228.4	18.1	12.5	1.0	2.0	36.4	20.6	6.0	3.8	9.8	14.7	9213.0	20.2
SA 0163	73.3	156.1	14.7	10.1	1.0	5.3	17.9	29.5	7.5	4.5	23.4	4.5	3692.0	12.4
SA 0240	60.7	125.3	11.3	7.2	1.0	5.0	8.4	25.0	6.4	2.8	21.1	4.9	350.0	4.3
SA 0311	56.7	144.1	10.4	7.8	1.0	2.0	12.1	22.0	5.2	2.0	19.7	13.6	4210.0	7.2
SA 0312	83.0	201.1	20.3	11.4	1.0	2.7	40.4	23.0	6.8	5.0	11.1	11.7	11956.0	25.5
SA 0317	75.3	212.7	18.7	11.9	1.0	2.3	39.1	18.0	5.0	4.5	24.8	9.1	11750.0	22.1
SA 0831a	72.3	185.1	16.0	9.8	1.0	3.0	35.8	24.4	6.1	4.3	14.6	6.7	10434.0	13.6
SA 1238	63.7	163.8	12.1	9.1	1.0	6.3	9.5	34.2	7.9	3.4	20.2	6.6	1503.0	6.2
SA 1242	84.7	202.5	15.8	11.7	1.0	3.0	27.6	22.7	5.5	2.2	9.8	11.2	8367.0	15.1
SA 1330	91.0	281.3	20.3	12.6	1.0	2.7	38.2	18.8	5.5	1.5	15.9	19.2	12865.0	21.7
SA 1382	81.7	269.1	14.2	11.1	1.0	2.7	27.7	34.0	8.0	4.6	13.6	13.6	8686.0	15.7
SA 160	88.0	232.8	16.9	10.9	1.0	2.0	30.0	31.9	8.1	4.1	12.5	15.5	9259.0	18.1
SA 1904	64.0	174.3	12.3	8.2	1.0	3.0	14.5	25.4	6.2	3.2	20.2	10.9	5462.0	8.1
SA 2012	81.7	217.7	16.0	11.7	1.0	3.3	29.1	24.0	6.6	3.1	12.2	11.4	10934.0	15.1
SA 2013	98.7	169.3	13.8	12.3	1.0	2.3	22.9	21.9	5.4	4.0	9.4	13.9	6044.0	11.4
SA 2014	66.0	234.8	14.4	9.1	1.0	4.3	22.6	31.7	7.6	5.8	24.2	5.0	4958.0	14.0
SA 2034	70.7	148.5	10.8	8.7	1.0	3.3	13.3	27.3	7.7	2.4	24.0	3.1	2482.0	8.3
SA 2036	72.3	178.4	16.1	9.9	1.0	2.3	22.3	22.2	5.6	2.3	19.4	17.4	5733.0	13.0
SA 2116	85.7	240.6	15.7	10.2	1.0	1.7	23.8	36.2	9.3	1.7	17.6	10.1	2300.0	15.3
SA 2193	100.0	239.6	19.0	14.9	1.0	3.3	44.2	13.4	4.4	2.9	16.3	10.6	16133.0	24.8
SA 2218	97.3	170.3	17.9	12.4	1.0	2.3	36.2	22.4	5.6	4.0	9.8	14.0	11000.0	22.2
SA 2249	86.3	212.8	14.6	10.5	1.0	3.3	24.6	22.5	5.9	2.3	13.2	13.2	8271.0	13.1
SA 2330a	88.7	215.2	15.6	11.8	1.0	2.3	31.0	22.6	5.7	2.5	12.4	13.4	9000.0	15.3
SA 2400	80.7	189.7	18.2	11.3	1.0	2.7	57.9	25.0	7.3	3.5	12.7	13.1	18022.0	30.7
SA 4476	82.3	223.2	18.4	10.1	1.0	4.3	26.6	28.7	7.1	3.0	14.3	13.8	9287.0	16.4
SA 4477	97.7	201.4	19.8	12.9	1.0	2.0	51.9	24.0	6.4	3.1	11.7	11.7	12778.0	26.1
SA 4479	70.0	242.0	18.6	11.6	1.0	4.7	46.3	30.8	8.3	8.7	20.5	7.2	15193.0	23.7
SA 4481	101.0	256.5	19.1	13.7	1.0	2.3	37.9	15.2	4.5	1.2	12.8	9.7	12444.0	20.9

Accession	DFL	PHT	SD	LPP	LS	PS	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
SA 4482	72.7	124.4	13.3	8.0	1.0	2.0	8.5	25.1	5.6	2.4	19.8	13.5	1098.0	4.9
SA 4483	82.3	203.4	19.7	10.7	1.0	3.0	34.5	22.8	5.7	4.1	12.0	10.9	11982.0	20.5
SA 4484	76.7	208.9	17.7	10.3	1.0	3.3	30.4	29.8	6.9	3.9	14.4	9.8	10883.0	14.3
SA 4485	86.3	236.5	19.9	11.7	1.0	2.0	40.6	25.9	6.3	3.6	14.6	11.9	14318.0	24.3
SA 4486	78.3	253.3	16.0	10.6	1.0	4.3	30.9	30.7	7.6	3.4	14.7	7.3	10887.0	16.0
SA 4488	70.3	207.7	16.5	10.5	1.0	3.7	30.3	24.7	6.3	4.0	13.9	13.1	10102.0	14.6
SA 4489	75.0	191.1	17.7	9.4	1.0	5.7	30.4	27.3	7.3	4.3	19.3	8.0	11167.0	15.8
SA 4490	90.3	284.1	21.4	14.1	1.0	3.3	54.8	31.2	8.0	2.9	14.5	10.6	20178.0	28.0
SA 4491	90.0	207.4	15.7	12.0	1.0	2.3	30.1	20.9	5.3	3.6	9.3	11.0	10222.0	15.1
SA 4494	68.7	291.6	19.1	9.4	4.0	6.0	37.8	30.4	8.7	3.6	20.6	7.1	11667.0	21.5
SA 4495	93.7	328.3	19.7	14.2	1.0	4.0	50.7	33.4	8.4	2.9	16.8	11.1	16778.0	26.4
SA 4501	89.7	244.7	20.3	12.5	1.0	2.7	40.2	26.6	7.2	3.4	12.6	10.2	13111.0	20.5
SA 4503	69.7	272.1	13.3	10.7	1.0	3.0	17.7	33.9	7.9	6.3	26.3	7.2	1804.0	13.7
SA 4506	79.3	248.6	18.7	11.5	1.0	2.7	38.5	25.3	6.4	2.9	16.3	13.4	12333.0	22.1
SA 4510	98.7	181.5	16.7	12.5	1.0	2.3	32.3	21.7	5.5	3.4	9.4	14.5	8491.0	14.6
SA 4520	92.0	223.2	17.2	12.2	1.0	2.0	33.9	25.5	6.6	3.5	12.6	14.4	13156.0	17.2
SA 4523	69.3	145.4	14.1	9.7	1.0	6.3	15.3	23.0	6.7	2.1	23.1	3.8	2440.0	8.0
SA 4524	72.0	265.0	20.8	12.4	1.0	4.0	51.9	32.0	8.4	5.1	19.6	9.3	12178.0	30.0
SA 4525	70.7	246.6	14.4	10.7	1.0	3.0	21.1	35.6	8.7	6.5	27.4	7.3	2489.0	13.8
SA 4526	82.7	253.0	19.3	11.9	1.0	2.3	45.8	27.4	7.1	4.1	16.9	17.1	11311.0	23.5
SA 4528	77.7	215.9	16.1	10.7	1.0	3.3	34.6	26.4	6.9	4.4	13.5	11.1	10400.0	14.2
SA 4534	84.3	242.9	19.1	11.9	1.0	2.0	42.8	31.9	8.1	3.9	12.2	13.5	10378.0	23.9
Mean	80.8	215.8	16.7	11.1	1.1	3.2	31.6	26.2	6.7	3.9	16.1	10.9	9500	17.3
L.S.D $_{0.05}$	2.23	5.54	2.83	0.97	0.23	0.76	2.34	3.12	0.75	0.37	1.65	1.69	1727.00	2.17
%C.V.	2.66	2.14	15.07	7.91	18.92	29.73	6.24	10.63	9.92	8.82	7.88	10.91	15.97	11.21

DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS= Pest damage score (%), SY = Stalk yield (t/ha), PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix (%), JV = Juice volume, BFW = Bagasse fresh weight (t/ha)

Table 3.10 Mean squares for seven quantitative traits under combined analysis, 2017

Source of Variation	DF	DFL	PHT	SD	LPP	LS	PS	SY
Rep	2	14.1	69.7	58.4	6.3	0.1	0.4	9.9
Rep.IB	12	19.8***	2335.8***	13.1**	1.3*	0.2***	2.5***	125.0***
Germplasm	49	320.8***	8760.7***	35.7***	11.7***	0.6***	2.4***	672.6***
Site	1	15566.4***	41809.3***	13.4ns	19.3***	0.0ns	280.3***	519.6***
Germplasm × Site	49	173.6***	747.0***	11.8***	1.5***	0.1***	2.0***	121.5***
Residual	186	3.8	23.7	6.2	0.7	0.0	0.4	4.2
Total	299	136.4	1806.9	12.6	2.8	0.2	2.0	139.6

DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS = Pest infestation score (%), SY = Stalk yield (t/ha), IB = Incomplete bloc

Table 3.11 Mean square for seven quantitative traits under combined analysis, 2017

Source of Variation	DF	PAL	PAW	PWT	GWT	BR	JV	BFW
Rep	2	2.3	3.8	0.1	4.6	2.9	15948918.0	26.9
Rep.IB	12	59.2***	4.8***	2.5***	12.6***	20.9***	30024593.0***	25.5***
Genotype	49	146.1***	6.6***	5.2***	89.8***	53.9***	118671546.0***	186.9***
Site	1	57.9***	3.1***	10.6***	1582.2***	2138.4***	3346112.0ns	18.5*
Germplasm × Site	49	11.1*	4.0***	5.5***	20.9***	16.4***	17372001.0***	37.0***
Residual	186	7.5	0.4	0.1	2.1	2.2	2300215.0	3.6
Total	299	33.0	2.2	2.0	25.3	20.9	25048538.0	40.2

PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix, JV = Juice volume, BFW = Bagasse fresh weight (t/ha)

3.6.5 Principal component analysis

The first four principal components (PC1, PC2, PC3 and PC4), which had an eigenvalue greater than 1.00 accounted for 79.12% of the total variation for all the traits studied. These were extracted and presented in Table 3.12. The first principal component (PC) accounted for 42.02% of the total variation and had an eigenvalue of 5.89. Traits which contributed to this variation were days to 50% flowering, plant height, stalk diameter, number of leaves per plant, stalk yield, %brix, juice volume and fresh bagasse weight. The second PC had the eigenvalue of 2.68 and accounted for 19% of the total variation. Traits which contributed to this variation were pest damage, panicle length, panicle width and panicle weight. The third PC contributed 10% of the total variation and had an eigenvalue of 1.44. 1000 grain weight was the only trait which contributed to this variation. Finally, the fourth PC accounted for 7.63% of the total variation and had an eigenvalue of 1.07 with lodging score as the main contributor to this variation.

Table 3.12 Principal component analysis for 14 quantitative traits indicating Eigenvectors, Eigenvalues and proportion of variation explained with the first four PC axes across two sites

Trait	Eigenvector			
	PC1	PC2	PC3	PC4
DFL	0.76	-0.32	-0.03	-0.11
PHT	0.68	0.47	0.01	0.26
SD	0.81	0.20	0.24	-0.23
LPP	0.87	-0.03	0.00	-0.25
LS	0.07	0.40	0.28	0.61
PS	-0.26	0.61	0.54	-0.31
SY	0.94	0.23	0.11	-0.03
PAL	-0.31	0.73	-0.28	0.18
PAW	0.14	0.71	-0.57	0.14
PWT	0.05	0.43	-0.58	-0.46
GWT	-0.58	0.44	0.46	-0.09
BR	0.57	-0.50	-0.16	0.35
JV	0.91	0.17	0.08	-0.06
BFW	0.93	0.21	0.16	0.09
Eigenvalue	5.89	2.68	1.44	1.07
Individual %	42.08	19.12	10.28	7.63
Cumulative %	42.08	61.20	71.48	79.12

PC= Principle component, DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS= Pest infestation score (%), SY = Stalk yield (t/ha), PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix (%), JV = Juice volume, BFW = Bagasse fresh weight (t/ha).

