Assessment of Sweet Sorghum Lines for Genetic Diversity Using Quantitative Traits and SSR Markers

By

Sombo Chinyama

BSc. Agricultural Science (University of Zambia)

A dissertation submitted in partial fulfilment of the requirements for the degree of Masters (MSc.) in Plant Breeding

School of Agricultural Sciences, Earth and Environmental Sciences
Faculty of Science and Agriculture
University of KwaZulu-Natal
Republic of South Africa

November 2016

General Abstract

The increase in energy demand, volatile oil prices and climate change has led South Africa to reduce its dependency on fossil fuels and promote biofuels. Sweet sorghum [*Sorghum bicolor* (L.) Moench] has been considered as one of the promising crops due to its sugar-rich stalk to supplement sugarcane which is the major feedstock for bioethanol. Establishing genotypic variability for biomass yield and sugar-related traits in sweet sorghum is therefore essential for developing superior cultivars. The objectives of the study were: (i) to assess sweet sorghum lines for agronomic performance and genetic diversity using quantitative morphological traits and (ii) to assess sweet sorghum lines for genetic diversity and interrelationships using simple sequence repeat (SSR) markers.

Twenty-five sweet sorghum lines collected from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and the African Centre for Crop Improvement (ACCI-South Africa) were evaluated during the 2015/2016 season in KwaZulu-Natal (KZN) province at Ukulinga Research Farm in Pietermaritzburg and Makhathini Research Station in Jozini. Seven agronomic traits; fresh biomass yield, fresh stalk yield, grain yield, plant height, stalk diameter, panicle length and days to 50% flowering, and six quality traits; fibre, dry matter, °brix, °total brix, total fermentable sugars and ethanol were recorded. The sweet sorghum lines revealed highly significant variations for the 13 quantitative characters assessed in this study. The extent of variation was highly influenced by environment and genotype by environment interaction. Genotypes designated as IS 2331, IESV 92008 DL, ICSV 700, AS 244, URJA and SS 27 were identified as suitable genotypes with high plant height, dry matter, fibre, °brix, °total brix, total fermentable sugars and ethanol.

The specified genotypes also exhibited medium to late maturity with relatively high fresh biomass and fresh stalk yield. Genotype 91018 LT showed the highest fresh biomass yield, fresh stalk yield, stalk diameter and relatively high grain yield. High levels of trait heritability were observed for fresh stalk yield (98%), stalk diameter (93%), fresh biomass yield (81%), panicle length (76%), fibre (73%) and plant height (66%). Heritability estimates were influenced by the environment and genotype by environment interaction. Principal component analysis resulted in the first three principal components showing 83% of the total variability among the genotypes. Ethanol, total fermentable sugars, °total brix, fresh stalk yield and °brix contributed mainly to PC 1, whereas fresh biomass yield and stalk diameter contributed mainly to PC 2. The dendrogram generated from cluster analysis divided the genotypes into two main clusters

and three singletons (ICSB 324, ICSB 654 and ICSV 700). Cluster I comprised 54% of the total germplasm and included only one ACCI genotype (SS 27), while cluster II comprised of 33% of the total variation.

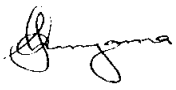
The morphological variability analysis of the genotypes was also complimented with the use of molecular markers. The 24 sweet sorghum lines were genotyped with 10 simple sequence repeat (SSR) markers and distance-based method was used to analyze the data. Variation was observed for all the markers with allelic size ranging from 1 to 36 bp. A total of 61 alleles were generated with an average of 6.1 alleles per locus. The polymorphism information content (PIC) values ranged from 0.32 to 0.86 with an overall mean value of 0.62, showing a high discriminating ability of the markers used. The largest genetic distance was observed for AS 244 (GD = 1.9), while IESV 92001 DL and IESV 92008 had the smallest genetic distance (GD = 0.50). The dendrogram generated from cluster analysis using SSR markers classified the 24 sweet sorghum lines into two major clusters. Cluster I comprised of 12.5% of the total genotypes which included URJA, SS 27 and ICSB 654. It was observed that all ACCI genotypes apart from AS 244 were grouped in Cluster I, with URJA and SS 27 being very closely related. Cluster II was observed to be the largest (87.5%) with 21 genotypes, which further formed 3 sub-clusters (A, B and C) and a singleton (AS 244). The results from molecular marker characterization were similar to those obtained using PCA analysis of morphological traits which grouped genotypes into four clusters, with the same type of genotypes in each group. The information obtained in this study coupled with phenotypic characterization can be used by plant breeders to select parents or pure lines that can be used in developing improved cultivars. This will therefore contribute to the production of sweet sorghum and promotion of its use for bioethanol in South Africa.

Declaration

I, Sombo Chinyama, declare that:

1. The research reported in this dissertation, except where otherwise indicated, is my original research.
2. This dissertation has not been submitted for any degree or examination at any other university.
3. This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This dissertation 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 quotation marks, and referenced.
5. 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 thesis and in the references sections.

Signed



Sombo Chinyama

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

……
Dr. Julia Sibiya (Supervisor)

……..……..
Dr. Alfred Odindo (Co-Supervisor)

Acknowledgements

I owe my deepest gratitude to Dr. Julia Sibiya and Dr. Alfred Odindo who have been the ideal project supervisors. Their sage advice, insightful criticisms and patient encouragements enabled me to develop a clear understanding of the subject. I am also heartily thankful to my research mentor, Prof. Marvellous Zhou for his guidance and steadfast support throughout the dissertation write up.

It is also with immense gratitude that I acknowledge the following organizations and people for making my research a success in innumerable ways;

- Alliance for a Green Revolution in Africa (AGRA), for funding this study.
- ICRISAT-Kenya, for providing the sweet sorghum germplasm used in the trials.
- All the academic staff under the School of Agriculture, Earth and Environmental Sciences, University of Kwa-Zulu Natal (UKZN). The high quality of this work was made possible through their individual contribution.
- The former administrator of the MSc. Plant Breeding Project, Jayshree Singh. Her endless support and care enabled me to forge on even during the hardest times.
- The technical assistants from the School of Agriculture, Earth and Environmental Sciences, UKZN (Ukulinga Research Farm) and Makhathini Research Station - Jozini. These technicians ensured that all the fields were well managed for successful data collection. I also offer my blessings to all those people who supported me in any aspect during the completion of my dissertation.

I am truly grateful for having such a supporting family. My loving husband, Chazya Musukuma, enabled me to believe in myself and fulfil my dreams. Lastly, I would like to give gratitude to God for giving me the wisdom and strength to forge on throughout my two year study.

Dedication

This dissertation is dedicated to my husband, Mr. Chazya Musukuma, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother (Mercy Kambi-Kambi), sister (Kasweka Chinyama Sinyinza) and all my friends who taught me that even the largest task can be accomplished if it is done one step at a time.

Table of Contents

List of Figures

LIST OF TABLES

Introduction

1. Biofuels in South Africa

South Africa is known to be the largest emitter of greenhouse gases in Africa due to its energy intensive economy (Meyer et al., 2005). The increase in energy demand, volatile oil prices and climate change have led the country to reduce its dependency on fossil fuels and promote biofuels (Cilliers, 2012). Biofuels are known to be sustainable solutions to global energy crisis, global warming and global pollution (Demirbas, 2009). The use of biofuels in the country will not only boost the socio-economic sectors, but will particularly create job opportunities for the rural community which is a priority for the South African government. Despite government intervention to provide technology and infrastructure needed in the biofuel industry, there is currently no medium or large scale biofuel producer in operation (Cilliers, 2012). Bioethanol is the most well-known biofuel and is an alcohol fuel produced by fermentation of sugars found in traditional food crops such as sugarcane, maize, sugar beet, sorghum and cassava (Johnson and Matsika, 2006).

2. Importance of sweet sorghum

Sweet sorghum [*Sorghum bicolor* (L.) Moench] is a multi-purpose crop species that provides grain and stalk for feed, fodder, fuel, paper making and fencing (Janssen et al., 2010; Kumar et al., 2010). The species has tremendous potential in the tropics and sub-tropics for sustainable production of first generation biofuels (Reddy et al., 2005). This is due to its high soluble sugars, cellulose and hemicellulose content in the stalks (Antonopoulou et al., 2008; de Vries et al., 2010). The soluble sugars are comprised of 70% sucrose, which is used for refining crystal sugar (Almodares and Hadi, 2009) and the sugar yields range between 1.6 to 13.2 t ha^{-1} (Zhao et al., 2009). However, there is significant variation due to the influence of the environment on the relative performance of genotypes (Zhao et al., 2009).

Variation in sugar content is known to exist along and across the stalk during the stages of plant growth. Glucose is found at higher concentrations in young and elongating internodes while in ripening internodes, sucrose has been observed to be dominant (Tarpley and Vietor, 2007). At physiological maturity, sweet sorghum consists of approximately 10% roots, 75% stalk, 10% leaves and 5% seeds by weight (Mutepe et al., 2012), with the stalk juice having about 10 to 25% sugar (Reddy et al., 2007). After harvest, the stalks are squeezed for the

sweet juice and can be turned into sugar or fermented to ethanol. The bagasse, which is the stalk material remaining after the sugar juice has been squeezed out, can be used as animal feed or pretreated, hydrolyzed and fermented to produce ethanol at low cost (Wang and Liu, 2009; Zhao et al., 2009).

3. Global production of sweet sorghum

The total production of sorghum ranks fifth among the most important cereal crops in the world and second in Africa after maize (Makanda, 2009; Ghani et al., 2015). Despite the importance of sweet sorghum in the bioethanol industry, there is currently no readily available information on total global production of sweet sorghum biomass yield. However, studies carried out by researchers in different countries and regions have shown potential in biomass yield production. Sweet sorghum cultivars developed specifically for stalk sugar production in Europe produced a stalk biomass yield from 50 to 140 t ha^{-1} (Claassen et al., 2004). Whereas, in India, sweet sorghum stalk biomass yield during *kharif* (rain) season can vary from 30 to 50 t ha^{-1} (Rao et al., 2013). On the other hand, research done by Makanda et al. (2009) on sweet sorghum varieties at Ukulinga Research Farm in South Africa, recorded a stalk biomass yield ranging between 3 to 50 t ha^{-1} (Figure 1-0). In Southern Africa (South Africa, Zimbabwe and Mozambique), research done on stalk sugar traits on dual purpose sorghum germplasm showed that stalk biomass yield can range from 8 to 76 t ha^{-1} (Makanda, 2009; Makanda et al., 2011). Generally low production of sweet sorghum is related to many factors such as lack of high yielding genotypes adapted to abiotic and biotic stresses (Reddy et al., 2006; Rao et al., 2009). Although South African yields are lower than those documented above, the introduction and production of improved sweet sorghum cultivars can boost the biofuel industry in the country.