3.6.6   Principal component biplot

The observed phenotypic diversity among sweet sorghum population under study is presented in a principal biplot in Figure 3.1. Narrow angles between dimension vectors in the same direction indicated positive and significant correlations of the variable traits in terms of discriminating the accessions. From Figure 3.1, stem diameter and juice volume were positively and significantly correlated. Accessions that performed better in a specific trait were plotted closer to the vector line and further in the direction of that particular vector, often on the vertices of the convex hull of the first principal component.
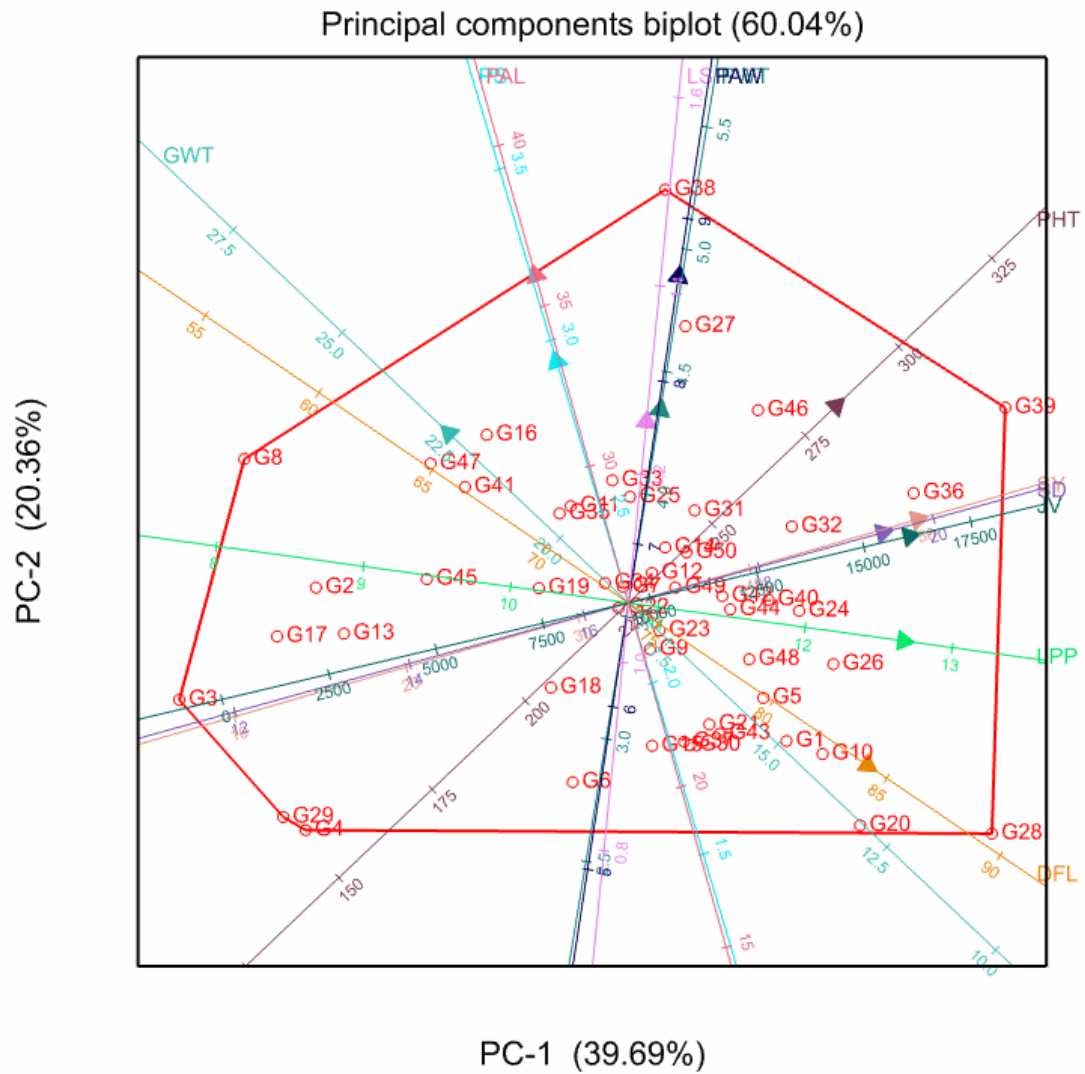
Figure 3-1 Principal component analysis biplot descrbing the overall variation among sweet sorghum accessions estimated using 14 quantitative traits across two sites.

3.7 Discussion

3.7.1 Genetic variation for the traits under combined analysis

Characterisation of germplasm reveals the level of genetic diversity and relatedness which exists amongst collections, and progress in a breeding programme depends largely on the extent of genetic diversity existing in the population (Zou et al., 2011).

Combined analysis of variance for the two environments showed highly significant G x E interaction for almost all traits. Days to 50% flowering for the combined analysis varied from 59 days to 100 days and these findings are within the range of flowering days reported by Lekgari and Dweikat (2014). The number of days to flowering is an important trait for categorising sweet sorghum into maturity ranges. Kudadjie et al. (2007) grouped sorghum into early maturing (less than 85 days flowering days), medium maturing (86-105 days) and late maturing (more than 105 days). Therefore, this study identified 48 accessions as early maturing and two accessions (SA 2193 and SA 4481) as medium maturing. As reported by Kisua et al. (2015), early maturity is an important trait which can result in drought escape. Therefore, the accessions identified in this study can be a good source for gene mining when breeding for earliness.

Plant height ranged from 140.9 cm for SA4482 to 328.3 cm for SA4495 with a grand mean of 215.8 cm. Similar variations were reported in previous studies by Ericksom et al. (2012) and Shukla et al. (2016). Plant height is an important trait, which is positively and significantly correlated to stalk yield and ultimately juice volume (Audilakshmi et al., 2010). However, taller plants are undesirable when selecting or breeding against lodging because most of them are susceptible to lodging (Murray et al., 2009). On the other hand, stalk thickness is an important trait when selecting against lodging because it compensates for height. Stalk diameter and plant height are also important traits which directly influence stalk yield and juice volume (Audilakshmi et al., 2010). Therefore, tall and thick stemmed genotypes must be selected for high yields and for resistance to lodging.

The mean stem borer score for the combined analysis was 6.0. Davis and Williams (1992), summarised pest infestation based on a scale of 0-9 as follows: 0-4 low infestation, 5-7 medium and 8-9 high infestation. This means there was medium pest infestation in the germplasm characterised. Stem borer damage reduces juice yield and purity (Long and Hensley, 1972). The authors also reported a sucrose yield reduction of 10 to 20% due to pest

damage. During larvae stage, the stalk borers tunnel through the stalk, causing plant breakage, stunting and ultimately low stalk yield (Rodriguez-del-Bosque et al., 1990).

The results also showed a wide variation in %brix amongst accessions. The values ranged from 3.1 to 19.2%. These figures are slightly closer to those reported by Reddy et al. (2005). The authors reported a minimum brix of 16.0% and a maximum of 23%. In another study, Woods (2000) reported brix values ranging from 11.0% to 18.5%. Significant phenotypic differences for brix were also reported by Wang et al. (2009). Brix is an important industrial trait, which makes sweet sorghum a useful source for bioethanol production (Calvino et al., 2008). Therefore, brix can be an important trait to consider for selection. However, brix is an additive trait as reported by Felderhoff et al. (2012), meaning that breeding for high sugar concentration in sweet sorghum hybrids requires both parents to have high brix.

Combined analysis of variance on juice volume revealed a wide genetic diversity amongst the accessions. Juice volume ranged from 437 l/ha for SA0240 to 22000 l/ha for SA4495 with an overall mean of 9500 l/ha. These findings were also consistent with Burks et al. (2015). Juice volume is an important trait for selection because it is positively correlated to sugar yield (Murray et al., 2008).

Fresh bagasse weight ranged from 4.7 t/ha for SA0240 to 31.1 t/ha for SA4495, with an overall mean of 17 t/ha. These results are similar to those reported by Holou and Stevens (2012). Sweet sorghum bagasse can be used as livestock feed and to generate power for the plant. The bagasse from sweet sorghum has a higher biological value than the bagasse from sugar cane when used as fodder for animals because it has high content micronutrients and minerals (Seetharama et al., 2002).

3.7.2   Genotype by environment (G x E) interaction

Results from the combined analysis of variance showed that G x E interaction was significant for all the traits studied. However, there was no crossover on the ranking of accessions between locations (graph not provided). In a G x E study by Holou and Stevens (2012) higeher juice and %brix yields were observed in loamy clay soils than in clay soils. In this study however, higher juice and %brix yields were observed at Ukulinga (clay soils) than at Potchefstroom (loamy clay soils). This could be because of stem borer damage which had a negative correlation to both juice and %brix yield at Potchefstroom. Significant G x E interactions for traits means that these characters cannot be selected simultaneously across environments. Selection for each location, therefore, must be carried out independently. This

complicates the full utilisation of genotype means across environments for selecting and advancing superior genotypes, as reported by Pham and Kang (1988).

3.7.3 Principal component analysis

In the principal component analysis components with eigenvalues less than 1.00 were eliminated as explained by Chatfield and Collins (1980). According to Hair et al. (1998), component loadings of ±0.03 are considered meaningful, therefore, they were picked. The eigenvalues decreased significantly from first component (5.89) to second component (2.68), but the values decreased non-significantly from third component (1.44) to the fourth component (1.07). A similar trend of values was reported by Kisua et al. (2015). Traits which contributed more to the first component accounted for 42.08% of the total variation and these were days to 50% flowering, plant height, stalk diameter, number of leaves per plant, %brix and juice volume. This implies that selection in the first component should focus on these traits. The first and second component, which had a cumulative variation of 61.2%, displayed most variation among the accessions indicating a high degree of association among the traits studied (Kisua et al., 2015).

High phenotypic variability observed among the sweet sorghum germplasm was also confirmed by the PCA biplot display. Small angles between dimension vectors in the same direction indicate significant correlation of the trait in discriminating accessions, as previously suggested by Mwadzingeni et al. (2016). Germplasm performing better for a specific trait were plotted closer to the vector line and further in the direction of that vector, usually on the vertices of the convex hull of the first two principal components. For instance, stem diameter and juice yield were positively and significantly correlated. This means that selection for high juice yield can be achieved through selecting for stem thickness.

3.8 Conclusion

The results revealed a high level of genetic diversity among the sweet sorghum germplasm studied. There were positive and significant correlations between plant height and biomass yield, biomass yield and juice volume and bagasse, suggesting that improvement for these traits can be carried out simultaneously. The G x E interaction was significant for almost all the traits studied, implying that selection for these traits cannot be done across sites. Overall, the study identified SA4490, SA2400, SA4495, SA2193 and SA4479 as superior accessions for juice yield.

References

Ali, M.L., J.F. Rajewski, P.S. Baenziger, K.S. Gill, K.M. Eskridge and A. Dweikat. 2008. Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Molecular Breeding 21: 497-509.

Atokple, I.D.K. 2003. Sorghum and Millet breeding in West Africa in Practice. Proceedings of AFRIPRO. European Union Research Framework International workshop on Sorghum and Millet Proteins, Intundla Game Lodge, South Africa. 2- 4 April 2003

Audilakshmi, S., A.K. Mall, M. Swarnalatha and N. Seetharama. 2010. Inheritance of sugar concentration in stalk (Brix), sucrose content, stalk and juice yield in sorghum. Biomass Bioenergy 34: 813-820.

Brinkman, H.-J., S. de Pee, I. Sanogo, L. Subran and M.W. Bloem. 2009. High food prices and the global financial crisis have reduced access to nutritious food and worsened nutritional status and health. The Journal of Nutrition 140: 153S-161S.

Burks, P.S., C.M. Kaiser, E.M. Hawkins and B. P.J. 2015. Genomewide association for sugar yield in sweet sorghum. Crop Science 55: 2138-2148.

Calvino, M., R. Bruggmann and M. J. 2008. Screen of genes linked to high sugar content in stems by comparative genomics. Rice 1: 166-176.

Chatfield, C. and A. Collins. 1980. Introduction to multivariate analysis. Chapman and Hall.

Cochran, W.G. and G.M. Cox. 1992. Experimental Designs. second edition ed. Wiley Interscience. Canada.

Davis, F.M., and W.P. Williams. 1992. Visual rating scales for screening whorl-stage corn for resistance to fall armyworm. Technical bulletin-Mississippi Agricultural and Forestry Experiment Station.

Ericksom, J.E., K.R. Woodhard and L.E. Sollenberger. 2012. Optimizing sweet sorghum production for biofuelin the southeastern USA through nitrogen fertilisation and top removal. Bioenergy Research 5: 86-94.

Felderhoff, T.J., S.C. Murray, P.E. Klein, A. Sharma, M.T. Hamblin, S. Kresovich, W. Vermerris and W.L. Rooney. 2012. QTLs for energy-related traits in a sweet × grain sorghum mapping population. Crop Science 52: 20-40.

Franco, J., J. Crossa, J.M. Ribaout, J. Betran, M.L. Warburton and M. Khairallah. 2001. A method for combining molecular markers and phenotype attributes for classifying plant genotypes. Theoretical and Applied Genetics 103: 944-952.

Hair, J., R. Andreson, R. Tatham and W. Black. 1998. Multivariate Data Analysis. 5th ed. Prentice-Hall Inc, Unites States of America.

Holou, R.A.Y. and G. Stevens. 2012. Juice, sugar and bagasse response of sweet sorghu (*Sorghum bicolor* (L.) Moench cv. M81E) to N fertilisation and soil type. GCB Bioenergy 4: 302-310.

IBPGR, I. 1993. Descriptors for sorghum [*Sorghum bicolor* (L.) Moench] International Board for Plant Genetic Resources, Rome, Italy.

Jessup, R. 2009 Development and status of dedicated energy crops in the United States. In Vitro Cellular & Developmental Biology 45: 282-290.

Kisua J., Mwikamba K., M. M. and M. A. 2015. Genetic diversity of sweet and grain sorghum using phenotypic markers. International Journal of Biosciences 6: 34-46.

Kudadjie, C., p. Struik, P. Richards, S. Offei and P. Atengdem. 2007. Understanding variation in sorghum through with-farmer experimentation. International Journal of Agricultural Sustainability 5: 124-139.

Lekgari, A. and I. Dweikat. 2014. Assessment of Genetic Variability of 142 Sweet Sorghum Germplasm of Diverse Origin with Molecular and Morphological Markers. Open Journal of Ecology 4: 371-393.

.Long, W.H. and S.D. Hensley. 1972. Insect pests of sugar cane. Annual Review of Entomology 17: 149-176.

Lynch, M. and B. Walsh. 1998 Genetics and Analysis of Quantitative Traits. Sinauer Associates.

Murray, S.C., A. Sharma, W.L. Rooney, P.E. Klein, J.E. Mullet, S.E. Mitchell. 2008. Genetic improvement of sorghum as a biofuel feedstock I: QTL for stem sugar and grain nonstructural carbohdrates. Crop Science 48: 2165-2179.

Murray, S.C., W.L. Rooney, M.T. Hamblin, S.E. Mitchell and S. Kresovich. 2009 Sweet Sorghum Genetic Diversity and Association Mapping for Brix and Height. The Plant Genome 2: 48-62.

Mwadzingeni, L., H. Shimelis, S. Tesfay and T.J. Tsilo. 2016. Screening of Bread Wheat Genotypes for Drought Tolerance Using Phenotypic and Proline Analyses. Frontiers in Plant Science 7: 1.12.

Pham, N.K. and M.S. Kang. 1988. Simultaneous selection for high yielding and stable crop genotypes. Agrononmy Journal 83: 161-165.

Reddy, B.V., S. Ramesh, P.S. Reddy, B. Ramaiah, M. Salimath and R. Kachapur. 2005. Sweet sorghum as potential alternate raw material for bio-ethanol and bio-energy. International Sorghum and Millets Newsletter 46: 79-86.

Reddy, B.V.S., S. Ramesh, S.P. Reddy and A.A. Kumar. 2009 Genetic enhancement for drought tolerance in sorghum. Plant Breeding Reviews 31: 189-222.

Rodriguez-del-Bosque, L.A., J.W. Smith and H.W. Browning. 1990. Feeding and pupation sites of *Diatraea lineolata*, *D. saccharalis*, and *Eoreuma loftini* (Lepidoptera: Pyralidae) in relation to corn phenology. Journal of Economic Entomology 83: 850-855.

Seetharama, N., D.B. Rao, C.V. Ratnavathi, S.M. Parwez, B. Mathew, K. Singh and B. Singh. 2002. Sweet sorghum-an ancillary sugar crop. Indian Farming 36: 7-8.

Semagn, K., C. Magorokosho, V. Ogugo, D. Makumbi and M.L. Warburton. 2010. Genetic relationships and structure among open-pollinated maize varieties adapted to eastern and southern Africa using microsatellite markers. Molecular Breeding 34: 1423-1435.

Sergio, L., B. Gianni. 2005. Molecular markers based analysis for crop germplasm preservation. FAO Meeting on the role of biotechnology for the characterisation and conservation of crops, forestry, animal and fishery genetic resources, Turin, Italy 5-7 March 2005.

Serna-Saldívar, S.O., C. Chuck-Hernández, E. Pérez-Carrillo and E. Heredia-Olea. 2012. Sorghum as a multifunctional crop for the production of fuel ethanol: Current status and future trends. Bioethanol. InTech. Open Science.

Shukla, S., A. Bhargava, A. Chatterjee, A. Srivastava and S.P. Singh. 2006. Genotypic variability in vegetableamaranth (*Amaranthus tricolor* L.) for foliage yield and its contributing traits over successive cuttings and years. Euphytica 151 : 103-110.

Shukla, S., T.J. Felderhoff, A. Saballos and W. Vermerris. 2016. The relationship between plant height and sugar accumulation in the stems of sweet sorghum (*Sorghum bicolor* (L) Moench). Field Crops Research 203: 181-191.

Singh, B. 2001. Plant breeding: Principles and Methods. Kalyani Pulisher, New Delhi.

Upadhyaya, H., C. Gowda and D. Sastry. 2008. Plant genetic resources management: collection, characterisation, conservation and utilization. Journal of Semi-Arid Tropics Agricultural Research 6: 1-16.

Wang, M.L., C. Zhu, N.A. Barkley, Z. Chen, J.E. Erpelding, S.C. Murray. 2009. Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection.Theoretical and Applied Genetics 120: 13-23.

Woods, J. 2000. Integrating Sweet Sorghum and Sugar cane for Bioenergy: Modelling The Potential for Electricity and Ethanol Production in SE Zimbabwe. University of London, King's College London.

Zou, G., S. Yan, G. Zhai, Z. Zhang, J. Zou and Y. Tao. 2011. Genetic variability and correlation of stalk yield-related and sugar concentration of stalk juice in a sweet sorghum (*Sorghum bicolor L. Moench*) population. Australian Journal of Crop Science 5: 1232-1238.

CHAPTER 4

Genetic and path coefficient analysis of juice yield and juice related traits in sweet sorghum

Abstract

Sweet sorghum is a potential crop for biofuel production due to its versatility, non-negative footprints on the environment and low production input requirement. For progress in breeding it is necessary to study genetic variability in the existing. Information on the correlation of traits and the extent to which desirable traits are heritable is important in sweet sorghum improvement programes. The objectives of this study were; i) to estimate variance components and heritability, and ii) to determine correlations and path analysis of juice yield and juice yield-related traits. Fourteen quantitative traits were studied on 50 sweet sorghum accessions. Variance components in the accessions were estimated using mean range, coefficient of variation, standard error, genotypic and phenotypic variance. Simple Pearson correlation coefficients were used to examine the extent and type of association between two quantitative traits. The correlations were further partitioned into direct and indirect effects. Higher phenotypic coefficient of variation (PCV) values were observed for all the traits studied compared to the genotypic coefficient of variation (GCV) values, suggesting influence of environmental effects on these traits. The estimated value of expected genetic advance expressed as percentage of the mean (GAM) at 5% proportion selected ranged from 25.7% to 131.2% for number of days to flowering and juice volume, respectively, across environments. This means that these traits are simply inherited and most likely the heritability is due to additive gene effects which means selection may be effective. Significant and positive correlations were observed between some morphological traits (number of days to flowering, plant height, stalk diameter and stalk yield) and industrial traits (%brix, juice volume and bagasse weight). However, path analysis of some of these traits did not show positive and direct contribution on juice yield. The study thus identified plant height, stalk diameter and number of leaves as primary traits with significant application for sweet sorghum improvement programme because they showed direct effects on juice yield. Therefore, direct and indirect selection for these traits would be effective as these are the best indices for sweet sorghum improvement programme.

Key words: correlation, genetic advance, heritability, juice yield, path analysis, selection.

4.1 Introduction

Sweet sorghum is a possible alternative raw material for ethanol production. The juice extracted from sweet sorghum cane is rich in sucrose and invert sugar that are fermented to produce biofuel (Prasad et al., 2007). According to Ali et al. (2008) and Murray et al. (2009) juice yield in sweet sorghum is one of the most useful morphological traits for selection apart from plant height, stalk diameter and brix. These traits, however, are quantitative and polygenically inherited in nature and thus complex to be manipulated directly in breeding procedure because of environmental noise (Ali et al., 2008). Therefore, to effectively improve these complex traits, there is a need to separate them into smaller morphological, physiological and yield components, which could be easily analysed and evaluated through correlation coefficient and path analysis.

Information on correlation (direction and magnitude) of traits is very useful in breeding programmes for further understanding of the genetic mechanism of stalk yield-related traits and sugar concentration of stalk juice (Kisua et al., 2015). This is critical in determining whether the traits are genetically controlled and heritable and thus transmitted from parent to offspring of the desired genotypes. In sweet sorghum, therefore, it is necessary to understand correlation of both agronomical traits such as plant height, stalk diameter, stalk yield, panicle mass and industrial traits such as juice extraction (%), reducing sugar content, stalk fibre ratio, sucrose percentage in a stalk and total soluble solute percentage (%brix).

Knowledge of the correlation between yield and its contributing traits and among the component traits themselves can improve the efficiency of selection in plant breeding (Lzge et al., 2006). Studies on correlations lead to better efficiency and reliability when combined with other analyses such as path analysis (Vendruscolo et al., 2016). As proposed by Wright (1921), the path analysis identifies direct and indirect variables that are significantly correlated with the basic variable in order to attain an efficient response. It enables partitioning of correlation coefficient into direct and indirect effects of several traits towards dependent variable and thus helps in evaluating the cause-effect association as well as effective selection (Dewey and Lu, 1959). In addition, path analysis is critical for better understanding of correlations among traits, which is a pathway for knowledge on specificity of the genetic material being studied (Dewey and Lu, 1959).

On the other hand, analysis of variability among the traits and the correlation of a trait in relation to other traits contributing to yield of a crop would be significant in planning a successful breeding programme (Mary and Gopalan, 2006). Development of high-yielding cultivars needs a comprehensive understanding of the genetic variation present for juice yield

and its components. The observed variability is a collective estimate of genetic and environmental causes of which only the former is inherited. However, estimates of heritability alone do not provide information about the expected gain in the succeeding generation, but must be considered together with estimates of genetic advance, that is, the change in mean value among successive generations (Shukla et al., 2006; Al-Tabbal and Al-Fraihat, 2011). The estimates of genetic advance (GA) help in understanding the type of gene action involved in the expression of various polygenic traits. High values of GA indicate additive gene action whereas low values are evident of non-additive gene action. This, therefore, implies that the heritability estimates will be useful if accompanied by high GA (Singh and Narayanan, 1993).

This study, therefore, aimed at analysing and determining traits having significant interrelationship with juice yield by using the correlation and path analysis, and to estimate the genotypic, non-genotypic variance components leading to indirect selection of yield in sweet sorghum.

4.2 Hypothesis

The hypothesis tested for this study was that there is no significant relationship between juice volume and other traits in sweet sorghum.

4.3 Materials and methods

Fifty sweet sorghum germplasm listed in chapter 3 (Table 3.1) were used in this study.

4.4 Data collection

Quantitative characters were measured based on sorghum descriptors (IBPGR, 1993) presented in Chapter 3, Section 3.3 under data collection.

4.5 Statistical analysis

4.5.1 Analysis of variance (ANOVA)

Initially, combined analysis of variance for the quantitative data for the two sites was performed in GenStat statistical package, 18th edition, using the unbalanced treatment structural design. This was aimed at determining the genetic diversity of the measured quantitative traits. Descriptive statistics (mean value, coefficient of variation (%CV), least significant difference at 5% level ($LSD_{0.05}$) were used to compare levels of variation for morphological and agronomic traits both within and between two sites.

4.5.2   Estimation of genetic and phenotypic variance components

Variability in the accessions was estimated using mean, range, coefficient of variation, standard error, genotypic and non-genetic (phenotypic) variance components, as previously done by Singh and Chaudhary (1979). Estimations were done on each individual trait as follows:

Genotypic variance, $\sigma_g = \dfrac{[MSt - MSe]}{r}$

Phenotypic variance, $\sigma_p^2 = \sigma_g^2 + \sigma_e^2$

Where;

$\sigma_g^2 =$ Genotypic variance

$\sigma_p^2 =$ Phenotypic variance

$\sigma_e^2 =$ Environmental variance

$MSt =$ Treatment Mean Square

$MSe =$ Residual Mean Square

$r =$ Number of replications

Genotypic coefficient of variation, $(GCV) = \dfrac{\sqrt{\sigma_g^2}}{mean} \times 100$

Phenotypic coefficient of variation, $(PCV) = \dfrac{\sqrt{\sigma_p^2}}{mean} \times 100$

Broad sense heritability $\left(H^2\right)$

Broad sense heritability, H^2 was estimated using the formula, previously suggested by Hanson et al. (1956), as follows:

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2} \times 100$$

H^2 = Broad sense heritability

σ_g^2 = Genotypic variance

σ_p^2 = Phenotypic variance

The expected Genetic Advance (GA) for the traits was calculated as follows

$$GA = K \times \left(\sqrt{\sigma_p^2}\right) \times H^2$$

Where

$K = 2.06$, at 5% selection intensity

σ_p^2 = Phenotypic variance for the trait

H^2 = Broad sense heritability of the trait

Genetic Advance as a percentage of the mean was calculated as

$$GAM = \frac{GA}{mean} \times 100$$

GA = Genetic Advance

4.5.3 Correlation coefficient and path analysis

Simple Pearson correlation coefficients were used to examine the extent and type of association between two quantitative traits using IBM SPSS Statistics 25 software (SPSS., 2006). The correlations were further partitioned into direct and indirect effects as previously done by Dewey and Lu (1959).

4.6 Results

4.6.1 Genotypic and phenotypic variability and heritability estimates

The results showed significant genotypic and phenotypic differences among the accessions (Table 4.1). In all traits, a large portion of the phenotypic variance was accounted for by the genetic component. For all the traits studied, genotypic variance (σ^2_g) and phenotypic variances (σ^2_p) were higher than environmental variances (σ^2_e). Vg values ranged from 0.20 for days to 50% flowering to 38 790 443.67 for juice volume. σ^2_g ranged from 0.24 for days to 50% flowering to 41 090 658.67 for juice volume, while σ^2_e ranged from 0.04 for days to 50% flowering to 2 300 215.00 for juice volume. The estimates for phenotypic coefficient of variation (PCV) were higher than those of genetic coefficient of variation (GCV) for all the traits. PCV values ranged from 12.95% for days to 50% flowering to 67.48% for juice volume. GCV values, ranged from 12.72% to 65.56% for days to 50% flowering and juice volume, respectively.

Table 4.1 Estimates of variance components of GCV and PCV for 14 quantitative traits across two sites

Trait	Mean	MSt	Error MS	δ^2g	δ^2e	δ^2p	GCV	PCV
DFL	80.8	320.8	3.8	105.6	3.8	109.5	12.7	12.9
PHT	215.8	8760.7	23.7	2912.3	23.7	2936.0	25.0	25.1
SD	16.7	35.7	6.2	9.8	6.2	16.0	18.7	23.9
LPP	11.1	11.7	0.7	3.7	0.7	4.4	17.3	19.0
LS	1.1	0.6	0.0	0.2	0.0	0.2	42.1	46.2
PS	3.2	2.4	0.4	0.6	0.4	1.1	25.2	32.6
SY	31.6	672.6	4.2	222.8	4.2	227.0	47.2	47.6
PAL	26.2	146.0	7.5	46.2	7.5	53.7	26.0	28.0
PAW	6.7	6.6	0.4	2.0	0.4	2.5	21.2	23.4
PWT	3.9	5.2	0.1	1.7	0.1	1.8	33.7	34.7
GWT	16.1	89.8	2.1	29.2	2.1	31.3	33.5	34.7
BR	10.9	53.9	2.2	17.2	2.2	19.4	38.0	40.4
JV	9500.0	118671546.0	2300215.0	38790443.7	2300215.0	41090658.7	65.6	67.5
BFW	17.3	186.9	3.6	61.1	3.6	64.7	45.3	46.6

MS_t = Means squares for treatments, $\delta2g$ = Genotypic variance (Vg), $\delta2p$ = Phenotypic variance (Vp), GCV = Genotypic coefficient of variation, PCV = Phenotypic coefficient of variation

4.6.2 Estimates of heritability and expected genetic advance (GA)

High broad sense heritability (H^2) was observed in all the traits under study (Table 4.2). The estimates ranged from 59.5% for pest infestation score to 99.2% for plant height. The expected genetic advance expressed as a percentage of mean (GAM) ranged from 25.7% for days to 50% flowering to 131.2 for juice volume.