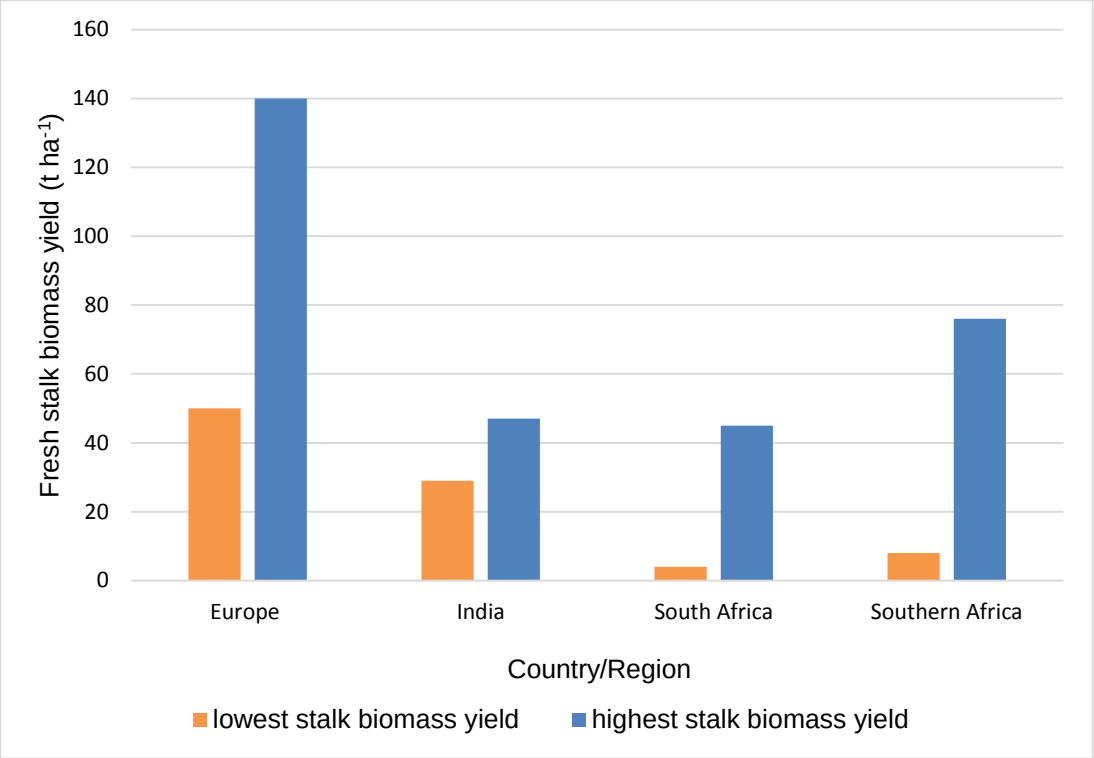


Figure 1-0: Sweet sorghum stalk biomass yield (t ha^{-1}) per country and region

4. Biomass yield and sugar-related traits

Phenotypic selection is one of the most successfully used method in plant breeding, especially for highly heritable traits. One of the many objectives of sweet sorghum breeding is to develop superior cultivars with high biomass and sugar-related traits. However, secondary traits have a major influence on response to selection (Ilker, 2011). Therefore, proper knowledge of the relationships among quantitative traits is important for assessing the feasibility of joint selection of two or more traits (Ezeaku and Mohammed, 2006). In sweet sorghum, performance of stalk biomass yield and sugar-related traits such as stalk sugar content and plant height are known to be affected by a number factors including the environment (Boćanski et al., 2009; Elangovan et al., 2014). Evaluating genotype performance across different locations is therefore essential when exploiting existing variability and development of improved cultivars (Faisal and Aisha, 2011; Abubakar and Bubuche, 2013). This is evidenced by several studies that were carried out to characterize the interaction pattern of stalk biomass yield and its components in different locations. The findings confirmed the presence of a significant interaction between the genotype and the environment as a consequence of the differential response of the genotypes to environmental changes (Makanda et al., 2009; Zou et al., 2011; Elangovan et al., 2014).

5. Genetic diversity in sweet sorghum

Increase in yield and desirable traits in different crops is practically impossible without genetic variability in the germplasm (Cholastova et al., 2013). Genetic diversity assessments are therefore important as they enable plant breeders to explore different techniques to ensure germplasm enhancement through the availability of genetic variability (Makumbi et al., 2011; Cholastova et al., 2013). It is also emphasized by Govindaraj et al. (2015) that the introgression of desirable traits and elimination of deleterious genes can only be achieved by genetic manipulation. Knowledge of genetic diversity is also essential in selecting elite parents from heterotic groups that are formed by the use of genetic distance (Legesse et al., 2008). Sweet sorghum cultivars with high biomass yield and sugar-related traits can be developed with this background and enhance the South African bioethanol industry.

Genetic diversity can be estimated based on morphological, biochemical and molecular markers (Mehmood et al., 2008; Amelework et al., 2015). Plant breeders have commonly used morphological traits for many years because these traits provide the simplest way of measuring genetic diversity while studying the performance of genotypes under normal growing conditions. However, this approach requires significant time and is unreliable because it is largely influenced by genotype by environment interaction (Pecina−Quintero et al., 2012; Govindaraj et al., 2015). Valuable alleles are also masked by negative alleles at loci and there is unknown genetic control of poly genetically inherited morphological traits (Assar et al., 2005; Geleta et al., 2005; Perumal et al., 2007).

Despite the low adoption rate of DNA molecular markers in developing countries due to high costs and technical demand (Ribaut et al., 2010), DNA-based molecular marker analysis has been reported to be the most reliable and stable method for assessment of genetic diversity. This is due to lack of environmental influences on the relative performance of genotypes(Turki et al., 2011), and analysis is carried out at any growth stage of the plant (Mehmood et al., 2008; Assar et al., 2009). Among the different DNA-based molecular markers, simple sequence repeats (SSRs) have commonly been used on sorghum to determine its genetic variability because of their ability to discriminate among closely related individuals and effectively identifying heterotic groups (Mofokeng et al., 2014; Olweny et al., 2014; Amelework et al., 2015). For effective selection in genetic diversity studies, a combination of morphological and molecular-marker tools are mostly used by plant breeders (Barata and Carena, 2006; Vieira et al., 2007), and this has been adopted for this study.

6. Genetic diversity grouping techniques

Multivariate grouping techniques are used to analyze genetic variability among breeding materials irrespective of the data (Bertan et al., 2007; Aremu, 2012). These methods include, principal component analysis (PCA), cluster analysis, principal coordinate analysis (PCoA), canonical correlation and multidimensional scaling (Aremu, 2012). Principal component analysis (PCA) and clustering analysis (dendrogram) are the most used grouping techniques in genetic diversity studies (Mohammadi and Prasanna 2003). Principal component analysis produces a 2 or 3 dimensional scatter plot that enables visualization of the differences among the samples (Mohammadi and Prasanna 2003). On the other hand, cluster analysis, differs from PCA in that it depends upon previous measurements of genetic distance, which rely on the overall distance of Euclidean and the Mahalanobis (Jolliffe, 1973; Ganapathy et al., 2012). The different techniques group genetically similar individuals together in clusters using different algorithms such as; centroids (UPGMA and UPGMC), complete linkage (CLCA), median linkage (MLCA) and single linkages (Aremu, 2011; Kubie, 2013). The most popularly used is the unweighted pair group method with arithmetic averages (UPGMA) as it provides more accuracy by discriminating closely related genotypes in various crops (Bertan et al., 2007; Trindade et al., 2010).

7. Rationale of the study

South Africa's energy intensive economy has led to government intervention to promote biofuel production. Sweet sorghum has therefore been considered as a crop for biofuel production mainly because of the following reasons: (i) using maize and grain sorghum could contribute to threats to food security, especially among smallholder farmers and (ii) the use of sugarcane is relatively costly in relation to the heavy investment costs in terms of management, and the area suitable for cultivation is confined to frost free areas with high rainfall. The use of sweet sorghum for biofuel production could be advantageous because of a shorter growing period, adaptation to dry environments which requires less water, and high fermentable sugar content (Tsuchihashi and Goto, 2008).

Even with all these benefits, there is limited information on the availability of sweet sorghum lines that have high stalk biomass and sugar-related traits in South Africa, mainly due to: (i) poor conservation and characterization of sweet sorghum lines; (ii) poor selection of appropriate parental lines for hybrid development; and (iii) unknown performance of these

genotypes across different agro-ecological zones. The study was therefore undertaken to evaluate genetic variability among a collection of sweet sorghum lines using morphological data and SSR markers. This was done to identify specific genotypes exhibiting high levels of biomass yield and sugar-related traits from various heterotic groups that will improve the genetic base and characterization of the current cultivated sweet sorghums (Singh and Singh, 2015).

8. Aims and objectives

The overall aim of this study was to determine the genetic diversity of sweet sorghum germplasm obtained from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and the African Centre for Crop Improvement (ACCI-South Africa). The generated information will contribute to the development of superior varieties that will be a potential source of feedstock for bioethanol production in South Africa.

The specific objectives were:

1. To assess sweet sorghum lines for agronomic performance and genetic diversity using quantitative morphological traits.
2. To assess sweet sorghum lines for genetic diversity and interrelationships using simple sequence repeat (SSR) markers.

9. Research hypotheses

The following hypotheses were tested:

1. There is no difference in performance, genetic diversity for biomass yield and sugar-related traits among the collected sweet sorghum lines.
2. There is no genetic diversity or interrelationships for biomass yield and sugar-related traits among the collected sweet sorghum lines using simple sequence repeat markers.

10. Outline of dissertation

Chapter two and three are written in AIMRD format that include Abstract, Introduction, Materials and Methods, Results and Discussion. All chapters have been written as

independent journal papers with a reference list and contain some repetition and overlap between chapters. The chapters are divided as follows:

1. Introduction to dissertation
2. Chapter 1: Literature review
3. Chapter 2: Assessment of sweet sorghum lines for agronomic performance and genetic diversity using quantitative morphological traits.
4. Chapter 3: Assessment of sweet sorghum lines for genetic diversity and interrelationships using simple sequence repeat (SSR) markers.
5. Chapter 4: General overview

11. References

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

Almodares, A., and M. Hadi. 2009. Production of bioethanol from sweet sorghum: A review. African Journal of Agricultural Research 4:772-780.