Table 4.2 Estimate of broad sense heritability, GA and GAM for 14 quantitative traits across two environments

Trait	GCV	PCV	H^2	GA	GAM (%)
DFL	12.7	12.9	96.5	20.8	25.7
PHT	25	25.1	99.2	110.7	51.3
SD	18.7	23.9	61.4	5.1	30.2
LPP	17.3	19	83.4	3.6	32.6
LS	42.1	46.2	83.1	0.8	79.1
PS	25.2	32.6	59.5	1.3	40
SY	47.2	47.6	98.1	30.5	96.3
PAL	26	28	86.1	13	49.6
PAW	21.2	23.4	82.5	2.7	39.8
PWT	33.7	34.7	94.1	2.6	67.3
GWT	33.5	34.7	93.3	10.8	66.7
BR	38	40.4	88.7	8.1	73.8
JV	65.6	67.5	94.4	12465.8	131.2
BFW	45.3	46.6	94.4	15.6	90.6

GCV = Genetic coefficient of variation, PCV = Phenotypic coefficient of variation, H^2 = Broad sense heritability, GA = Genetic advance, GAM = Genetic advance as percentage of mean

4.6.3 Correlation coefficients for quantitative traits across sites

There was a high and positive correlation amongst some traits in the germplasm under study, as presented in (Table 4.3). Positive and significant associations were observed for days to

50% flowering, plant height, stalk diameter, number of leaves, stalk yield, brix, juice volume and bagasse weight. Negative and significant associations were observed for 50% days to flowering and 1000 grain weight, number of leaves and 1000 grain weight, pest infestation and %brix, stalk yield and 1000 grain weight, and panicle length and %brix.

Positive moderate associations were observed for plant height and panicle weight, stem diameter and %brix, number of leaves and %brix. Negative moderate associations were observed on stalk diameter and 1000 grain weight, number of leaves and panicle length. Positive weak associations were observed on plant height and number of leaves, plant height and %brix, stalk diameter and number of leaves, number of leaves and panicle weight, panicle weight and juice volume. Weak negative associations were observed on days to 50%flowering and panicle weight, plant height and 1000 grain weight, stalk yield and panicle length, panicle weight and fresh bagasse weight.

Table 4.3   Phenotypic correlation coefficient for 14 traits of sweet sorghum germplasm among various pairs across two sites

	DFL	PHT	SD	LPP	LS	PS	SY	PAL	PAW	PWT	GWT	BR	JV	BFW
DFL	1													
PHT	0.33*	1												
SD	0.53**	0.49**	1											
LPP	0.75**	0.59**	0.67**	1										
LS	-0.08	0.26	0.09	-0.12	1									
PS	-0.39**	-0.04	0.15	-0.23	0.26	1								
SY	0.58**	0.72**	0.84**	0.80**	0.13	-0.07	1							
PAL	-0.42**	0.23	-0.2	-0.29*	0.09	0.30*	-0.17	1						
PAW	-0.09	0.36*	0.13	0.04	0.22	0.07	0.23	0.62**	1					
PWT	-0.09	0.09	0.09	0.09	-0.03	0.12	0.09	0.18	0.49**	1				
GWT	-0.59**	-0.13	-0.31*	-0.44**	0.07	0.54**	-0.37**	0.39**	-0.05	-0.1	1			
BR	0.39**	0.21	0.32*	0.34*	-0.05	-0.55**	0.41**	-0.43**	-0.15	-0.19	-0.62**	1		
JV	0.56**	0.63**	0.80**	0.73**	0.12	-0.04	0.94**	-0.24	0.22	0.12	-0.43**	0.46**	1	
BFW	0.60**	0.77**	0.78**	0.79**	0.16	-0.11	0.96**	-0.13	0.2	-0.04	-0.32*	0.42**	0.86**	1

* Correlation is significant at the 0.05 level (2-tailed), * *Correlation is significant at 0.01 level (2-tailed)

DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS= Pest infestation score (%), SY = Stalk yield (t/ha), PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix (%), JV = Juice volume, BFW = Bagasse fresh weight (t/ha).

4.6.4   Path coefficient analysis

Path-coefficient analysis (Table 4.4) was studied at phenotypic level considering juice yield as the dependent trait. Independent traits were days to 50% flowering, plant height (cm), stalk diameter (mm), number of leaves per plant, lodging score (%), pest infestation score (%), stalk yield (t/ha), panicle length (cm), panicle width (cm), panicle weight (t/ha), 1000 grain weight (g), brix (%), juice volume and bagasse fresh weight (t/ha). The phenotypic correlations were partitioned into direct and indirect effects on juice yield.

The highest positive and direct effect was found for bagasse weight, followed by panicle width, brix, stalk diameter, plant height and number of leaves. The negative and direct effects were found for panicle length, 1000 grain weight, lodging score and days to 50% flowering.

Bagasse weight, number of leaves per plant, stalk diameter, days to 50% flowering and brix contributed highly positively and indirectly to juice yield via plant height. Number of leaves per plant, days to flowering, plant height, and brix had a highly positive and indirect effect on juice yield through stalk diameter. Bagasse weight, days to 50% flowering, stalk diameter, plant height and brix had a highly positive and indirect effect on juice yield via number of leaves per plant. Bagasse weight, days to 50% flowering, number of leaves per plant, stalk diameter and plant height had a highly positive and indirect effect on juice yield through brix. Number of leaves per plant, stalk diameter, plant height, days to 50% flowering and brix had a highly positive and indirect effect on juice yield via bagasse weight.

Table 4.4   Phenotypic path coefficient analysis indicating direct (in bold) and indirect effects of components traits on juice yield

	DFL	PHT	SD	LPP	LS	PS	PAL	PAW	PWT	GWT	BR	BFW	PATH
DFL	**-0.00007**	0.040	0.077	0.023	0.005	-0.084	0.110	-0.019	-0.004	0.057	0.061	0.295	0.560**
PHT	-0.00002	**0.120**	0.071	0.018	-0.016	-0.009	-0.060	0.077	0.004	0.013	0.033	0.379	0.630**
SD	-0.00003	0.059	**0.146**	0.020	-0.006	0.032	0.052	0.028	0.004	0.030	0.050	0.384	0.800**
LPP	-0.00005	0.071	0.098	**0.031**	0.008	-0.050	0.076	0.009	0.004	0.043	0.053	0.389	0.730**
LS	0.00001	0.031	0.013	-0.004	**-0.063**	0.056	-0.024	0.047	-0.001	-0.007	-0.008	0.079	0.120
PS	0.00003	-0.005	0.022	-0.007	-0.016	**0.216**	-0.079	0.015	0.006	-0.052	-0.085	-0.054	-0.040
PAL	0.00003	0.028	-0.029	-0.009	-0.006	0.065	**-0.262**	0.133	0.008	-0.038	-0.067	-0.064	-0.240
PAW	0.00001	0.043	0.019	0.001	-0.014	0.015	-0.162	**0.215**	0.023	0.005	-0.023	0.098	0.220
PWT	0.00001	0.011	0.013	0.003	0.002	0.026	-0.047	0.105	**0.047**	0.010	-0.029	-0.020	0.120
GWT	0.00004	-0.016	-0.045	-0.013	-0.004	0.117	-0.102	-0.011	-0.005	**-0.097**	-0.096	-0.158	-0.430**
BR	-0.00003	0.025	0.047	0.010	0.003	-0.119	0.113	-0.032	-0.009	0.060	**0.155**	0.207	0.460**
BFW	-0.00004	0.092	0.114	0.024	-0.010	-0.024	0.034	0.043	-0.002	0.031	0.065	**0.492**	0.860**

* Correlation is significant at the 0.05 level (2-tailed), * *Correlation is significant at 0.01 level (2-tailed)DFL = Days to 50% flowering, PHT = Plant height (cm), SD = Stalk diameter (mm), LPP = Number of leaves per plant, LS = Lodging score (%), PS= Pest infestation score (%), PAL = Panicle length (cm), PAW = Panicle width (cm), PWT = Panicle weight (t/ha), GWT = 1000 grain weight (g), BR = Brix (%), JV = Juice volume, BFW = Bagasse fresh weight (t/ha).

4.7 Discussion

4.7.1 Variance components

Analysis of variability among the traits and the association of a trait in relation to other traits contributing to yield of a crop would be important in planning a successful breeding programme (Mary and Gopalan, 2006). Development of high-yielding cultivars needs comprehensive information of the genetic variation present for juice yield and its contributing traits. The observed variability is a collective estimate of genetic and environmental causes, of which only the former is inherited. However, estimates of heritability alone do not provide information on the expected gain in the next generation, but must be considered together with estimates of genetic advance, the change in mean value among successive generations (Shukla et al., 2016; Al-Tabbal and Al-Fraihat, 2011).

For all traits studied, variance components showed higher genotypic and phenotypic variance estimates than environmental variance estimates. This implies that expression of these traits was genetic, which can be exploited in breeding programs. This corroborates the report by Al-Tabbal and Al-Fraihat (2011). The estimates for GCV were high for juice volume (65.6) and low for days to 50% flowering (12.7). The trend was similar for PCV where juice volume had 67.5 and days to flowering had 12.8. However, GCV values for days to 50% flowering, plant height and stalk yield were closer to their respective PCV values. Similar results were also reported by Kalpande et al. (2014). The author reported that the environment had little influence on the phenotypic expression of these traits. Genotypic coefficient of variation reveals the extent of genetic variability present in the genotypes for various traits (Singh, 2000). Deshmukh et al. (1986), reported that GCV and PCV values above 20% are regarded as high while those below 10% are regarded as low. In this study therefore, all GCV and PCV values for the traits were high.

4.7.2 Heritability estimates and genetic advance

Broad sense heritability (H^2) estimates were high for all the traits measured. The highest was 98.1%. Similar results were reported by Mahajan et al. (2011) and Amare et al. (2015). Heritability is useful for genetic improvement because of its predictive role to indicate the reliability of the phenotypic value as a guide to breeding value (Falconer and Mackay, 1996). Several traits had H^2 values greater than 80%. This implies that selection for these traits can be effective, because there would be high response to direct selection as reported by Singh (2001) and Shadakshari et al. (1995).

Highest genetic advance (GA) values were observed in juice volume (12465.8 l/ha), while stalk yield and %brix recorded GA values of 13.0 t/ha and 8.1%, respectively. Johnson et al. (1955), reported that high heritability coupled with high genetic advance is an important factor for predicting the resultant effect for selecting the best individuals. From the values observed in this study, it implies that if the top performing 5% were selected as parents, the mean juice volume, brix and stalk yield for progenies would be improved by 12465.8 l/ha, 8.1% and 13.0 t/ha, respectively. The estimates of GA help in understanding the type of gene action controlling the expression of several polygenic traits. High values of GA indicate additive gene action while low values are evident of non-additive gene action (Singh and Narayanan, 1993). This, therefore, implies that the heritability estimates will be useful if accompanied by high GA.

From the results, high expected genetic advance expressed as percentage of the mean (GAM) were observed in juice volume (131.2%), stalk yield (96.3%), fresh bagasse weight (90.6%), and brix (73.8%). According to a study by Al-Tabbal and Al-Fraihat (2011), high values indicate that these traits are simply inherited, and most likely the heritability is due to additive gene effects and selection may be effective in early generations for these traits. Therefore, selection based on these traits with high GAM will result in improved performance of the genotypes for these traits.

4.7.3   Correlations between juice and juice related traits

Sweet sorghum breeding programmes aim at developing superior cultivars for biofuel production. Since several agronomic and industrial traits are related to biofuel production, it is necessary to determine the level of association, which can help breeders establish better selection criteria in the breeding programmes (Lombardi et al., 2015).

Results from the combined analysis revealed several correlations of the traits under study. Number of days to 50% flowering was positively and highly significantly correlated to plant height, number of leaves, stalk yield, juice volume, %brix and bagasse weight. Plant height and stalk diameter exhibited a positive and significant correlation with number of leaves per plant, stalk diameter, stalk yield, juice volume, and fresh bagasse weight. However, plant height displayed a negative and significant correlation with panicle, length, panicle width, panicle weight and 1000 grain weight. These results agree with those reported by Kisua et al. (2015), Audilakshmi et al. (2010) and Murray et al. (2008). Plant height also displayed a weak positive correlation with brix. These positive correlations among traits such as days to 50% flowering, plant height, stalk yield, juice volume, brix and bagasse weight, suggest that tall plants that flower later have ample time to accumulate photosynthates and more time to accumulate stem biomass, in which to accumulate sugars (Ferraris and Charles-Edwards,

1986; Ritter et al., 2008). According to Shukla et al. (2006), positive correlation between plant height and sugar accumulation complicates the development of short sweet and lodging resistant inbred lines. Knowledge of correlation is important to obtain the expected response of other traits when selection is applied to a trait of interest in a breeding programme. Therefore, as suggested by Audilakshmi et al. (2010), selection for stalk yield, juice volume and bagasse weight in sweet sorghum can be achieved through indirect selection for plant height and stalk diameter. A 1000-grain weight was negatively and significantly correlated to stalk yield, brix and juice volume. These results are consistent with those reported by Zou et al. (2011) and Makanda et al. (2009). The negative and highly significant correlation between these traits suggests grain yield penalty as the improvement is focused on biomass and juice volume. As reported by Makanda et al. (2009), this association suggests that genes controlling grain yield and juice volume are antagonistic to those controlling grain yield. In other studies, Fisher and Wilson (1976) and Broadhead and Freeman (1980) reported that lack of seed development increases sugar accumulation because of changes in assimilate partitioning, with the stem predominantly becoming the sink. Panicle length was negatively correlated to brix and juice volume and ultimately bagasse. This suggests the existence of a tradeoff between panicle length against brix, juice volume and bagasse. These results however, contradict the findings reported by Zou et al. (2011).

Correlation estimates enable the evaluation of the behaviour of one or more traits using the behaviour of the other. Therefore, a detailed understanding of inter-trait associations makes it feasible to conduct indirect selection. This is particularly significant for traits that are not easily quantified and that are associated with low heritability (Lombardi et al., 2015). Positively correlated traits suggest gene linkage and pleiotropic effects (Kisua et al., 2015; Kearsey and Pooni, 1998). Kisua et al. (2015), also reported that strong positive correlations among genotypes suggest that such traits are heritable and genetically controlled, which means they can be transmitted to the desired genotypes.