Amelework, B., H. Shimelis, P. Tongoona, M. Laing, and F. Mengistu. 2015. Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. Plant Genetic Resources 13:131-141.

Antonopoulou, G., H.N. Gavala, I.V. Skiadas, K. Angelopoulos, and G. Lyberatos. 2008. Biofuels generation from sweet sorghum: fermentative hydrogen production and anaerobic digestion of the remaining biomass. Bioresource Technology 99:110-119.

Aremu, C. 2011. Genetic diversity: a review for need and measurements for intraspecies crop improvement. Journal Microbiology Biotechnology Research 1:80-85.

Aremu, C. 2012. Exploring statistical tools in measuring genetic diversity for crop improvement. INTECH Open Access Publisher.

Assar, A., R. Uptmoor, A. Abdelmula, M. Salih, F. Ordon, and W. Friedt. 2005. Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). Crop Science 45:1636-1644.

Assar, A.H.A., R. Uptmoor, A.A. Abdelmula, C. Wagner, M. Salih, A.M. Ali, et al. 2009. Assessment of sorghum genetic resources for genetic diversity and drought tolerance

using molecular markers and agro-morphological traits. University of Khartoum. Journal of Agricultural Science 17:1-22.

Barata, C., and M. Carena. 2006. Classification of North Dakota maize inbred lines into heterotic groups based on molecular and testcross data. Euphytica 151:339-349.

Bertan, I., F. Carvalho, and A.d. Oliveira. 2007. Parental selection strategies in plant breeding programs. Journal of Crop Science and Biotechnology 10:211-222.

Boćanski, J., Z. Srećkov, and A. Nastasić. 2009. Genetic and phenotypic relationship between grain yield and components of grain yield of maize (*Zea mays* L.). Genetika 41:145-154.

Cholastova, T., M. Soldanova, and R. Pokorny. 2013. Random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) marker efficacy for maize hybrid identification. African Journal of Biotechnology 10:4794-4801.

Cilliers, B.L. 2012. An industry analysis of the South African biofuels industry. PhD dissertation, North-West University, South Africa.

Claassen, P., T. De Vrije, M. Budde, E. Koukios, A. Glynos, and K. Réczey. 2004. Biological hydrogen production from sweet sorghum by thermophilic bacteria. Proceedings 2nd World Conference on Biomass for Energy, Rome. p. 1522-1525.

de Vries, S.C., G.W. van de Ven, M.K. van Ittersum, and K.E. Giller. 2010. Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion techniques. Biomass and Bioenergy 34:588-601.

Demirbas, A. 2009. Political, economic and environmental impacts of biofuels: A review. Applied Energy 86:S108-S117.

Elangovan, M., N. Seetharama, and J. Patil. 2014. Genetic diversity and heritability characters associated in sweet sorghum [*Sorghum bicolor* (L.) Moench]. Sugar Technology 16:200-210.

Ezeaku, I., and S. Mohammed. 2006. Character association and path analysis in grain sorghum. African Journal of Biotechnology 5:1337-1340.

Faisal, E., and O. Aisha. 2011. Genotype x seed production environment interaction on the performance of sorghum (*Sorghum bicolor* [L.] Moench) under irrigation. Agriculture and Biology Journal for North America 2:2151-7517.

Ganapathy, K., S. Gomashe, S. Rakshit, B. Prabhakar, S. Ambekar, R. Ghorade, et al. 2012. Genetic diversity revealed utility of SSR markers in classifying parental lines and elite

genotypes of sorghum (*Sorghum bicolor* L. Moench). Australian Journal of Crop Science 6:1486-1493.

Geleta, L., M. Labuschagne, and C. Viljoen. 2005. Genetic variability in pepper (*Capsicum annuum* L.) estimated by morphological data and amplified fragment length polymorphism markers. Biodiversity & Conservation 14:2361-2375.

Ghani, A., M. Saeed, D. Hussain, M.M. Shafique, M. Arshad, and S.A.S. Shah. 2015. Evaluation of different sorghum (*Sorghum bicolor* L. Moench) varieties for grain yield and related characteristics. Science Letters 3:72-74.

Govindaraj, M., M. Vetriventhan, and M. Srinivasan. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics Research International Journal 2015.

Ilker, E. 2011. Correlation and path coefficient analyses in sweet corn. Turkish Journal of Field Crops 16:105-107.

Janssen, R., D. Rutz, S. Braconnier, B. Reddy, S. Rao, R. Schaffert, et al. 2010. Sweet sorghum–an alternative energy crop. Journal of Environments 4:6.

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.

Jolliffe, I.T. 1973. Discarding variables in a principal component analysis. II: Real data. Applied Statistics 22: 21-31.

Kubie, L. 2013. Evaluation of Genetic Diversity in Garden Egg (*Solanum Aethiopicum*) Germplasm in Ghana. PhD dissertation, University of Ghana, Ghana.

Kumar, A.A., B. Reddy, M. Blümmel, S. Anandan, Y.R. Reddy, C. Raviender Reddy, et al. 2010. On-farm evaluation of elite sweet sorghum genotypes for grain and stover yields and fodder quality. Animal Nutrition and Feed Technology S 10:69-78.

Legesse, B., A.A. Myburg, K. Pixley, S. Twumasi-Afriyie, and A.-M. Botha. 2008. Relationship between hybrid performance and AFLP based genetic distance in highland maize inbred lines. Euphytica 162:313-323.

Makanda, I. 2009. Combining ability and heterosis for stem sugar traits and grain yield components in dual-purpose sorghum (*Sorghum bicolor* L. Moench) germplasm. PhD dissertation, University of KwaZulu-Natal, South Africa.

Makanda, I., J. Derera, P. Tongoona, and J. Sibiya. 2011. Development of sorghum for bio-energy: A view from the stakeholders and priorities for breeding dual purpose varieties. African Journal of Agricultural Research 6:4477-4486.

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.

Makumbi, D., J.F. Betrán, M. Bänziger, and J.-M. Ribaut. 2011. Combining ability, heterosis and genetic diversity in tropical maize *(Zea mays* L.) under stress and non-stress conditions. Euphytica 180:143-162.

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

Meyer, F., P. Strauss, M. Cutts, J. de Beer, G. du Toit, T. Funke, et al. 2005. Bio-ethanol production in South Africa: An objective analysis. http://www. bfap. co. za/pdfs/BFAP% 20report% 20on. (Accessed on 13 March, 2016).

Mofokeng, A., H. Shimelis, P. Tongoona, and M. Laing. 2014. A genetic diversity analysis of South African sorghum genotypes using SSR markers. South African Journal of Plant and Soil 31:145-152.

Mutepe, R.D., S. Marx, and P. Van der Gryp. 2012. Ethanol production from sweet sorghum. MSc thesis, North-West University, Potchefstroom Campus, South Africa.

Olweny, C., J. Jamoza, M. Dida, W. Kimani, J. Njuguna, D. Githae, et al. 2014. High genetic diversity for improvement of sweet sorghum (*Sorghum bicolor* (L.) Moench) genotypes for sugar and allied products. Molecular Plant Breeding 5:29-35.

Pecina−Quintero, V., J.L. Anaya−López, A. Zamarripa−Colmenero, N. Montes−García, C. Nuñez−Colín, J.L. Solis−Bonilla, et al. 2012. Genetic diversity of sweet sorghum germplasm in Mexico using AFLP and SSR markers. Pesquisa Agropecuária Brasileira 47:1095-1102.

Perumal, R., R. Krishnaramanujam, M.A. Menz, S. Katilé, J. Dahlberg, C.W. Magill, et al. 2007. Genetic diversity among sorghum races and working groups based on AFLPs and SSRs. Crop Science 47:1375-1383.

Rao, S., J. Patil, A. Umakanth, J. Mishra, C. Ratnavathi, G.S. Prasad, et al. 2013. Comparative performance of sweet sorghum hybrids and open pollinated varieties for millable stalk yield, biomass, sugar quality traits, grain yield and bioethanol production in tropical Indian condition. Sugar Technology 15:250-257.

Rao, S., S. Rao, N. Seetharama, A. Umakath, P.S. Reddy, B. Reddy, et al. 2009. Sweet sorghum for biofuel and strategies for its improvement. Information Bulletin No. 77.

International Crops Research Institute for the Semi-Arid Tropics, Patancheru - 502 324, A.P, India.

Reddy, B.V., A.A. Kumar, and W.D. Dar. 2006. Overview of sweet sorghum breeding at ICRISAT: Opportunities and constraints. International Crops Research Institute for the Semi-Arid Tropics, Patancheru - 502 324, A.P, India.

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

Reddy, B.V.S., 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.

Ribaut, J., M. De Vicente, and X. Delannay. 2010. Molecular breeding in developing countries: challenges and perspectives. Current Opinion in Plant Biology 13:213-218.

Ribeiro Trindade, A.P., R.J. Barth Pinto, A.T.d. Amaral Júnior, C.A. Mangolin, S. Machado, and C.A. Scapim. 2010. Genetic diversity of breeding popcorn lines determined by SSR markers. Electronic Journal of Biotechnology 13:4-5.

Singh, B., and A.K. Singh. 2015. Marker-assisted plant breeding: principles and practices. Springer, India.

Tarpley, L., and D.M. Vietor. 2007. Compartmentation of sucrose during radial transfer in mature sorghum culm. BMC Plant Biology 7:33-42.

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.

Turki, D.A., M. Al Jamali, and A. Kanbar. 2011. Genetic variation of sorghum (*Sorghum bicolor* L. Moench) varieties assessed by ISSR markers. Advances in Environmental Biology 5:3504-3511.

Vieira, E.A., F.I.F.d. Carvalho, I. Bertan, M.M. Kopp, P.D. Zimmer, G. Benin, et al. 2007. Association between genetic distances in wheat (*Triticum aestivum* L.) as estimated by AFLP and morphological markers. Genetics and Molecular Biology 30:392-399.

Wang, F., and C.-Z. Liu. 2009. Development of an economic refining strategy of sweet sorghum in the Inner Mongolia region of China. Energy & Fuels 23:4137-4142.

Zhao, Y.L., A. Dolat, Y. Steinberger, X. Wang, A. Osman, and G.H. Xie. 2009. Biomass yield and changes in chemical composition of sweet sorghum cultivars grown for biofuel. Field Crops Research 111:55-64.

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

1 Literature Review

1.1 Introduction

There is an urgent need for alternative plant species that can produce large volumes of biomass for conversion into cost effective bioethanol on a large industrial scale. This is due to the rapid rise of fossil fuels which increase carbon dioxide levels in the atmosphere and contribute to global warming. The purpose of this literature review is to provide a current understanding on the importance of sweet sorghum production and its use as an alternative crop that can be used in the South African biofuel industries. The chapter gives an overview on global biofuel production and its importance by elaborating on the challenges with regards to fossil fuels, the impact of biofuels and highlighting the top producing countries in the world. It further describes sweet sorghum as an alternative source for bioethanol production by giving an outline on its origin and botany, ecology and production constraints, types of sugar in sweet sorghum stalks and benefits. The review further gives an overview on sweet sorghum challenges regarding bioethanol production in South Africa. The chapter also provides literature on morphological and molecular markers as the basis to assess genetic diversity in sweet sorghum. The different methods used to explore genetic diversity, such as cluster and principal component analysis are also outlined.

1.2 Biofuel production

1.2.1 Global challenges with regard to fossil fuels

Fossil fuels are non-renewable primary sources of energy globally. For decades, fossil fuels have been used in the energy industries because of easy accessibility and affordability. However, the rapid rise in industry development and population growth has led to an increase in energy demand around the world. This has further caused major fluctuations in fuel prices. This type of energy accounts for approximately 65% of the worlds' energy consumption. Coal, petroleum and natural gas being the main fossil fuel sources are responsible for 45%, 35%, and 20% of global greenhouse gas emissions (Covert et al., 2016). The use of coal has led to extensive environmental degradation, causing destruction of wild lands, acidification and soil erosion. It is reported by Almodares and Hadi (2009) that 11.2 million ha year $^{-1}$ of forest is lost due to the use of fossil energy and is a serious concern. When burned, high levels of carbon dioxide are produced and released into the atmosphere, contributing towards global warming (Archer, 2005; Mutepe et al., 2012). There is therefore need to look for alternative renewable sources of energy to supplement fossil fuel supplies.

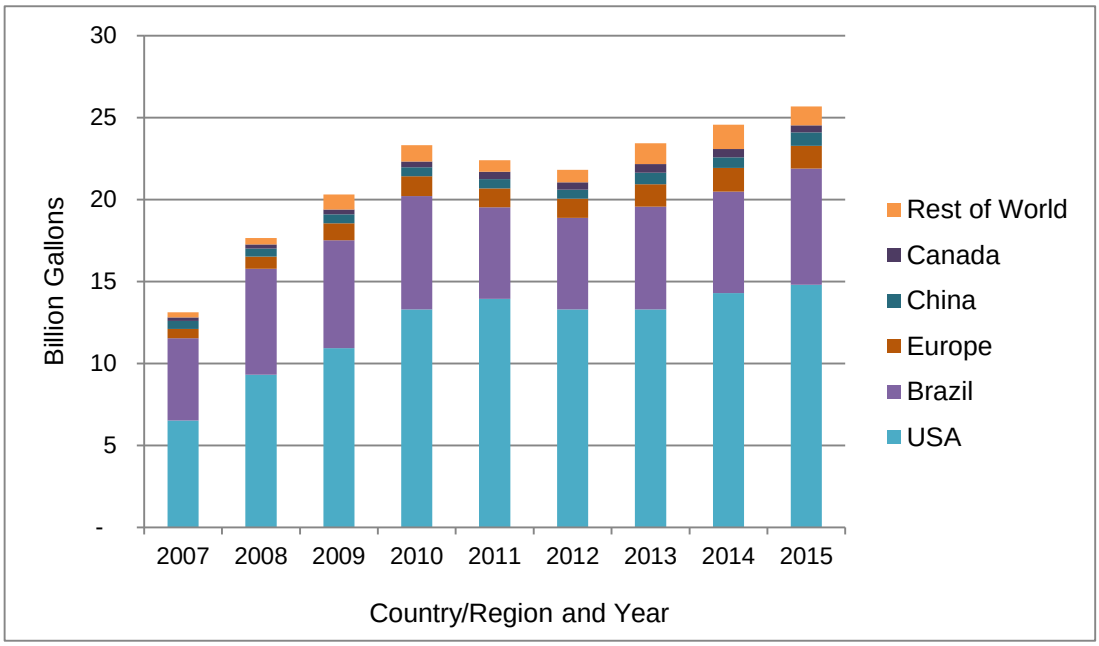
1.2.2 Biofuel as an alternative source of energy

Biofuels have been globally used as energy sources for heating and cooking for centuries. Nonetheless, the rise in industry development and population growth has led to an increase in global energy demand. A Biofuel is a type of energy that is produced from renewable biological resources (Demura and Ye, 2010). It is defined by Demirbas (2009) as a liquid or gaseous fuel that is predominantly produced from biomass. Unlike other renewable resources, biomass can be converted directly into burnable fuel which becomes biofuel (Demirbas, 2009). According to Hood (2016), ethanol and biodiesel are the most widely used biofuels and are both found in liquid form. This type of energy is known to be a sustainable solution to global energy crisis, global warming and global pollution. For this reason, renewable fuel sources are being promoted and are the most promising alternatives of fossil fuels (Meyer et al., 2005). According to Cilliers (2012), plans for a number of biofuel projects have been established in South Africa despite the lack of commercial feedstock needed in the biofuel industries (Figure 1-1).

Figure 1-1: Location of biofuel projects across provinces in South Africa: North-Western (Silversands), Free State (Ethanol Africa), KwaZulu-Natal (Siyanda, First In Spec. BDC), Eastern Cape (Arengo, biofusion, Phytoenergy, Green Tech., Biobreen, RNRF)

1.2.3 Bioethanol production

Ethanol is an alcohol fuel produced by fermentation of sugars found in crops such as sorghum, sugarcane and maize (Johnson and Matsika, 2006). To reduce air pollution, which is the main disadvantage of fossil fuels, different countries and regions have resorted to producing bioethanol from biofuel crops. These include; United States of America (USA), Brazil, Europe, China, Canada and the rest of the world (Johnson et al., 2007; Timilsina and Shrestha, 2011). USA and Brazil are shown to be the top producing countries, producing 14,806 and 7,093 million gallons respectively in 2015 as shown in Figure 1-2.



Source: Renewable fuel association (2016)

Figure 1-2: Global ethanol production by Country/Region and Year

It is reported by Johnson and Matsika (2006) that Sub-Saharan Africa has the greatest potential for bioethanol production due its underutilized vast areas of land, sub-tropical climate and low productivity levels. According to Reddy et al. (2007), ethanol production is 5000 litres ha $^{-1}$ year $^{-1}$ in Sub-Saharan Africa. South Africa is known to have energy intensive sectors and is the largest emitter of greenhouse gases in Africa. This had led the country to voluntarily opt for a reduction in fossil fuel dependency and promote biofuels (Meyer et al., 2005). According to Meyer et al. (2005), sugarcane produces about 5500 l ha^{-1}. On the hand, the use of maize for bioethanol production enabled the establishment of Ethanol-Africa located in Bothaville, Bloemfontein in 2006 by GrainSA

and SA biofuels Association (SABA). The plant was to produce 470,000 l day $^{-1}$ of ethanol from 1125 tons of maize. However, maize being a food crop, it was prohibited for use in bioethanol production due to food security concerns. The group, therefore, identified sweet sorghum as a potential biofuel crop for bioethanol production due to the high fermentable sugars in its stalks and its ability to thrive in marginal envrionments (Tsuchihashi and Goto, 2008). This is supported by Prasad et al. (2007) who emphasised that using improved sweet sorghum cultivars in South Africa as feedstock could double the size of the current sorghum market if yields of new varieties are much higher than the current average. This will, in turn, improve the livelihoods of rural communities who will be key players in its production. Despite the many benefits, there is currently no sweet sorghum commercial production units in South Africa (Musango and Brent, 2011).

1.3 The use of sweet sorghum for bioethanol production

1.3.1 Sorghum origin and domestication

In order to develop improved sweet sorghum cultivars for biofuel production, there is need to understand and appreciate the origin and characteristics of the crop. Sorghum [*Sorghum bicolor* (L.) Moench] is a cereal crop species belonging to the family Poaceae (Bryan, 1990). Even though the geographical root still remains unclear, according to different literature, Ethiopia and surrounding countries are considered as the geographical area of origin (Dillon et al, 2007). Sorghum is known to have a wide genetic diversity ranging from 20 to 30 species, showing potential for crop improvement (Assar et al., 2005). This could have been as a result of the distribution of sorghum races across Africa and around the world due to various tribal and trade movements (Doggett, 1988; Acquaah, 2007). There are five cultivated (*Sorghum bicolor* subspecies *bicolor*) sorghum races; bicolor, caudatum, durra, guinea and kafir. The races are differentiated based on panicle morphology, grain size, yield potential and adaptation (Acquaah, 2007). Sorghum is known to be classified into four major groups based on the applications; sweet sorghum, grain sorghum, broom and grass sorghum. Sweet sorghum and grain sorghum are the commonly grown types. The main difference between the two types is that sweet sorghum is mainly produced for the sucrose-filled stalks while grain sorghum is used as a staple food in most tropical areas of Africa and Asia. However, Ratnavathi et al. (2010) reported that the two types of sorghum originated from the same species [*Sorghum bicolor* (L.) Moench].

1.3.2 Sweet sorghum botany

Sweet sorghum has an extensive root system that helps in drought tolerance. The stalk height depends on its number of nodes and internodes, and may go up to 4 m high (Kunyuga, 2013). Sweet sorghum has characteristically juicy stalks which are also associated with stay-green (non-senescence) characteristics of the plant. According to (Pande et al., 1989), such stalks have been observed to show resistance against damage by termites and charcoal rot. Studies done by Carter et al. (1989) also observed distinct differences between sweet and grain sorghum. Although sweet sorghum is similar in appearance with grain sorghum, it is believed to produce more stalks and more finely branched roots due to a few gene differences which regulate plant height, amount of juice in stalks and sugar content. Even though there is a large emphasis on the rich sugar stalks of sweet sorghum in different literature (Antonopoulou et al., 2008; de Vries et al., 2010), the biofuel crop also produces about 1,500-7,500 kg ha $^{-1}$ of grain with a diameter ranging from 3 to 4 mm.