4.7.4 Path coefficient analysis

Path analysis is a standard partial regression coefficient which facilitates partitioning of correlation coefficient into direct and indirect effects of several traits towards dependent variables. It helps in evaluating the cause-effect relationship and effective selection (Bello et al, 2010).

Results of path analysis of component traits of juice yield showed maximum positive direct effect of bagasse weight, followed by panicle width, brix, stalk diameter, plant height and number of leaves. These results agree with the earlier report by Sandeep et al. (2011). On the

other hand, bagasse weight, number of leaves per plant, days to 50% flowering (which had earlier showed highly negative and direct contribution on juice) and stalk diameter had a positive indirect effect on juice via plant height. Number of leaves per plant, days to flowering, plant height, bagasse weight and brix had a highly positive and indirect effect on juice yield via stalk diameter. Bagasse weight, days to flowering, number of leaves per plant, plant height and brix had a highly positive and indirect effect on juice yield via number of leaves per plant. These results agree with reports by Mallikarjun et al. (1998) and Kachapur and Salimath (2009). In general, these results indicated the indirect contribution of several traits resulted in their positive correlation with juice yield, as earlier suggested by Sandeep et al. (2011). Results on association of juice yield with its attributing traits indicated importance of plant height, stalk diameter, number of leaves per plant, bagasse weight and brix, because these traits exhibited a direct relation with juice yield. As suggested by Sandeep et al. (2011), improvement on these traits automatically improves juice yield.

4.8   Conclusion

High heritability was observed in plant height, stalk yield, days to 50% flowering and juice volume, suggesting that selection for these traits can be effective. GAM for juice volume, bagasse yield, stalk yield and %brix was high, indicating that these traits are controlled by additive gene action, and that they are easily inherited. There were positive and significant correlations between plant height and biomass yield, biomass yield and juice volume and bagasse, suggesting that improvement for these traits can be carried out simultaneously. Bagasse weight, brix, stalk diameter, plant height and number of leaves had a highly positive and direct contribution on juice yield. Several traits had a highly positively and indirect contribution on juice yield via these traits which had a direct contribution. The most desirable lines of sweet sorghum should have less number of days to 50% flowering, more biomass yield, and be tall.

References

Al-Tabbal, J.A. and A.H. Al-Fraihat. 2011. Genetic Variation, Heritability, Phenotypic and Genotypic Correlation Studies for Yield and Yield Components in Promising Barley Genotypes. Journal of Agricultural Science 4: 193-210.

Ali, M.L., J.F. Rajewski, P.S. Baenziger, K.S. Gill, K.M. Eslridge and I. Dweikat. 2008. Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Molecular Breeding 21: 497-509.

Amare, K., H. Zeleke and G. Bultosa. 2015. Variability for yield, yield related traits and association among traits of sorghum (*Sorghum bicolor* (L.) Moench) varieties in Wollo, Ethiopia. Journal of Plant Breeding and Crop Science 7: 125-133.

Audilakshmi, S., A.K. Mall, M. Swarnalatha and N. Seetharama. 2010. Inheritance of sugar concentration in stalk (Brix), sucrose content, stalk and juice yield in sorghum. Biomass Bioenergy 34: 813-820.

Broadhead, D.M. and K.C. Freeman. 1980. Stalk and sugar yield of sweet sorghum as affected by spacing. Agronomy Journal 72: 523-524.

Cochran, W.G. and G.M. Cox. 1992. Experimental Designs. second edition ed. Wiley Interscience. Canada.

Deshmukh, S., M. Basu and P. Reddy. 1986. Genetic variability, character association and path coefficients of quantitative traits in Virginia bunch varieties of groundnut. Indian Journal of Agricultural Sciences 56: 816-821.

Dewey, D.R. and K.H. Lu. 1959. A correlation and path coefficient analysis of components of crested wheatgrass seed production. Agronomy Journal 51: 515-518.

Falconer, D.S. and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Benjamin Cummings, England.

Ferraris, R. and D. Charles-Edwards. 1986. A comparative analysis of sweet and forage sorghum crops II: accumulation of soluble carbohydrates and nitrogen. Australian Journal of Agricultural Research 37: 513-533.

Fisher, K.S. and G.L. Wilson. 1976. Studies of grain production (*Sorghum bicolor* (L) Moench). VII: contribution of plant parts of canopy photosynthesis and grain yield in field situations. Australian Journal of Agricultural Research 27.

Hanson, C., H. Robinson and R. Comstock. 1956. Biometrical studies of yield in segregating populations of Korean Lespedeza. Agronomy Journal 48: 268-272.

IBPGR, I. 1993. Descriptors for sorghum (*Sorghum bicolor* (L.) Moench) International Board for Plant Genetic Resources, Rome, Italy.

Johnson, H.W., H.E. Robinson and R.E. Comstock. 1955. Estimates of genetic and environmental variability in soybean. Agronomy Journal 47: 314-318.

Kachapur, R. and P. Salimath. 2009. Genetic studies on correlation and character association in sweet sorghum (*Sorghum bicolor* (L.) Monech). Green Farming 2: 343-346.

Kalpande, H.V., S.K. Chavan, A.W. More, V.S. Patiland and P.B. Unche. 2014. Character association, genetic variability and component analysis in sweet sorghum [*Sorghum bicolor* (L. Moench)]. Journal of Crop and Weed 10: 108-110.

Kearsey, M.J. and H.S. Pooni. 1998. Genetic Analysis of Quantitative Traits Taylor and Francis, New York.

Kisua J., Mwikamba K., M. M. and M. A. 2015 Genetic diversity of sweet and grain sorghum using phenotypic markers. International Journal of Biosciences 6: 34-46.

Lombardi, G.M.R., Nunes J.A.R., Parrella R.A.C., Teixeira D.H.L., Bruzi A.T., Durães N.N.L. 2015. Path analysis of agro-industrial traits in sweet sorghum. Genetics and Molecular Research. 14: 16392-16402.

Lzge, A.U., A. Kadams and D.T. Gungula. 2006. Studies on character association and path analysis of certain quantitative characters among parental lines of pearl millet (*Pennisetum glaucum* L.) and their F1 hybrids in a diallel cross. African Journal of Agricultural Research 1: 194-198.

Mahajan, R., P. Wadikar, S. Pole and M. Dhuppe. 2011. Variability, correlation and path analysis studies in sorghum. Research Journal of Agricultural Sciences 2: 101-103.

Makanda, I., P. Tongoona and J. Derera. 2009. Quantification of genotypic variability for stem sugar accumulation and associated traits in new sweet sorghum varieties. African Crop Science Conference Proceedings 9: 391-398.

Mallikarjun, H., S. Khanure and M. Kachapur. 1998. Correlation and path analysis for juice quality parameters in sweet sorghum genotypes. Madras Agricultural Journal 85: 207-208.

Mary, S.S. and A. Gopalan. 2006. Dissection of genetic attributes yield traits of fodder cowpea in F3 and F4. Applied Science Research Journal 2: 805-808.

Murray, S.C., W.L. Rooney, M.T. Hamblin, S.E. Mitchell and S. Kresovich. 2009 Sweet sorghum genetic diversity and association mapping for brix and height. The Plant Genome 2: 48-62.

Murray, S.C., A. Sharma, W.L. Rooney, P.E. Klein, J.E. Mullet, S.E. Mitchell, and S. Kresovich. 2008. Genetic improvement of sorghum as a biofuel feedstock I: QTL for stem sugar and grain non-structural carbohydrates. Crop Science 48: 2165-2179.

Prasad, S., A. Singh, N. Jain and H.C. Joshi. 2007. Ethanol production from sweet sorghum syrup for utilization as automotive fuel in India, Energy & Fuels. Plant Science 21: 2415-2420.

Ritter, K.B., D.R. Jordan, Chapman S.C., Godwin I.D., E.S. Mace and C. McIntyre. 2008. Identification of QTL for sugar related traits in sweet x grain sorghum [*Sorghum bicolor* (L) Moench] recombinant inbred population. Molecular Breeding 22: 367-384.

Sandeep, R., M.G. Rao, B.V. Bhat, S. Rao, R. Kulkarni, S. Hittalmani and C.A.S. Murthy. 2011. Research Note Inter-relationship between sugar yield and its component characters in two segregating populations of sweet sorghum [*Sorghum bicolor* (L.) Moench]. Electronic Journal of Plant Breeding 2: 244-247.

Shadakshari, Y.G., K. Virupahshappa and G. Shivasharikar. 1995 Genetic variability studies in the germplasm collections of sesamum (*Sesamum indicum* L.). Mysore Journal of Agricultural Science 29: 133 -137.

Shukla, S., A. Bhargava, A. Chatterjee, A. Srivastava and S.P. Singh. 2006. Genotypic variability in vegetable amaranth (*Amaranthus tricolor* L.) for foliage yield and its contributing traits over successive cuttings and years. Euphytica 151: 103-110.

Shukla, S., T.J. Felderhoff, A. Saballos and W. Vermerris. 2016. The relationship between plant height and sugar accumulation in the stems of sweet sorghum (*Sorghum bicolor* (L) Moench). Field Crops Research 203: 181-191.

Singh, B. 2001. Plant breeding: Principles and methods. Kalyani Polisher, New Delhi.

Singh, B.D. 2000. Plant breeding: principles and methods Kalyani Publishers, New Delhi.

Singh, P. and S.S. Narayanan. 1993. Biometrical techniques in plant breeding. Kalyani, Publishers, New Delhi.

Singh, R.K. and B.D. Chaudhary. 1979. Biometrical methods in quantitative genetic analysis. Kalyani publishers, Ludhiana, New Delhi India.

SPSS. 2006. Statistical Package for Social Sciences-Users guide. Chicago.

Vendruscolo, T.P.P., M.A.A. Barelli, M.A.S. Castrillon and R.S. da Silva. 2016. Correlation and path analysis of biomass sorghum production. Genetic and Molecular Research 15: 1-10.

Wright, S. 1921. Correlation and causation. Journal of Agricultural Research 20:557-585.

Zou, G., S. Yan, G. Zhai, Z. Zhang, J. Zou and Y. Tao. 2011. Genetic variability and correlation of stalk yield-related and sugar concentration of stalk juice in a sweet sorghum (*Sorghum bicolor* L. Moench) population. Australian Journal of Crop Science 5: 1232-1238.

CHAPTER 5

Assessment of genetic diversity in sweet sorghum accessions using single nucleotide polymorphism (SNP) markers

Abstract

Sweet sorghum is an alternative feed stock for ethanol production. Knowledge on genetic diversity of sweet sorghum is key in identifying groups of similar genotypes for conservation purpose and proper use of genetic resources, further to the protection of property rights. The use of molecular markers in genetic diversity studies has revolutionized the pace and precision and has facilitated implementation of molecular breeding of crops. In this study 53 sweet sorghum accessions were studied using single nucleotide polymorphisms (SNP) to examine the level of genetic diversity. Kompetitive Allele Specific Polymorphism (KASP) genotyping was done using 137 SNP markers. Three populations were generated from the analysis. The expected heterozygosity (H_e) values ranged from 0.236 to 0.291 with a mean of 0.266. The observed heterozygosity (H_o) ranged from 0.007 to 0.038 with an overall mean of 0.021. The mean of effective alleles across populations was 1.438. The percentage of polymorphic loci ranged from 80.29% to 91.24% with a mean of 86.86%. Dissimilarity indices ranged from 0.000 to 0.583 with a mean of 0.296. The highest dissimilarity index was observed between SA 2193 and SA 2014, which implied a considerable amount of genetic diversity. Accessions were clustered into three main groups based on dissimilarity indices. The study provided information on presence of genetic diversity among the sweet sorghum accessions, which can be utilised by plant breeders and geneticists for gene mining and introgression for future breeding programmes for development of superior genotypes for biofuel production in South Africa.

Key words: alleles, cluster analysis, dissimilarity matrix, heterozygosity, single nucleotide polymorphism

5.1 Introduction

Sweet sorghum (*Sorghum bicolor* L.) has been recognized broadly as the potential alternative source of biofuel because of high fermentable sugar content in the stalk (Ali et al., 2007; Prasad et al., 2007). It is tolerant to water stress due to several morphological and physiological features, including a widespread root system, waxy bloom on leaves that reduces water loss, ability to stop growth during water stress and resume it when the conditions are favourable, and a C_4 photosynthetic pathway (Balole and Legwaila, 2005).

Knowledge of the genetic diversity of sweet sorghum is key in identifying groups of similar genotypes for conservation purposes and proper use of genetic resources, in addition to the protection of property rights (Cox et al., 1985). Genetic diversity assessment can be achieved using morphological and molecular markers (Karp et al., 1997). Morphological markers are phenotypic traits such as flower colour, seed colour and shape, growth habits, pigmentation, texture, maturity, yield, and pest and disease resistance (Franco et al., 2001). However, morphological or pedigree evaluations are influenced by the environment, management practices, and developmental stage of the plant, among other drawbacks (Semagn, et al., 2006). Thus molecular markers are now preferred for genetic diversity analysis.

Molecular markers are nucleotide sequences found naturally in populations, which show neutral sites of variation at DNA sequence level (Jiangfeng et al., 2014). They are biological features determined by allelic forms of genes or genetic loci and can be inherited by succeeding generations. They can be used as experimental probes or tags to keep track of an individual, a tissue, a cell, a nucleus, a chromosome or a gene. The use of molecular markers allows plant breeders and geneticists to locate and understand the basics of several gene interactions which determine complex traits (Haussmann et al., 2000).

Depending on detection method and throughput, all molecular markers can be grouped as low-throughput, hybridization-based markers such as restriction fragment length polymorphisms (RFLPs) (Botstein et al., 1980), medium-throughput, PCR-based markers that include random amplification of polymorphic DNA (RAPD), simple sequence repeats (SSRs) (Welsh and McClelland, 1990; Jacob et al., 1991) and high-throughput (HTP), sequence-based markers, under which single nucleotide polymorphisms (SNPs) fall (Wang et al., 1998). Simple sequence repeats are more abundant in non-coding regions than in coding regions. They are regarded as the most efficient markers compared to other markers because they are highly polymorphic (Ngugi and Onyango, 2012).