1.3.3 Sweet sorghum ecology

Sweet sorghum is described as a C4 annual grass with high photosynthetic efficiency (Olson et al., 2012). The "stay-green" genes possessed by some cultivars enables the crop to always carry out photosynthesis (Borrell et al., 2000). It is mainly grown between 40^{o}N and 40^{o}S in arid, semi-arid tropics and subtropics and can also be grown at an altitude of up to 3000 m above sea level (Reddy et al., 2012; Elangovan et al., 2015). Sweet sorghum requires about 350-700 mm of annual rainfall per growing season which is dependent on the length of the growing cycle. Being a short day crop, it requires about 90-140 days to mature at a temperature range of 27-30℃ depending on climatic conditions and type of cultivar (Malala, 2010). However, the best yields are obtained at a temperature range of 24-27℃ (Kamuntu, 2010). According to Almodares and Hadi (2009), the crop exhibits a good adaptation in a wide range of soils; from clay to light sandy soils with a pH range of 5.0-8.5. It is also known to be adapted to a wide range of environmental conditions due to its genetic variation in the response of photoperiod and temperature (Gnansounou et al., 2005).

1.3.4 Production constraints

Biotic and abiotic factors can cause substantial yield losses in sweet sorghum. According to Reddy et al. (2012), pests such as stem borer and shoot fly, and diseases such as Striga are the major biotic factors that cause serious yield losses in Sub-Saharan Africa. Stem borer (particularly, *Busseola fusca* and *Chilo partellus*) and shoot fly (*Antherigona soccata*) have

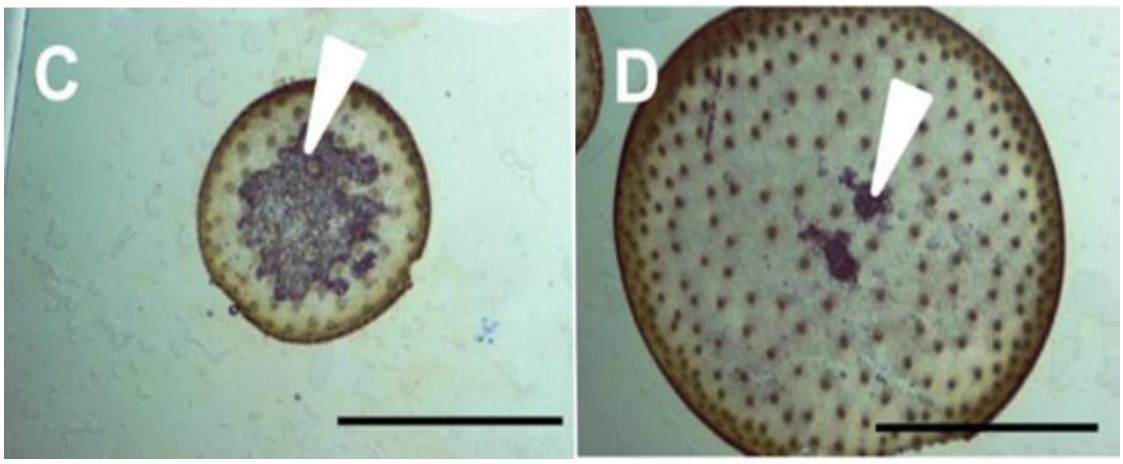
been identified to cause serious economic yield losses in sweet sorghum. Teka (2014) also reported that weeds like *Striga* spp. can cause serious yield losses of up to 50%. On the other hand, Reddy et al. (2006) emphasized that extreme drought, high temperatures, low soil fertility and acidic soils are abiotic stresses that could also contribute to yield losses in sweet sorghum. In order to prevent severe losses, there is need to consider the various production constraints and implement correct management practices.

1.3.5 Types of sugar in sweet sorghum stalks

Sweet sorghum stalks have two main types of non-structural carbohydrates (NSC); (i) saccharin-type; used for refining crystal sugar as it mainly contains sucrose (70%) and (ii) syrup-type; used for producing syrup as it mainly contains glucose (20%). The main economic product of sweet sorghum is the saccharin-type (Almodares and Hadi, 2009). Variation in sugar content is known to exist along and across the stalk during the plants growth stages. Studies done by Rose and Botha (2000) have shown that along the stalk, the lower tissue (six to nine internodes) of sorghum stalk contains less sucrose content than the upper tissue (three to six internodes) because metabolic processes involving respiration and growth rate occurs in the lower tissue of the plant. It is also evidenced by Tarpley and Vietor (2007) that glucose is found at higher concentrations in young and elongating internodes while in ripening internodes, sucrose in observed to be dominant. The sugar content across the stalk is high in the inner pith with sucrose and glucose possessing 67% and 4% respectively. On the other hand, the outer part of the stalk (bark) is believed to possess 32% and 2% of sucrose and glucose respectively (Rose and Botha, 2000; Makanda, 2009).

Different studies have shown a difference between grain and sweet sorghum in non-structural carbohydrates (NSC) partitioning. Sweet sorghum accumulates more of NSC than grain sorghum during all stages of plant growth. Grain sorghum partitions carbohydrates to one sink which are the grains while sweet sorghum partitions to two sinks; grains and stalk. Therefore grain sorghum has been shown to have less carbohydrates due to increased portioning to the grains in the apical panicle. This is also validated by the high number of dead-air-filled cells in the center of the stalk of grain sorghum as compared to sweet sorghum which has a juicy stalk with little to no dead-air-filled cells (Figure 1-3) (Blum et al., 1997). Light interception and stay-green trait in sweet sorghum can also have a major impact on sucrose concentrations in the internodes. Additionally, stay-green varieties of sweet sorghum usually have higher stalk sugar concentrations than senescing lines (Borrell et al., 2000). Duncan et al. (1981) indicated that

this may be due to the reduced need for re-mobilizing stem sucrose in addition to prolonged photosynthetic capacity.



Source: Blum et al. (1997)

Figure 1-3: High number of dead-air-filled cells in the center of the stalk of grain sorghum (C) as compared to sweet sorghum (D) which has a juicy stalk with little to no dead-air-filled cells

1.3.6 Benefits of using sweet sorghum for bioethanol production

In South Africa, the cultivation of sweet sorghum for bioethanol has been underutilized. This is despite its many benefits such as; having a sugar-rich stalk almost like sugarcane, rapid growth, high sugar accumulation, biomass production potential and wide adaptability (Tsuchihashi and Goto, 2008). Harvey and Pilgrim (2011) indicated that water and land availability are becoming major constraints to agricultural production in the coming years. This will, in turn, cause serious challenges in the cultivation of sugarcane which has been the major feedstock for bioethanol production in the country (Cilliers, 2012). Sweet sorghum is therefore an alternative biofuel crop that can be grown with less inputs and management as compared to sugarcane. Since sweet sorghum thrives on marginal lands, there is less deforestation or need for proper production cropland. The benefits of sweet sorghum compared to sugarcane are summarized in Table 1-1.

Table 1-1: Comparative characterisitcs of sweet sorghum to sugarcane

Parameter	Sweet Sorghum	Sugarcane
Propagation	Seed	Vegetative
Crop Duration	4 months	12-13 months
Water requirements (m^3 ha^{-1})	8000	36000
Crop management	Easy management, low fertilizer	Good management, high fertilizer
Millable stalk (t ha^{-1})	45 - 65	60 - 85
Ethanol productivity (l ha^{-1})	2475 - 3500	4350 - 7000
Grain yield (t ha^{-1})	2- 6	0
Harvesting	Difficult and laborious	Very simple

Source: Reddy et al. (2005), Vinutha et al. (2014)

1.4 Sweet sorghum challenges regarding bioethanol production in South Africa

South Africa's energy intensive economy has led to government intervention to promote biofuel production. According to Meyer et al. (2005), the technology and infrastructure required for the fuel industry in the country is believed to be well established. Ethanol Africa and Silversands Ethanol are examples of the various biofuel projects that have been put in place in recent years (Musango and Brent, 2011). There is therefore need to have readily available feedstock required by the industries for biofuel production. However, even with all the available biofuel crops in South Africa, the feedstock is still not enough to meet the rising demand. Sweet sorghum has therefore been identified to complement the existing feedstock material (Tsuchihashi and Goto, 2008). However, sweet sorghum is currently being cultivated mainly by smallholder farmers who only have access to local varieties. Its suitability for the production of ethanol has not been proven on a large scale due to lack of commercial material on the market (Cilliers, 2012). It is evidenced by Prasad et al. (2007) that the current sorghum market could be doubled if superior varieties are readily accessible by stakeholders. Therefore, it is imperative to assess and identify genotypes that will be utilized to develop superior sugar-rich

commercial cultivars. The cultivation of commercial material will not only boost the biofuel industry, but will provide rural stimulation by fully engaging in improved sweet sorghum production and hence increase their income generation. This is strongly supported by Amigun et al. (2008) through his similar findings.

1.5 Testing sweet sorghum lines for bioethanol production in South Africa

1.5.1 Biomass yield and sugar-related traits

Selection is known to be the most successfully used method in plant breeding. One of the many objectives of a sweet sorghum plant breeder is to develop superior cultivars with high biomass and sugar-related traits. However, it is evidenced by Ilker (2011) that secondary traits have a major influence on response to selection. Therefore, proper knowledge on the relationships among quantitative traits is important for assessing the feasibility of joint selection of two or more traits (Ezeaku and Mohammed, 2006). In sweet sorghum, performance of stalk biomass yield and sugar-related traits such as stalk sugar content and plant height are known to be affected by a number factors such as the environment (Boćanski et al., 2009; Elangovan et al., 2014). Evaluating genotype performance across different locations is therefore essential when exploiting existing variability and developing improved cultivars (Faisal and Aisha, 2011; Abubakar and Bubuche, 2013). This is evidenced by several studies that were carried out to characterize the interaction pattern of stalk biomass yield and its components in different locations. The findings confirmed the presence of a significant interaction between the genotype and the environment as a consequence of the differential response of the genotypes to environmental changes (Makanda et al., 2009; Elangovan et al., 2014).

Temperature and seasonal length have been reported to have a major influence on the relative performance of sweet sorghum genotypes. A study carried out by Zou et al. (2011) reported that temperature and photoperiod differed significantly between locations and that no two locations are exactly the same. Higher temperatures and short day length hastened the flowering heading date and plants where observed to be significantly shorter with lower stalk biomass yield and °brix readings. These findings are supported by various researchers who also observed genotypic differences in stalk biomass yield and sugar-related traits due to differences in climatic conditions between locations (Zou et al., 2011; Elangovan et al., 2014; Olweny et al., 2014). On the other hand, lower temperatures and longer seasonal length have been shown to cause a positive relationship between stalk biomass yield and °brix readings,

indicating that taller plant cultivars require a longer growth cycle and in turn increase the stalk sugar content.

1.5.2 Correlation between stalk biomass and sugar-related traits

According to Malik et al. (2005), correlation analysis is a technique which helps to explain the degree of linear relationships between plant quantitative traits without regard to cause and effect. Stalk biomass yield is a trait that is controlled by polygenes, with each different gene contributing a small effect to its expression (Figure 1-4). There is therefore a relationship between stalk production and plant height; stalk diameter and juice yield; sugar yield, °brix values and sugar content (Gutjahr et al., 2013). Sweet sorghum in known to have a strong sink available for sugars at soft and hard dough stages. Thick stalk, late maturing and tall sweet sorghum genotypes are known to have more juice extracted from their stalks due to higher biomass and stalk yield production (Disasa et al. (2016). An evaluation done by Rao et al. (2009) in India (ICRISAT) on sweet sorghum, recorded a positive and very highly significant correlation between fresh stalk yield with fresh biomass and juice yields respectively ($r = 0.80$ and $r = 0.84$; $p \leq 0.01$). On the other hand, Shiringani et al. (2010) reported a positive and highly significant correlation between sugar content and sucrose content ($r = 0.99$, $p \leq 0.01$), while ^{o}brix showed a positive and highly significant correlation with sucrose content ($r = 0.61$, $p \leq 0.01$). These findings are consistent with other studies, suggesting that ^{o}brix could be used as a surrogate trait for measuring total fermentable sugars and theoretical ethanol yields in screening large numbers of breeding materials and segregating populations (Saha et al., 2013).

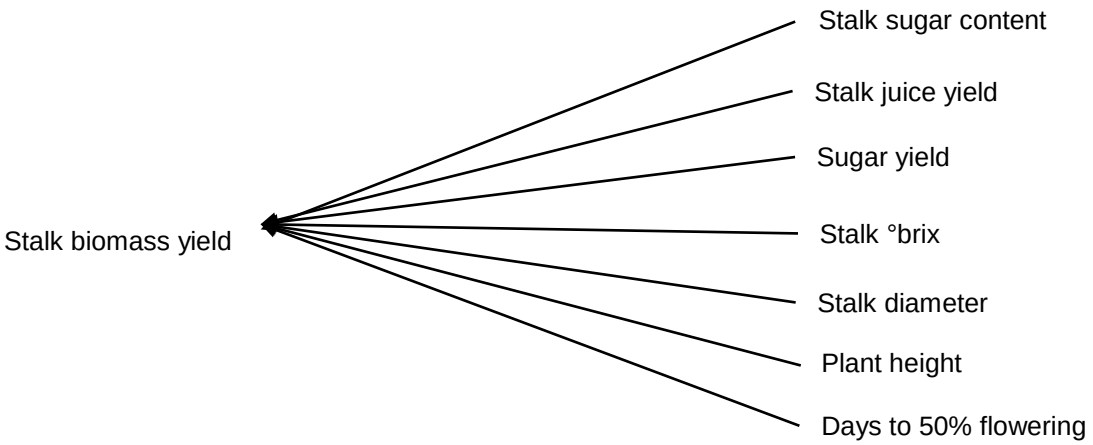


Figure 1-4: Relationship between biomass yield and sugar-related traits

1.5.3 Genetic diversity

Genetic diversity study in crop species is the underlying principle for advancement and conservation. Genetic diversity refers to the total number of genetic characteristics in the genetic makeup of species while genetic variation occurs within alleles both within and among species. Selection of desirable traits in a breeding program is determined by genetic variation which further improves the genetic diversity of plant genetic resources (Cholastova et al., 2013). Therefore, the biofuel industry in South Africa can only be improved by identifying genetically diverse segregating parents from sweet sorghum germplasm. However, the most common germplasm found in the country are landraces which are widely grown by smallholder farmers (Mofokeng et al., 2014). This is as a result of sweet sorghum not being considered as an important crop. The rise in energy demand has however led to governments intervention in promoting improved sweet sorghum production due its high sugar content in its juicy stalks (Meyer et al., 2005; Cilliers, 2012). Therefore, it is prudent to acquire knowledge on genetic diversity to facilitate in the effective selection of superior germplasm possessing desirable characteristics for hybridization on the basis of divergent analysis (Ganesamurthy et al., 2010; Tomar et al., 2012; Govindaraj et al., 2015).

Various studies across the world have been documented on sweet sorghum diversity. Their findings validate the importance of genetic diversity studies for; (i) identification and selection of desirable lines for hybrid development and conservation (Tomar et al., 2012; Singh and Singh, 2015), (ii) characterization of individuals in determining duplications for germplasm collections (Disasa et al., 2016), and (iii) selecting parents from heterotic groups formed from cluster analysis (Zhan et al., 2012). In South Africa, a study done by Mofokeng et al. (2014) has been documented on sorghum genetic diversity of 103 diverse landraces and breeding lines. The study proved the existence of considerable genetic diversity among sorghum germplasm. However, readily available improved sweet sorghum germplasm is still limited on the market. The sweet sorghum genotypes obtained from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and the African Centre for Crop Improvement (ACCI-South Africa) in this study will therefore help breeders to determine genetic diversity among sweet sorghum lines for use in new hybrid development suitable for biofuel production in South Africa.

1.5.3.1 Phenotypic selection

Characterization and evaluation of genetic diversity has been traditionally based on morphological traits for many years (Assar et al., 2005). This is proven by Mendel who explored phenotypic selection based on visible traits in the progeny of sexual crosses (Bateson and Mendel, 2013). Qualitative and quantitative traits are widely used in selection because of being regarded as a cheap technology and such traits are necessary in the understanding and formulation of ideotypes (Bänziger et al., 2006). Similarly, Mujaju et al. (2011) also pointed out that the use of morphological traits is the classical way of assessing genetic diversity in Southern Africa where resource limitation for molecular markers is prevalent. However, phenotypic selection is unreliable because it is influenced by environmental factors as documented by Govindaraj et al. (2015). Another shortcoming is that this type of selection is impeded by low polymorphism and requires plants to grow to full maturity prior to identification (Patil et al., 2014). Phenotypic selection is also known to be a slow process and expensive in the long run because it requires many locations and seasons to efficiently select desirable traits. Therefore, there is a need to complement it with marker-assisted selection.

1.5.3.2 Marker-assisted selection

Marker-assisted selection has been successfully used in genetic diversity studies (Assar et al., 2005; Aremu, 2012). According to Schulman (2007), molecular markers are nucleotide sequences corresponding to a physical position in the genome and their polymorphisms between accessions allow the pattern of inheritance to be easily traced. DNA-based marker selection has gained popularity among breeders due to its high polymorphic property and quick generation of quality data. The lack of environmental influences has also enabled DNA-based markers to be detected in all tissues at all developmental stages (Kumar et al., 2009; Govindaraj et al., 2015). On the other hand, the markers are known to be reliable, stable and have the ability to discriminate between homozygotes and heterozygotes (Govindaraj et al., 2015). However, in developing countries, the main cause of low adoption rate has been due to high costs and technical demand (Ribaut et al., 2010).

The various types of DNA-based markers that have been used to measure genetic diversity include; (i) RFLP-restriction fragment length polymorphism, (ii) AFLP-amplified fragment length polymorphism, (iii) RAPD-random amplification of polymorphic DNA, (iv) SSR-simple sequence repeats and (v) SNP-single nucleotide polymorphism (Aremu, 2012). According to Govindaraj et al. (2015), these markers are brought about by mutations in the DNA segment

through insertions, deletions, duplications, inversions or translocation. The markers also differ in principle, application, amount of polymorphism detected and time required. However, Geleta et al. (2006) emphasized that SSRs have proven to be more variable and are used in most genetic diversity studies. SSR markers are described by Govindaraj et al. (2015) as short tandem repeats which range from 1 to 10 bp. These markers are popularly utilized due to their high degree of polymorphism, co-dominance and multiple allele properties (Geleta et al., 2006). This is also evidenced by various researchers who have used SSRs in sorghum genetic diversity studies (Gurmu et al., 2009; Kunyuga, 2013; Mofokeng et al., 2014).

1.5.3.3 Conventional versus molecular breeding

Despite molecular markers providing a precise assessment of results, the analysis covers the entire genome instead of targeting the regions with desired traits (Benin et al., 2012). Most breeders are also known to use morphological data in genetic diversity studies because the same type of information is obtained from characterizations, adaptability and yield potential measurements (Bertan et al., 2007). Therefore, the use of both morphological and molecular markers is prudent in providing more representative sampling of the genome (Franco et al., 1997; Mohammadi and Prasanna, 2003). This is evidenced by Barata and Carena (2006) who reported that extensive field assessment is required to complement SSR markers in genetic diversity studies. These findings were later supported by Lekgari and Dweikat (2014) who reported that the use of morphological and molecular marker data is the solution to any successful genetic diversity study. The use of morphological data in the evaluation of 142 sweet sorghum lines in their study formed distinct cluster groups, while SSR clusters further narrowed the groups based on the origin. Therefore, phenotyping and SSRs will be used to assess genetic diversity in this study.

1.5.3.4 Genetic distance

Genetic distance is an important concept for distinguishing among plant materials. According to Nei (1973), genetic distance (GD) is defined as that difference between two entities that can be described by allelic variation. It was later eloborated by Nei (1987) that genetic distance is a way to predict genetic variability in crops by grouping similar germplasm into heterotic groups. Closely related species are known to have simliar alleles with smaller genetic distance than distantly related species. This is due to the low genetic variation within the population which leads to inbreeding depression and loss of desired traits (Losa et al., 2012). According to literature, genetic distance can be estimated based on morphological, biochemical and

molecular markers through six important steps. These include; i) choosing genotypes to be examined, ii) data production and formatting, iii) choosing the distance definition or measurement to be used for the estimations, iv) choosing clustering or plotting procedure to be used, v) analysis of the degree of distortion caused by the clustering/plotting procedure used and vi) interpreting the data. Even though genetic distance can be measured without phenotyping the germplasm material, phenotypic traits are the basis for estimating genetic distance by the main multivariate technique (Bertan et al., 2007).

1.5.3.5 Measures of genetic distance

There are various techniques used to estimate genetic distance (GD) between and within crop populations. In each method, multivariate and multiple data are essential in achieving accurate and desired results. Multivariate data obtained from morphological data exhibits various variables such as discrete and continuous while multiple data obtained from both morphological and molecular data displays the strengths and constraints in the choice of each of the data sets (Aremu, 2012). Literature documented by Bertan et al. (2007) emphasized that Euclidean (S^2) and the Mahalanobis (D^2) distances are the most common statistical procedures used to estimate genetic distance. These methods measure both morphological and molecular based marker data sets. However, Mahalanobis distance has some benefits over Euclidean distance because it takes into account the environmental influences and obtains correlations between characters. Even so, the drawback of using Mahalanobis distance is that it can only be estimated from data obtained from more than one replication (Bertan et al., 2007). The available data is then displayed in a symmetrical matrix once the distance estimate is acquired and analysis is followed (Bertan et al., 2007).