Although SNPs are less polymorphic than SSR markers because of their biallelic nature, they are abundant; they have a wide genomic coverage, and are amenable to high- and ultra-high-throughput automation (Mammadov et al., 2012). Single nucleotide polymorphisms can be used to study genetic diversity, create genetic maps and in marker assisted selection (MAS) breeding in many crop species, and they are termed as "the marker of choice" (Graves et al., 2016). These markers have been progressively used for QTL mapping studies, mainly, because they are ubiquitous in the genomes and, therefore, they can provide the uppermost map resolution compared to other marker systems (Bhattramakki et al., 2002).

In maize, SNP markers have enabled the dissection of complex traits such as flowering time (Buckler et al., 2009). Using a set of 5000 RILs, which represent the nested association mapping population and genotyping with 1,200 SNP markers, the authors reported that the genetic architecture of anthesis time is governed by small additive QTL rather than a single large effect QTL. In wheat genomics, recent advances have led to the implementation of high-density SNP genotyping. Gene-based SNP markers were developed that confer resistance to leaf rust, stripe rust, and powdery mildew diseases. These markers serve as efficient tools for marker assisted selection (MAS) of disease resistant wheat lines (Akhunov et al., 2009).

In a study by Yu et al. (2011) on QTL analysis in rice for yield and three-yield-contributing traits (number of tillers per plant, number of grains per panicle and grain weight) the authors compared a SNP-based map to that of a previous RFLP/SSR-based QTL map generated using the same mapping population. Using the ultra-high-density SNP map, the authors reported that this map was more powerful and had a high resolution unlike the RFLP/SSR map. As suggested by Mammadov et al. (2012), beside the power and the resolution, maps based on high-density SNP markers are also appropriate for fine mapping and cloning of QTL and often, SNPs on these maps are also functionally associated with the natural variation in the trait. Progress in rice genomics has led to mapping and cloning of various genes and QTL influencing expression of agronomically important traits, enabled routine use of SNP markers for MAS, gene pyramiding and MAB (Ashikari and Matsuoka, 2006; Jena and Mackill, 2008).

In a genetic diversity and association mapping study of sweet sorghum using SNP markers, Murray et al. (2009) identified three main genetic groupings of sweet sorghums. Based on observed phenotypes and known backgrounds, sweet sorghum accessions were clustered into three groups as historical and modern syrup, modern sugar/energy types, and amber types. Using information on population structure and relatedness, association mapping was performed for height and stem sugar (brix) traits. Three significant associations for height were also observed. Therefore, the objectives of the study were to examine the genetic diversity

and relatedness among the sweet sorghum genotypes using single nucleotide polymorphism (SNP) markers, and to assess the genetic variability for juice and sugar content among sweet sorghum genotypes.

5.2 Hypothesis

For this study, the hypothesis tested was that there is no genetic diversity among South African sweet sorghum germplasm.

5.3 Materials and Methods

5.3.1 Plant materials

Fifty-three sweet sorghum accessions were used in this study as presented in Table 5.1. these accessions were obtained from the Agricultural Research Council-Grain Crops (ARC-GC) in south Africa.

Table 5.1   Sweet sorghum accessions used in the study

Entry No.	Accession Designation	Entry No.	Accession Designation
1	SA 0029	28	SA 4477
2	SA 0096	29	SA 4478
3	SA 0163	30	SA 4479
4	SA 0240	31	SA 4481
5	SA 0307	32	SA 4482
6	SA 0311	33	SA 4483
7	SA 0312	34	SA 4484
8	SA 0317	35	SA 4485
9	SA 0831a	36	SA 4486
10	SA 1238	37	SA 4488
11	SA 1242	38	SA 4489
12	SA 1330	39	SA 4490
13	SA 1382	40	SA 4491
14	SA 160	41	SA 4494
15	SA 1904	42	SA 4495
16	SA 2012	43	SA 4501
17	SA 2013	44	SA 4503
18	SA 2014	45	SA 4506
19	SA 2034	46	SA 4510
20	SA 2036	47	SA 4520
21	SA 2116	48	SA 4523
22	SA 2193	49	SA 4524
23	SA 2195	50	SA 4526
24	SA 2218	51	SA 4528
25	SA 2249	52	SA 4534
26	SA 2330a	53	SA 4553
27	SA 2400		

5.3.2   Growth environment

These accessions were grown in the growth chamber in pots. Vermiculite was used as the growth media. Five seeds from each accession were planted per pot.

5.3.3   DNA Sampling and Isolation

Five weeks after planting, ten leaf discs were harvested from all the five plants per pot (two leaf discs per plant). These were used for the DNA extraction. The sampling kit was obtained from LGC Genomics Laboratory, United Kingdom. This included a 96-well plate, cutting mat and leaf cutting tool. The leaf was placed on a cutting mat. The leaf cutting tool was then held vertically over the leaf, pushed into the leaf tissue and twisted to cut and pick the leaf disc up in the cutting tool. One leaf punch (disc) was collected at a time. The end of the cutting tool was then inserted starting from the second well of the 96-well storage rack. The plunger was then depressed to dispense the leaf disc. The first and the last strips were deliberately left unfilled, according to LGC protocol. Leaf samples from the same genotype were placed in a specific well of the 96-well storage rack. After collecting the required number of discs from each genotype, the cutting tool was washed by placing the end of the cutting tool into a container of clean distilled water. The plunger was depressed 5-10 times. After washing, the cutting tool was then flicked for a while until it was completely dry. The steps were repeated until all the genotypes were sampled. The cutting tool was continuously washed in between sampling of individual genotypes.

Each strip of the tube within the rack was sealed using perforated trip cap. The desiccant sachet was placed directly on top of the strip cap-sealed tubes and the plastic lid was replaced on top. The storage rack was secured using an elastic band and was placed inside a labelled sealable plastic bag. Excess air was forced out of the sealable bag. The sealed bag was placed into the plant kit box. The samples were then shipped to LGC Genomics Laboratory in the United Kingdom for DNA extraction and genotyping.

5.3.4   SNP selection and amplification

In compliance with the protocol supplied by LGC genomics laboratory, Kompetitive Allele Specific Polymerase Chain Reaction (KASP) genotyping assays were used. These were based on competitive allele-specific PCR and enable bi-allelic scoring of single nucleotide polymorphisms (SNPs) and insertion and deletions (Indels) at specific loci.

The SNP-specific KASP Assay mix and the universal KASP Master mix (supplied at 2X concentration) were used. KASP Master Mix contains Taq polymerase enzyme and passive reference dye, 5-carboxy-X-rhodamine, succinimidyl ester (ROX) and $MgCl_2$ in an optimized buffer solution. The two mix were added to DNA samples then a thermal cycling was performed, followed by an end-point fluorescent read. Allele-specific primers each harbouring a unique tall sequence that correspond with a universal fluorescence resonant energy transfer (FRET) cassette; one labelled with FAMTM dye and the other with HEXTM dye were used. During thermal cycling, the relevant allele-specific primer would bind to the template and elongate, thus attaching the tail sequence to the newly synthesized strand. The complement of the allele-specific tail sequence was then generated during subsequent rounds of PCR, enabling the FRET cassette to bind to the DNA. Bi-allelic discrimination was achieved through the competitive binding of the two allele-specific forward primers. If the genotype was heterozygous, a mixed fluorescent was generated. If the genotype at a given SNP was homozygous, only one of the two possible fluorescent signal was generated. One hundred and thirty-seven assays were used in this study.

1) Assay components:

KASP uses three components: test DNA with the SNP of interest; KASP Assay Mix containing two different, allele-specific, competing forward primers with unique tail sequences and one reverse primer; the KASP Master mix containing FRET cassette plus Taq polymerase in an optimised buffer solution.
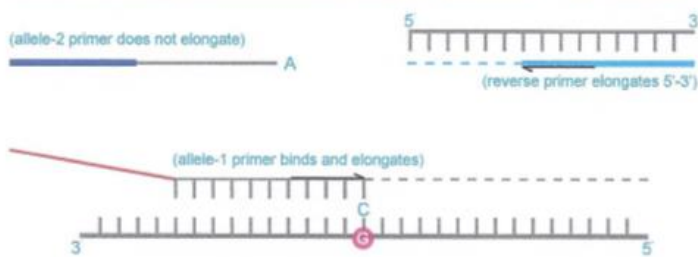
A) KASP Assay mix

Allele-specific forward primers:

allele-1
5'――――――――― C 3'

allele-2
5'――――――――― A 3'

Reverse primer:
3'――――――――― 5'

B) KASP Master mix

F Q   H Q
5' 3'   5' 3'

3' 5'   3' 5'

C) DNA template (sample)

5' ―――――――――――――――― 3'
3' ―――――――――――――――― 5'
[G/T]

2) Denatured template and annealing components – PCR round 1:

(allele-2 primer does not elongate)
―――――――― A

5' ――――――――― 3'
(reverse primer elongates 5'-3')

(allele-1 primer binds and elongates)
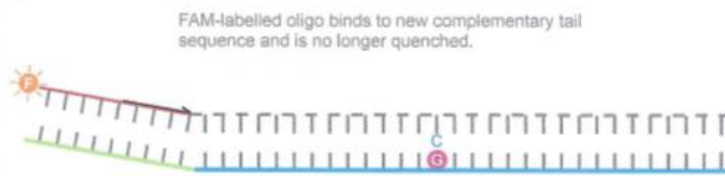
C
G
3' ――――――――――――――――― 5'

In the first round of PCR, one of the allele-specific primers matches the target SNP and, with the common reverse primer, amplifies the target region.

3) Complement of allele-specific tail sequence generated – PCR round 2:

C
G
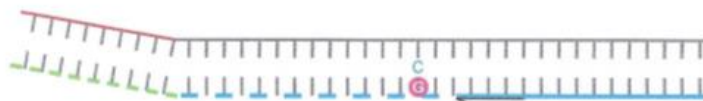
(Reverse primer binds, elongates and makes a complementary copy of the allele-1 tail.)

Legend

Allele-1 tail FAM-labelled oligo sequence

Allele-2 tail HEX-labelled oligo sequence

Common reverse primer

F   FAM dye

H   HEX dye

Target SNP

Q   Quencher

4) Signal generation – PCR round 3:

FAM-labelled oligo binds to new complementary tail sequence and is no longer quenched.

F

C
G

In further rounds of PCR, levels of allele-specific tail increase. The fluor labelled part of the FRET cassette is complementary to new tail sequences and binds, releasing the fluor from the quencher to generate a fluorescent signal.

Q   H Q
5' 3'   5' 3'

3' 5'   3' 5'

Fluor for incorporated G allele no longer quenched.

Fluor for non-incorporated T allele remains quenched.

Figure 5-1   Illustration of KASP procedure (https://www.integratedbreeding.net)

5.4 Data analysis

The SNP data from LGC Genomics grid report was recoded across all the 53 genotypes by assigning (1) for the more frequent allele, (2) for the less frequent allele and (-1) for missing data. There was a significant amount of missing data for genotype SA2034 and this entry was excluded from analysis altogether. Simple matching molecular genetic dissimilarity indices (genetic distances were determined among 52 sorghum genotypes using DARwin version 6.0.12 software (Perrier and Jacquemoud-Collet, 2006).

The genetic distance between any two sorghum genotypes was determined as:

$$d_{ij} = \frac{1}{L} \sum_{l=1}^{L} \frac{ml}{\pi}$$

Where,

d_{ij} = dissimilarity (genetic distance) between genotypes i and j

L = number of loci

π = ploidy

ml = number of matching alleles for locus l.

Analysis of molecular variance (AMOVA) was carried out using GenAlEx version 6.503 software (Peakall and Smouse, 2006; 2012).

5.4.1 Genetic diversity analysis

Genotypic data were subjected to analyses with various measures of genetic diversity within and among accessions using GenAlex software version 6.5 (Peakall and Smouse, 2012). Genetic diversity parameters such as total number of alleles per locus (N_a), number of effective alleles per locus (N_e), observed heterozygosity (H_o), Shannon's Information Index (I) and gene diversity (He) were determined using the protocol devised by Nei and Li (1979). To examine the degree of population differentiation, other genetic parameters such as differentiation (FST), gene flow (Nm), Nei's unbiased genetic distance (GD) and identity (GI) were estimated using GenAlex. The partitioning of total genetic variation into within and among populations was done with a molecular analysis of variance (AMOVA) procedure using GenAlex. For each

locus, number of alleles, number of effective alleles, information index, observed heterozygosity, expected heterozygosity and fixation index were determined.

5.4.2 Cluster analysis

The dissimilarity indices were used to cluster the genotypes using the Unweighted Neighbour joining method. The 52 genotypes were first structured by clustering using DARwin into 4 groups (populations) designated I, II, III, and IV. However, group IV consisted of only one member and this would not allow AMOVA to proceed in GenAlEx so this group was added onto group III, and in total three groups of genotypes Group I (18 individuals), Group II (18 individuals) and Group III (16 individuals) were analysed for within and between group diversity. To examine the genetic relationships among accessions, genetic distances between all pairs of individual accessions were estimated to draw a dendrogram.

5.5 Results

5.5.1 Genetic parameters

Information on sample size (N), number of alleles (N_a), number of effective alleles (N_e), information index (I), observed heterozygosity (H_o), expected and unbiased expected heterozygosity (uH_e), and fixation Index (F) is presented in Table 5.1. This shows the top 30 markers in order of observed heterozygosity.

Table 5.2 Genetic diversity within and among 52 sweet sorghum accessions based on 52 SNP marker analysis for top 30 markers based on H_e

Locus	N	N_a	N_e	I	H_o	H_e	uH_e	F
SB_10_012	17	2.000	1.710	0.606	0.235	0.415	0.428	0.433
SB00128_2	15	2.000	1.867	0.657	0.200	0.464	0.480	0.569
SB_04_088	15	2.000	1.385	0.451	0.200	0.278	0.287	0.280
SB_04_095	16	2.000	1.205	0.311	0.188	0.170	0.175	-0.103
SB_05_011	17	2.000	1.486	0.508	0.176	0.327	0.337	0.460
SB_05_095	17	2.000	1.637	0.578	0.176	0.389	0.401	0.547
SB_08_018	17	2.000	1.637	0.578	0.176	0.389	0.401	0.547
SB_03_012	18	2.000	1.314	0.403	0.167	0.239	0.246	0.303
SB_04_095	18	2.000	1.737	0.615	0.167	0.424	0.437	0.607
SB_10_087	18	2.000	1.600	0.562	0.167	0.375	0.386	0.556
SB_07_022	13	2.000	1.899	0.666	0.154	0.473	0.492	0.675
SB_07_022	14	2.000	1.960	0.683	0.143	0.490	0.508	0.708
SB00135_1	15	2.000	1.923	0.673	0.133	0.480	0.497	0.722
SB_06_015	15	2.000	1.642	0.580	0.133	0.391	0.405	0.659
SB_08_095	15	2.000	1.642	0.580	0.133	0.391	0.405	0.659
SB_06_078	15	2.000	1.991	0.691	0.133	0.498	0.515	0.732
SB_06_078	16	2.000	1.969	0.685	0.125	0.492	0.508	0.746
SB_03_012	16	2.000	1.280	0.377	0.125	0.219	0.226	0.429
SB_10_047	16	2.000	1.753	0.621	0.125	0.430	0.444	0.709
SB00165_1	17	2.000	1.710	0.606	0.118	0.415	0.428	0.717
SB00214_1	17	2.000	1.562	0.546	0.118	0.360	0.371	0.673
SB_01_054	17	2.000	1.993	0.691	0.118	0.498	0.513	0.764
SB_07_107	17	2.000	1.262	0.362	0.118	0.208	0.214	0.433
SB_08_092	17	2.000	1.562	0.546	0.118	0.360	0.371	0.673
SB_09_106	17	2.000	1.125	0.224	0.118	0.111	0.114	-0.063
SB_09_017	18	2.000	1.246	0.349	0.111	0.198	0.203	0.438
SB_05_114	18	2.000	1.906	0.668	0.111	0.475	0.489	0.766
SB_05_138	18	2.000	1.385	0.451	0.111	0.278	0.286	0.600
SB_07_079	18	2.000	1.906	0.668	0.111	0.475	0.489	0.766
SB_08_040	18	2.000	1.246	0.349	0.111	0.198	0.203	0.438

N= Sample size, N_a= Number of Alleles, N_e= Number of Effective Alleles, I= Fixation Index, H_o= Observed Heterozygosity, H_e= Expected Heterozygosity, uH_e= Unbiased Expected Heterozygosity, I= Fixation Index.

The number of alleles amplified pre locus (Table 5.3) ranged from 1.803 to 1.92 with a grand of 1.869, number of effective alleles amplified per locus varied from 1.384 to 1.489 with an overall mean of 1.438. Information Index had a grand mean 0.409, observed heterozygosity

was 0.021, expected heterozygosity was 0.266, unbiased expected heterozygosity had a grand mean of 0.274 and Fixation index had a grand mean of 0.893 and these results are summarized in Table 5.2.

Table 5.3   Mean and standard error over loci for each population

Pop		N	N_a	N_e	I	H_o	H_e	uH_e	F
Pop1	Mean	17.394	1.912	1.489	0.443	0.007	0.291	0.300	0.948
	SE	0.127	0.024	0.029	0.019	0.002	0.014	0.015	0.016
Pop2	Mean	16.460	1.891	1.441	0.416	0.038	0.269	0.278	0.842
	SE	0.151	0.027	0.028	0.019	0.005	0.014	0.014	0.020
Pop3	Mean	15.175	1.803	1.384	0.367	0.018	0.236	0.245	0.889
	SE	0.148	0.034	0.029	0.020	0.003	0.015	0.015	0.021
Total	Mean	16.343	1.869	1.438	0.409	0.021	0.266	0.274	0.893
	SE	0.093	0.017	0.017	0.011	0.002	0.008	0.009	0.011

5.5.2   Polymorphism of SNP markers

Table 5.4 summarises information on SNP marker polymorphism. Results indicated that 91.24% of SNP markers (125 assays out of 137) in population 1 were polymorphic while 9% (12 assays out of 137) were monomorphic. In population 2, 89.05% of marker alleles (122 assays out of 137) were polymorphic while 10.95% (15 assays out of 137) were monomorphic. Similarly, in population 3, 80.29% (110 assays out 137) were polymorphic while 19.71% (27 assays out 137) were monomorphic. The grand mean across population was 86.86%, meaning that 119 assays out of 137 were polymorphic while 18 assays out of 137 were monomorphic.

Table 5.4   Percentage of polymorphic loci

Population	%PL
Pop1	91.24%
Pop2	89.05%
Pop3	80.29%
Mean	86.86%
SE	3.34%

PL= Polymorphic loci.

5.5.3   Analysis of molecular variance (AMOVA)

A summary statistics for each marker (locus) including heterozygosity which measures allelic diversity at a locus, is presented in Table 5.5. there was no variance observed among populations (0%). The variance among individuals was significantly high (94%). The variance within individuals was low (6%).

Table 5.5   Summary of analysis of molecular variance (AMOVA)

Source	df	SS	MS	Est. Var.	%variation	F-Statistics
Among Pops	2	94.112	47.056	0.026	0%	0.381
Among Indiv	49	2260.965	46.142	22.388	94%	0.001
Within Indiv	52	71.000	1.365	1.365	6%	0.001
Total	103	2426.077		23.780	100%	

DF= degree of freedom, SS= sum of squares, MS= mean sum of squares, Est. var. = estimated variance, Per. Var. = percentage variation

5.5.4   Allelic patterns across populations

Figure 5-2 summarises the mean allelic patterns across the populations. The mean allelic in populations were 1.912, 1.891 and 1.803 for population 1, 2 and 3, respectively.
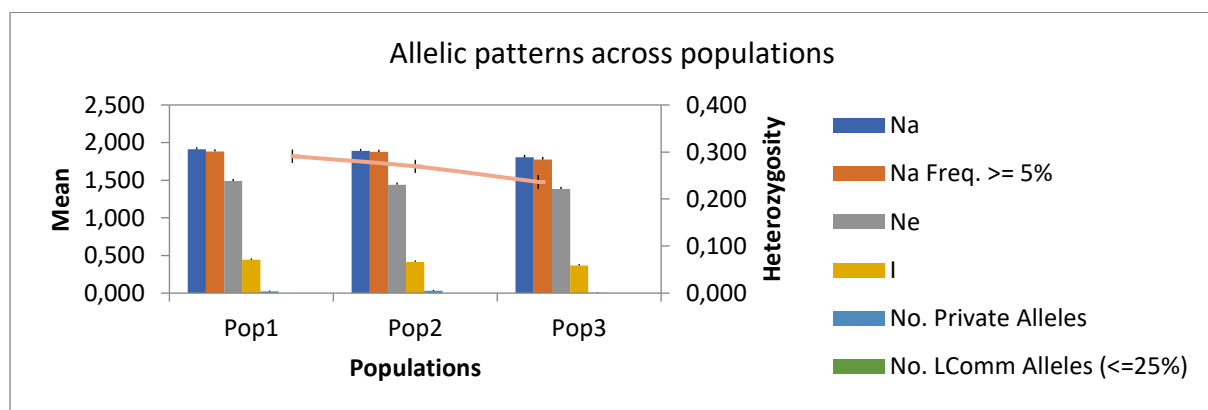

Figure 5-2   Allelic patterns across populations

5.5.5 Genetic distance and cluster analysis

The lowest dissimilarity index (Table not presented) observed was 0.00 between SA 4220 and SA 2012, SA 2013 and SA 4477, SA 4510 and SA 2013 amongst others. The highest dissimilarity index of 0.5833 was observed between SA 2193 and SA 2014. The overall mean for dissimilarity matrix was 0.296.

Cluster analysis discriminated the 52 sweet sorghum accessions into four groups (populations) designated I, II, III, and IV (Figure 5-3). However, group IV consisted of only one member and this group was added onto group III, and in total three groups of genotypes Group I (18 individuals), Group II (18 individuals) and Group III (16 individuals) were analysed for within and between group diversity. The branches in the cluster are graphical estimates of genetic distance between the accessions, which ultimately indicates genetic relationships between these sweet sorghum accessions.
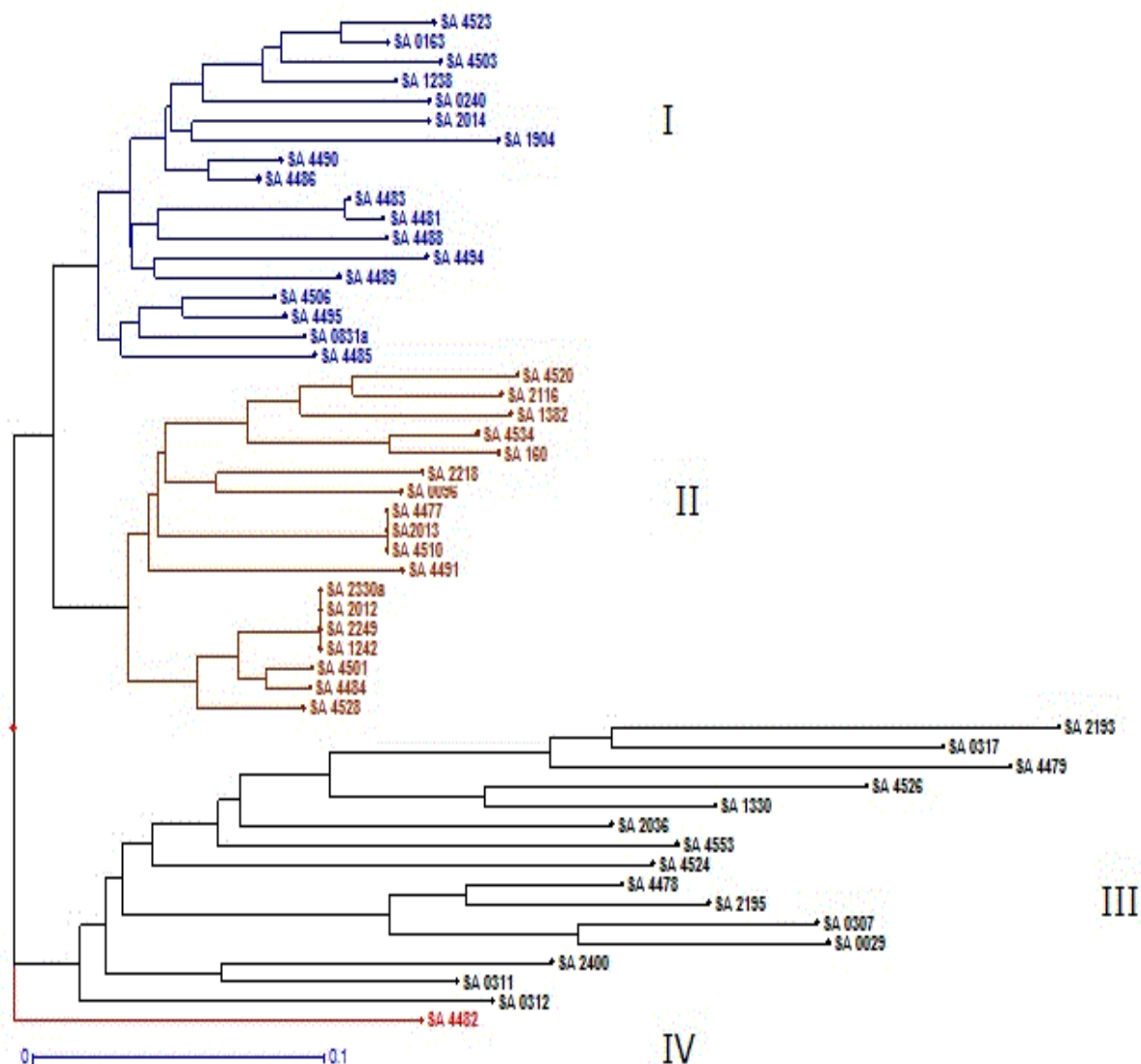
Figure 5-3   SNP cluster analysis showing genetic relationship among sweet sorghum accessions discriminated by different colours

5.6   Discussion

The use of molecular markers in plant breeding allows plant breeders and geneticists to locate and understand the basics of several gene interactions, which determine complex traits. It also provides information on genetic relatedness amongst accessions and this information is convenient in crop improvement programmes for management and utilization of germplasm (Haussmann et al., 2000). Genetic diversity on the other hand provides a platform for selection

of parental material to be used in breeding programmes. Genetic diversity also accelerates progress in crop improvement programmes.

In this study expected heterozygosity (H_e) ranged from 0.236 to 0.291 with an overall mean of 0.266. This overall mean is greater than the mean value reported for Tongkatali (*Eurycoma longifolia*), 0.216 by Osman et al. (2003), However, the mean value in this study is less than the mean H_e of 0.537 reported by Zulkifli. et al. (2008) based on simple sequence repeats (SSRs). The higher H_e detected by SSR markers is possibly due to the multi-allelic nature of the SSR markers. However, these results support the use of SNP markers in genetic diversity studies for sweet sorghum accessions. The mean values for observed heterozygosity (H_o) 0.021 was lower than the expected heterozygosity (H_e) 0.266). However, the mean H_o in this study is higher than the mean H_o (0.018) reported by Galeano et al. (2012). The H_o mean value was also lower than the fixation index (F) 0.893. These results concur with previous reports by Osman et al. (2003), Zulkifli. et al. (2008) and Islam et al. (2015). Low H_O and a very high F in this study implies that majority of the accessions were different and homozygous. This is sustained through selfing. A collection of diverse homozygous individuals can give rise to a heterogenious population which is a suitable gene pool for selection for hybridisation towards sweet sorghum improvement programme.

Singh et al. (2013) and Mvuyekure et al. (2018) reported that information on polymorphism frequencies is an important tool in assessing the value of the marker for germplasm characterisation. In this study therefore, a total 119 markers provided informative polymorphism in genetic diversity assessment for the sweet sorghum accessions under study.