1.5.4 Analysis for genetic diversity

Established multivariate techniques are used to analyze genetic distance among breeding materials irrespective of the data (Bertan et al., 2007; Aremu, 2012). These methods include, principal component analysis (PCA), cluster analysis, principal coordinate analysis (PCOA), canonical correlation and multidimensional scaling (Aremu, 2012; Lekgari and Dweikat, 2014). Among the mentioned techniques, cluster analysis (dendrogram) and principal component analysis (PCA) will be used in this study.

1.5.4.1 Cluster analysis

Cluster analysis is a multivariate technique that groups genetically similar genotypes together through a repetitive process that results in cluster formation (Bertan et al., 2007). According to Mohammadi and Prasanna (2003), a successful classification should show high within cluster homogeneity and high between cluster heterogeneity. Through geometrical visualization, genotypes within a cluster are closer together and those in different clusters are observed to be further apart. The two broadly used cluster methods by plant breeders include distance-based and model-based methods. The distance-based method is known to group genotypes by a repetitive process and forms a tree-like structure (dendrogram) without concern with the number of clusters generated. On the other hand, model based method generates groups according to a fixed clustering criterion which is mutually exclusive (Bertan et al., 2007). Additionally, distance-based method can be classified into two groups; hierarchical and nonhierarchical. In plant breeding, hierarchical clustering methods are mostly used to assess the genetic diversity in crop species. Among the various agglomerative hierarchical methods such as; unweighted pair group method with arithmetic averages (UPGMA), neighbor-joining method and Ward's method (Govindaraj et al., 2015), UPGMA is mostly used as it provides more accuracy by discriminating closely related genotypes in various crops (Bertan et al., 2007; Trindade et al., 2010). Various researchers including Assar et al. (2009), Zhan et al. (2012) and Kunyuga (2013) have used UPGMA cluster analysis to separate different breeding materials into major and sub-clusters. This is in order to efficiently differentiate, select and discriminate among breeding materials for cultivar improvement.

1.5.4.2 PCA analysis

Principal component analysis is known to produce a 2 or 3 dimensional scatter plot that enables visualization of the differences among the samples (Mohammadi and Prasanna 2003). According to Wiley (1981), PCA is defined as a method of data reduction to clarify the relationships between two or more characters and to divide the total variance of the original characters into a limited number of uncorrelated new variables. Mohammadi and Prasanna (2003) further explained that the reduction in data is obtained by a linear transformation of the original variables into a new set of uncorrelated variables known as principle components (PCs). The interpretation of PCs is to identify which variables are strongly associated with each component. The first PC is a variable system for the data set that summarizes the greatest variance of the data relative to all remaining PCs. The subsequent PCs are known to summarize most of the variability not summarized by the previous PC (Mohammadi and Prasanna, 2003). The eigenvectors, which are the locations along each component are then

associated with values across all variables. On the other hand, the component's eigenvalue is the association between the components and original values (Jolliffe, 2002).The major difference between cluster and PCA analysis is that the former depends on previous measures of genetic distance while the later uses data obtained from evaluated genotypes directly without using genetic distance measurements to create a graphical representation (Bertan et al., 2007).

1.6 Conclusion

Fluctuating oil prices and uncertainties regarding future reserves as well as the phenomenon of global warming has led South Africa to consider alternative means of energy generation. Sweet sorghum has been considered as a crop for biofuel production mainly in South Africa because maize and grain sorghum could contribute to the threats to food security, especially among smallholder farmers. The use of sugarcane is also relatively costly in relation to the heavy investment costs in terms of management and the use of water which is a scarce resource. Currently, there is limited information on the availability of sweet sorghum lines that have high stalk biomass and sugar-related traits in South Africa.There is also lack of readily accessible sweet sorghum cultivars on the market that can supplement other biofuel crops in the biofuel industry. The study was therefore undertaken to evaluate genetic diversity among a collection of sweet sorghum lines using morphological data and SSR markers in order to identify specific genotypes exhibiting high levels of biomass yield and sugar-related traits from various heterotic groups that will improve the genetic base and characterization of the current cultivated sweet sorghum. The biofuels produced from sweet sorghum will provide sustainable and eco-friendly energy options that foster environmental sustainability, and offer enormous opportunities to improve the income generation of smallholder farmers who depend on agriculture for their livelihoods.

1.7 References

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

Almodares, A., and M. Hadi. 2009. Production of bioethanol from sweet sorghum: A review. African Journal of Agricultural Research 4:772-780.

Acquaah, G. 2007. Principles of plant breeding and genetics. Malden, MA USA: Blackwell Publishing:509-518.

Alhajturki, D., M. Aljamali, A. Kanbar, and F. Azmah. 2012. Potential of some sweet sorghum (*Sorghum bicolor* L.) genotypes under two water regimes for sugar and bio-ethanol production. Sugar Technology 14:376-382.

Almodares, A., and M. Hadi. 2009. Production of bioethanol from sweet sorghum: A review. African Journal of Agricultural Research 4:772-780.

Amelework, B., H. Shimelis, P. Tongoona, M. Laing, and F. Mengistu. 2015. Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. Plant Genetic Resources 13:131-141.

Amigun, B., R. Sigamoney, and H. von Blottnitz. 2008. Commercialisation of biofuel industry in Africa: a review. Renewable and Sustainable Energy Reviews 12:690-711.

Antonopoulou, G., H.N. Gavala, I.V. Skiadas, K. Angelopoulos, and G. Lyberatos. 2008. Biofuels generation from sweet sorghum: fermentative hydrogen production and anaerobic digestion of the remaining biomass. Bioresource Technology 99:110-119.

Archer, D. 2005. Fate of fossil fuel CO_2 in geologic time. Journal Geophysical Research, 110, C09S05. doi:10.1029/2004JC002625.

Aremu, C. 2011. Genetic diversity: a review for need and measurements for intraspecies crop improvement. Journal Microbiology Biotechnology Research 1:80-85.

Aremu, C. 2012. Exploring statistical tools in measuring genetic diversity for crop improvement. INTECH Open Access Publisher.

Assar, A., R. Uptmoor, A. Abdelmula, M. Salih, F. Ordon, and W. Friedt. 2005. Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). Crop Science 45:1636-1644.

Assar, A.H.A., R. Uptmoor, A.A. Abdelmula, C. Wagner, M. Salih, A.M. Ali, et al. 2009. Assessment of sorghum genetic resources for genetic diversity and drought tolerance using molecular markers and agro-morphological traits. University of Khartoum. Journal of Agricultural Sciences. 17:1-22.

Bänziger, M., P.S. Setimela, D. Hodson, and B. Vivek. 2006. Breeding for improved abiotic stress tolerance in maize adapted to southern Africa. Agricultural Water Management 80:212-224.

Barata, C., and M. Carena. 2006. Classification of North Dakota maize inbred lines into heterotic groups based on molecular and testcross data. Euphytica 151:339-349.

Bateson, W., and G. Mendel. 2013. Mendel's principles of heredity. Courier Corporation.

Benin, G., G. Matei, A.C. de Oliveira, G. Silva, T. Hagemann, C.L. da Silva, et al. 2012. Relationships between four measures of genetic distance and breeding behavior in spring wheat. Genetics and Molecular Research 11:2390-2400.

Bertan, I., F. Carvalho, and A.d. Oliveira. 2007. Parental selection strategies in plant breeding programs. Journal of Crop Science and Biotechnology 10:211-222.

Blum, A., G. Golan, J. Mayer, and B. Sinmena. 1997. The effect of dwarfing genes on sorghum grain filling from remobilized stem reserves, under stress. Field Crops Research 52:43-54.

Boćanski, J., Z. Srećkov, and A. Nastasić. 2009. Genetic and phenotypic relationship between grain yield and components of grain yield of maize (*Zea mays* L.). Genetika 41:145-154.

Borrell, A.K., G.L. Hammer, and A.C. Douglas. 2000. Does maintaining green leaf area in sorghum improve yield under drought? I. Leaf growth and senescence. Crop Science 40:1026-1037.

Carter, P., D. Hicks, E. Oplinger, J. Doll, L. Bundy, R. Schuler, et al. 1989. Grain Sorghum (Milo). Alternative field crops manual. University of Wisconsin-Extension, Cooperative Extension. Online publication. Available, email: yyang2@ unl. edu, or http://www. hort. purdue. edu/newcrop/afcm/. (Accessed on 18 January, 2016).

Cholastova, T., M. Soldanova, and R. Pokorny. 2013. Random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) marker efficacy for maize hybrid identification. African Journal of Biotechnology 10:4794-4801.

Cilliers, B.L. 2012. An industry analysis of the South African biofuels industry. Doctoral dissertation. North-West University, South Africa.

Claassen, P., T. De Vrije, M. Budde, E. Koukios, A. Glynos, and K. Réczey. 2004. Biological hydrogen production from sweet sorghum by thermophilic bacteria. Proceedings 2nd World Conference on Biomass for Energy, Rome. p. 1522-1525.

Covert, T., M. Greenstone, and C.R. Knittel. 2016. Will We Ever Stop Using Fossil Fuels? The Journal of Economic Perspectives 30:117-137.

de Vries, S.C., G.W. van de Ven, M.K. van Ittersum, and K.E. Giller. 2010. Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion techniques. Biomass and Bioenergy 34:588-601.

Demirbas, A. 2009. Political, economic and environmental impacts of biofuels: A review. Applied Energy 86:S108-S117.

Demura, T., and Z.-H. Ye. 2010. Regulation of plant biomass production. Current Opinion in Plant Biology 13:298-303.

Disasa, T., T. Feyissa, and B. Admassu. 2016. Characterization of Ethiopian Sweet Sorghum Accessions for 0Brix, Morphological and Grain Yield Traits. Sugar Technology:1-11.

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

Duncan, R., A. Bockholt, and F. Miller. 1981. Descriptive comparison of senescent and nonsenescent sorghum genotypes. Agronomy Journal 73:849-853.

Elangovan, M., K. Ganesamurthy, K. Sankarapandian, and P.K. Babu. 2015. GIS mapping of sorghum landraces collected from Tamil Nadu, India. International Journal of Advanced Research 3:125-131.

Elangovan, M., N. Seetharama, and J. Patil. 2014. Genetic diversity and heritability characters associated in sweet sorghum [*Sorghum bicolor* (L.) Moench]. Sugar Technology 16:200-210.