The analysis further showed that no common alleles with a frequency of 25% or 50% were observed. In a study by de Oliveira et al. (2010), the authors reported that a rare allele often causes bias in covariance between markers and the population structure. This increases chances of committing a type I error in marker-trait association. The suitable population structure, without rare alleles, coupled with the use of a combination of randomly selected and candidate gene SNP markers, offers a favourable platform for marker-trait association, irrespective of a relatively small SNP panel (Ong et al., 2015).

Analysis of molecular variance (AMOVA) revealed a high significant variation among individuals, which contributed to 96% of the total variation. Wright (1965) reported that the F-statistics measures the extent of genetic variability among and within populations. The highly significant F-statistics among individuals in this study therefore indicate genetic diversity among sweet sorghum accessions.

SNP markers were able to discriminate sweet sorghum accessions into three main cluster groups. Group IV contained one accessions therefore, it was added to group III to allow AMOVA to proceed in GenAlEx. These accessions might have been clustered based on geographical origin or because they share some common ancestors in their pedigree information. Ali et al. (2007) also support this claim.

The lowest dissimilarity index of 0 observed among some accessions may suggest that these accessions are duplicates. These accessions have a very low genetic base hence not suitable for selection for breeding programme because progress in breeding programme depends largely on the extent of genetic diversity existing in a population (Zou et al., 2011). The highest dissimilarity index of 0.5833 was observed between SA 2193 and SA 2014.

Accessions with a low similarity mean indicate a considerable level of genetic variability. If selected for hybridisation programme, these may give a broad spectrum of variability in segregating generation, according to a report by Rohman et al. (2004). Accession SA 2193 and SA 2014 recorded the highest dissimilarity index. Therefore this implies a considerable genetic base which is an ideal platform for parental selection. If selected as parents for hybridisation programme, these may give a broad spectrum of variability in segregating generation (Rohman et al., 2004).

5.7 Conclusion

There is a considerable level of genetic diversity in the collection of accessions used in this study. Knowledge of the genetic diversity of sweet sorghum is key in identifying groups of similar genotypes for the purposes of conservation and proper use of genetic resources, in addition to the protection of property rights. SNP analysis is ideal for genetic diversity study analysis and marker assisted selection.

References

Akhunov, E., C. Nicolet and J. Dvorak. 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina Golden Gate assay. Theoretical and Applied Genetics 119: 507-517.

Ali, M.L., J.F. Rajewski, P.S. Baenziger, K.S. Gill, K.M. Eskridge, and L. Dweikat 2007. Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Molecular Breeding 21: 497-509.

Ashikari, M. and M. Matsuoka. 2006. Identification, isolation and pyramiding of quantitative trait loci for rice breeding. Trends in Plant Science 11: 344-350.

Balole, T.V. and G.M. Legwaila. 2005. *Sorghum bicolor* (L.) Moench.) PROTA Wageningen, Netherlands.

Bhattramakki, D., M. Dolan and M. Hanafey. 2002. Insertion-deletion polymorphisms in 3 regions of maize genes occur frequently and can be used as highly informative genetic markers. Plant Molecular Biology 48: 539-547.

Botstein, D., R.L. White, M. Skolnick and R.W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics 32: 314-331.

Buckler, E.S., J.B. Holland and P.J. Bradbury. 2009. The genetic architecture of maize flowering time. Science 325 714-718.

Cox, T.S., Y.T. Kiang, M.B. Gorman and D.M. Rodgers. 1985. Relationship between coefficient of parentage and genetic similarity indices in the soybean. Crop Science 25 529-532.

de Oliveira, B.T.C., Brondani RPV, F. Breseghello and A.S.G. Coelho. 2010. Association mapping for yield and grain quality traits in rice (*Oryza sativa* L.). Genetics and Molecular Biology 133: 515-524.

Franco, J., J. Crossa, J.M. Ribaout, J. Betran, M.L. Warburton and M. Khairallah. 2001 A method for combining molecular markers and phenotype attributes for classifying Plant Genotypes. Theoretical and Applied Genetics 103: 944-952.

Galeano, C.H., A.J. Cortés, A.C. Fernández, Á. Soler, N. Franco-Herrera, G. Makunde, J. Vanderleyden and M.W. Blair. 2012. Gene-based single nucleotide polymorphism markers for genetic and association mapping in common bean. Biomedical Central Genetics 13: p 48.

Graves, H., A.L. Rayburn, J.L. Gonzalez-Hernandez and G. Nah. 2016. Validating DNA polymorphism using KASP assay in prairie cordgrass (*Spartina pectinata* Link) populations in the U.S. Frontiers in Plant Science 6: pp 1271.

Haussmann, B.I.G., A.B. Obilana, P.O. Ayiecho, A. Blum, W. Schipprack and H.H. Gieger. 2000. Yield and yield stability of four population types of grain sorghum in a semi-arid area of Kenya. Crop Science 40: 319-329.

Islam, M.S., G.N. Thyssen, J.N. Jenkins and D.D. Fang. 2015. Detection, validation and application of genotype-by-sequencing based single nucleotide polymorphism in upland cotton. The Plant Genome 8: 1-10.

Jacob, H.J., K. Lindpaintner and S.E. Lincoln. 1991. Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. Cellular and Molecular Biology Letters 67: 213-224.

Jena, K.K. and D.J. Mackill. 2008. Molecular markers and their use in marker-assisted selection in rice. Crop Science 48: 1266-1276.

Jiangfeng, H.E., Z. Xiaoqing, A. Laroche, L. Zhen-Xiang, L. Hongkui and L. Ziqin. 2014. Genotyping by sequence (GBS), an ultimate marker assisted selection (MAS) tool to accelerate plant breeding. Plant Science 5: 1-8.

Karp, A., S. Kresovich, K.V. Bhatt, W.G. Ayad and T. Hodgkin. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies IPGR technical bulletins International Plant Genetics Resources Institute, Rome, Italy.

Mammadov, J., R. Aggarwal, R. Buyyarapu and S. Kumpatla. 2012. SNP Markers and their impact on plant breeding. International Journal of Plant Genomics 2012: 388-413.

Murray, S.C., W.L. Rooney, M.T. Hamblin, S.E. Mitchell and S. Kresovich. 2009 Sweet sorghum genetic diversity and association mapping for brix and height. The Plant Genome 2: 48-62.

Mvuyekure, S.M., J. Sibiya, J. Derera, J. Nzungize and G. Nkima. 2018. Assessment of genetic diversity of rice based on SNP markers for selection of parents for sheath rot (*Sarocladium oryzae*) resistance breeding. South African Journal of Plant and Soil 35: 51-59.

Nei, M. and W.H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences 76: 5269-5273.

Ngugi, K. and C.M. Onyango. 2012. Analysis of the molecular diversity of Kenyan sorghum germplasm using microsatellites. Crop Science Biotechnology 15: 189-194.

Ong, P.W., I. Maizura, N.A.P. Abdullah, M.Y. Rafii, L.C. Ooi, and E.T.L. Low. 2015. Development of SNP markers and their application for genetic diversity analysis in the oil palm (*Elaeis guineensis*). Genetics and Molecular Research 14: 12205-12216.

Osman, A., B. Jordan, P.A. Lessard and N. Muhammad. 2003. Genetic diversity of *Eurycoma longifolia* inferred from single nucleotide polymorphisms. Plant Physiology 131: 1294-1301.

Peakall, R. and P.E. Smouse. 2006. GenAlEx 6: Genetic Analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Resources 6: 288-295.

Peakall, R. and P.E. Smouse. 2012. GenAlEx 6.5: Genetic Analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539.

Perrier, X., Jacquemoud-Collet, J.P. 2006. DARwin software http://darwin.cirad.fr/darwin.

Prasad, S., A. Singh, N. Jain and H.C. Joshi. 2007. Ethanol production from sweet sorghum syrup for utilization as automotive fuel in India, Energy & Fuels. Plant Science 21: 2415-2420.

Rohman, M.M., M.A. Hakim, N.A. Sultana, M.E. Kabir, M.H. Hasanuzzan and M. Ali. 2004. Genetic divergence analysis in sorghum (*Sorghum bicolor* L.). Asian Journal of Plant Sciences 3: 211-214.

Semagn, K., A. Bjornstad and M.N. Ndjiondjop. 2006. An overview of molecular marker methods for plants. African Journal of Biotechnology 5: 2540-2568.

Singh, N., D.R. Choudhury, A.K. Singh, S. Kumar, K. Srinivasan, R. Tyagi, N.K. Singh and R. Singh. 2013. Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. PLoS One 8: e84136.

Wang, D.G., J.B. Fan and C.J. Siao. 1998. Large-scale identification, mapping, and genotyping of single nucleotide polymorphisms in the human genome. Science 280: 1077-1082.

Welsh, J. and M. McClelland. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18: 7213-7218.

Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19: 395-420.

Yu, H., W. Xie and J. Wang. 2011. Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. PLoS ONE 6.

Zou, G., S. Yan, G. Zhai, Z. Zhang, J. Zou and Y. Tao. 2011. Genetic variability and correlation of stalk yield-related and sugar concentration of stalk juice in a sweet sorghum (*Sorghum bicolor L.* Moench) population. Australian Journal of Crop Science 5: 1232-1238.

Zulkifli., Y., I. Maizura and R. Singh. 2008. Molecular screening of oil palm germplasm collections (*Elaeis guineensis*) using EST-SSR markers. Proceedings of the 3rd Seminar on Performance of MPOB PS1 and PS2 Materials and Elite Germplasm. Malaysian Palm Oil Board pp 168-185.

https://www.integratedbreeding.net/attachment/471/kasp_explanation.pdf

CHAPTER 6

Research overview

6.1 Introduction

Sweet sorghum has the potential to become a reliable biofuel crop because of high sugar accumulation in its juicy stalk, low production input requirements, rapid growth rate and effective conversion of atmospheric carbon dioxide into sugar (compared to maize and sugar cane), amongst others. Sweet sorghum research has intensified to meet the increased demand for ethanol, driven by the need to blend ethanol with fossil fuel.

6.2 Research objectives

Success in breeding programme is dependent on availability of genetic diversity. Therefore, knowledge of genetic diversity in sweet sorghum is very significant for breeders to develop varieties, which can also contribute to widening of the genetic base of sweet sorghum during parental selection. Therefore, the specific objectives of this research were:

- To estimate the level of variability among sweet sorghum germplasm using agro-morphological traits.
- To estimate genotypic and non-genotypic variance components and to determine correlations and path analysis of juice yield and juice yield-related traits.
- To assess the extent of genetic diversity using single nucleotide polymorphism (SNP) markers.

6.3 Summary of the research findings

Fifty sweet sorghum accessions were phenotyped under two environments (Potchefstroom, North West province Ukulinga Research farm in Pietermaritzburg, in KwaZulu-Natal Province), during the 2016-2017 growing season.

6.3.1 Assessment of genetic variability using agro-morphological traits under two environments

- ANOVA revealed high significant differences among the accessions for most of the traits across sites.
- Mean performance results of traits studied across the two sites revealed that sweet sorghum accessions performed better at Ukulinga than Potchefstroom.

- Four principal components contributed 42.08%, 19.12%, 10.28% and 7.63% of the total variation. Much of this was contributed by traits such as days to 50% flowering, plant height, stalk diameter, number of leaves per plant, stalk yield, brix, juice volume and fresh bagasse weight.
- The study identified SA4490, SA2400, SA4495, SA2193 and SA4479 as superior accessions in juice yield.

6.3.2 Genetic and path coefficient analysis of juice yield and juice related traits in sweet sorghum

- High heritability was observed in plant height, stalk yield, days to 50% flowering and juice volume.
- High GAM values were observed for juice volume, bagasse yield, stalk yield and %brix.
- There were positive and significant correlations between plant height and biomass yield, biomass yield and juice volume and bagasse.
- Bagasse weight, %brix, stalk diameter, plant height and number of leaves had a highly positive and direct contribution on juice yield.
- Several traits had a highly positively and indirect contribution on juice yield via these traits which had a direct contribution.

6.3.3 Assessment of genetic diversity using single nucleotide polymorphism (SNP) markers

- The grand mean for polymorphism of SNP markers across populations was 86.86% (119 assays out of a total of 137 were polymorphic) while 13.14% of the markers (18 assays out of 137) were monomorphic.
- The observed heterozygosity values ranged from 0.007 to 0.038 with a grand mean of 0.021 and its estimation (H_e) values ranged from 0.236 to 0.291 with an overall mean of 0.266.
- The observed heterozygosity ranged from 0.002 to 0.038 with a grand mean of 0.021
- The mean of effective alleles across populations was of 1.438.
- The percentage of polymorphic loci ranged from 80.29% to 91.24% with a mean of 86.86%. Dissimilarity indices ranged from 0.000 to 0.5833 with a mean of 0.296.
- The highest dissimilarity index was observed between SA 2193 and SA 2014 which implied a considerable amount of genetic diversity.
- Accessions were clustered into three main groups based on dissimilarity indices.

6.4 Implications of the research findings on breeding

There is an opportunity for parent selection among these accessions for increased juice yield production and ultimately increased sugar production per hectare due to the availability of genetic diversity. High heritability and high genetic advance for most of the traits showed the presence of additive genes in the traits and this suggests reliable sweet sorghum improvement through selection of the traits. Significant G x E interaction for the traits across sites implies that selection for each location must be done independently. Positive and significant correlations among juice related traits suggest simultaneous improvement of juice and other traits through selection. A highly positively and indirect contribution on juice yield by some traits via major traits which had a direct contribution, implies that indirect selection can be directed on these traits to improve juice yield. The majority of accessions in this study were different and there was a significant level of homozygosity.

6.5 Recommendations

Sweet sorghum accessions having the highest dissimilarity index should be considered in the selection and development of high juice yielding cultivars which are adapted local environments. Sweet sorghum improvement programme should aim at developing cultivars that are early to medium maturing, with tall and thick stalks, which are resistant to lodging. Research should also focus on identifying accessions which have a high juice extraction ratio, this will ensure development of high juice yielding sweet sorghum cultivars which retain little juice in the bagasse. There is a need for further research on genotype by sowing date interaction (under South African environment), for juice yield and %brix traits. This will identify high yielding and stable or specifically adapted accessions for cultivation across the year, which shall ensure a continuous supply of the feedstock to the distillery.