Ezeaku, I., and S. Mohammed. 2006. Character association and path analysis in grain sorghum. African Journal of Biotechnology 5:1337-1340.

Faisal, E., and O. Aisha. 2011. Genotype x seed production environment interaction on the performance of sorghum (*Sorghum bicolor* [L.] Moench) under irrigation. Agriculture and Biology Journal for North America 2:2151-7517.

Ganapathy, K., S. Gomashe, S. Rakshit, B. Prabhakar, S. Ambekar, R. Ghorade, et al. 2012. Genetic diversity revealed utility of SSR markers in classifying parental lines and elite genotypes of sorghum (*Sorghum bicolor* L. Moench). Australian Journal of Crop Science 6:1486-1493.

Ganesamurthy, K., D. Punitha, and M. Elangovan. 2010. Genetic diversity among the land races of sorghum collected in Tamil Nadu. Electronic Journal of Plant Breeding 1:1375-1379.

Geleta, L., M. Labuschagne, and C. Viljoen. 2005. Genetic variability in pepper (*Capsicum annuum* L.) estimated by morphological data and amplified fragment length polymorphism markers. Biodiversity & Conservation 14:2361-2375.

Geleta, N., M.T. Labuschagne, and C.D. Viljoen. 2006. Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers. Biodiversity & Conservation 15:3251-3265.

Ghani, A., M. Saeed, D. Hussain, M.M. Shafique, M. Arshad, and S.A.S. Shah. 2015. Evaluation of different sorghum (*Sorghum bicolor* L. moench) varieties for grain yield and related characteristics. Science Letters 3:72-74.

Gnansounou, E., A. Dauriat, and C. Wyman. 2005. Refining sweet sorghum to ethanol and sugar: economic trade-offs in the context of North China. Bioresource Technology 96:985-1002.

Govindaraj, M., M. Vetriventhan, and M. Srinivasan. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics Research International Journal 2015.

Gurmu, F., H. Mohammed, and G. Alemaw. 2009. Genotype x Environment interactions and stability of soybean for grain yield and nutrition quality. African Crop Science Journal 17:87-99.

Gutjahr, S., M. Vaksmann, M. Dingkuhn, K. Thera, G. Trouche, S. Braconnier, et al. 2013. Grain, sugar and biomass accumulation in tropical sorghums. I. Trade-offs and effects of phenological plasticity. Functional Plant Biology 40:342-354.

Harvey, M., and S. Pilgrim. 2011. The new competition for land: food, energy, and climate change. Food Policy 36:S40-S51.

Hood, E.E. 2016. Plant-based biofuels. F1000 Faculty Reviews 5:185. doi: 10.12688/f1000research.7418.1.

Ilker, E. 2011. Correlation and path coefficient analyses in sweet corn. Turkish Journal of Field Crops 16:105-107.

Janssen, R., D. Rutz, S. Braconnier, B. Reddy, S. Rao, R. Schaffert, et al. 2010. Sweet sorghum–an alternative energy crop. Journal of Environments 4:6.

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.

Johnson, J.M., M.D. Coleman, R. Gesch, A. Jaradat, R. Mitchell, D. Reicosky, et al. 2007. Biomass-bioenergy crops in the United States: A changing paradigm. American Journal of Plant Sciences-Biotechnology 1:1-28.

Jolliffe, I. 2002. Principal component analysis.Wiley Online Library.

Jolliffe, I.T. 1973. Discarding variables in a principal component analysis. II: Real data. Applied Statistics 22:21-31.

Kamuntu, S.P. 2010. Phenotypic and genetic diversity of sweet sorghum (*sorghum bicolor* I.) varieties and their suitability for ethanol production in Tanzania. Morogoro. Sokoine MSc thesis, University of Agriculture, Tanzania.

Kubie, L. 2013. Evaluation Of Genetic Diversity In Garden Egg (*Solanum Aethiopicum*) Germplasm In Ghana. PhD dissertation, University of Ghana, Ghana.

Kumar, A.A., B. Reddy, M. Blümmel, S. Anandan, Y.R. Reddy, C. Raviender Reddy, et al. 2010. On-farm evaluation of elite sweet sorghum genotypes for grain and stover yields and fodder quality. Animal Nutrition and Feed Technology S 10:69-78.

Kumar, P., V. Gupta, A. Misra, D. Modi, and B. Pandey. 2009. Potential of molecular markers in plant biotechnology. Plant Omics 2:141-162.

Kunyuga, P.W. 2013. Assessment of genetic diversity of sorghum (*sorghum bicolor)* accessions from Tanzania using ssr markers: implications for conservation. MSc thesis, Kenyatta University, Kenya.

Legesse, B. A.A. Myburg, K. Pixley, S. Twumasi-Afriyie, and A.-M. Botha. 2008. Relationship between hybrid performance and AFLP based genetic distance in highland maize inbred lines. Euphytica 162:313-323.

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.

Losa, A., H. Hartings, A. Verderio, and M. Motto. 2012. Assesment of genetic diversity and relationships among maize inbred lines developed in Italy. Maydica 56.95-104.

Makanda, I. 2009. Combining ability and heterosis for stem sugar traits and grain yield components in dual-purpose sorghum (Sorghum *bicolor* L. Moench) germplasm. PhD dissertation, University of KwaZulu-Natal, South Africa.

Makanda, I., J. Derera, P. Tongoona, and J. Sibiya. 2011. Development of sorghum for bio-energy: A view from the stakeholders and priorities for breeding dual purpose varieties. African Journal of Agricultural Research 6:4477-4486.

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

Makumbi, D., J.F. Betrán, M. Bänziger, and J.-M. Ribaut. 2011. Combining ability, heterosis and genetic diversity in tropical maize *(Zea mays* L.) under stress and non-stress conditions. Euphytica 180:143-162.

Malala, T.J. 2010. Evaluation and selection of 20 sorghum [*Sorghum bicolor* (L.) Moench] genotypes for drought tolerance. MSc thesis, University of Pretoria, South Africa.

Malik, H., S.I. Malik, M. Hussain, S. Chughtai, and H.I. Javed. 2005. Genetic correlation among various quantitative characters in maize (*Zea mays* L.) hybrids. Journal of Agriculture & Social Sciences 3:262-265.

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

Meyer, F., P. Strauss, M. Cutts, J. de Beer, G. du Toit, T. Funke, et al. 2005. Bio-ethanol production in South Africa: An objective analysis. Online: http://www. bfap. co. za/pdfs/BFAP% 20report% 20on 20. (Accessed on 13 March, 2016).

Mofokeng, A., H. Shimelis, P. Tongoona, and M. Laing. 2014. A genetic diversity analysis of South African sorghum genotypes using SSR markers. South African Journal of Plant and Soil 31:145-152.

Mohammadi, S., and B. Prasanna. 2003. Analysis of genetic diversity in crop plants—salient statistical tools and considerations. Crop Science 43:1235-1248.

Mujaju, C., A. Zborowska, G. Werlemark, L. Garkava-Gustavsson, S. Andersen, and H. Nybom. 2011. Genetic diversity among and within watermelon (*Citrullus lanatus*) landraces in Southern Africa. The Journal of Horticultural Science and Biotechnology 86:353-358.

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

Mutepe, R.D., S. Marx, and P. Van der Gryp. 2012. Ethanol production from sweet sorghum. MSc thesis, North-West University, Potchefstroom Campus, South Africa.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences 70:3321-3323.

Nei, M. 1987. Genetic distance and molecular phylogeny. Population Genetics and Fishery Management. Seattle. USA:193-223.

Olson, S.N., K. Ritter, W. Rooney, A. Kemanian, B.A. McCarl, Y. Zhang, et al. 2012. High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops. Biofuels, Bioproducts and Biorefining 6:640-655.

Olweny, C., J. Jamoza, M. Dida, W. Kimani, J. Njuguna, D. Githae, et al. 2014. High genetic diversity for improvement of sweet sorghum (*Sorghum bicolor* (L.) Moench) genotypes for sugar and allied products. Molecular Plant Breeding 5:28-35.

Pande, S., L. Mughogho, N. Seetharama, and R. Karunakar. 1989. Effects of nitrogen, plant density, moisture stress and artificial inoculation with Macrophomina phaseolina on charcoal rot incidence in grain sorghum. Journal of Phytopathology 126:343-352.

Patil, A., V. Premi, V. Sahu, M. Dubey, G. Sahu, and G. Chandel. 2014. Identification Of Elite Rice Germplasm Lines For Grain Protein Content, SSR Based Genotyping And DNA Fingerprinting.International Journal of Plant, Animal and Environmental Sciences 4:127-136.

Pecina−Quintero, V., J.L. Anaya−López, A. Zamarripa−Colmenero, N. Montes−García, C. Nuñez−Colín, J.L. Solis−Bonilla, et al. 2012. Genetic diversity of sweet sorghum germplasm in Mexico using AFLP and SSR markers. Pesquisa Agropecuária Brasileira 47:1095-1102.

Perumal, R., R. Krishnaramanujam, M.A. Menz, S. Katilé, J. Dahlberg, C.W. Magill, et al. 2007. Genetic diversity among sorghum races and working groups based on AFLPs and SSRs. Crop Science 47:1375-1383.

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

Rao, S., J. Patil, A. Umakanth, J. Mishra, C. Ratnavathi, G.S. Prasad, et al. 2013. Comparative performance of sweet sorghum hybrids and open pollinated varieties for millable stalk yield, biomass, sugar quality traits, grain yield and bioethanol production in tropical Indian condition. Sugar Technology 15:250-257.

Rao, S., S. Rao, N. Seetharama, A. Umakath, P.S. Reddy, B. Reddy, et al. 2009. Sweet sorghum for biofuel and strategies for its improvement. Information Bulletin No. 77. International Crops Research Institute for the Semi-Arid Tropics, Patancheru - 502 324, A.P, India.

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

Reddy, B.V., A.A. Kumar, and W.D. Dar. 2006. Overview of sweet sorghum breeding at ICRISAT: Opportunities and constraints. International Crops Research Institute for the Semi-Arid Tropics, Patancheru - 502 324, A.P, India.

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

Reddy, B.V., C.R. Reddy, and A.A. Kumar. 2012. Sweet Sorghum Crop Production and Management Practices. International Crops Research Institute for the Semi-Arid Tropics, Patancheru - 502 324. Andhra Pradesh, India.

Reddy, B.V.S., 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.

Renewable fuel association. 2016. Industry Statistics: World Fuel Ethanol Production. http://www.enatholrfa.org/resources/industry/stastitics/#1454098996479-8715d404-e546. (Accessed on 1 September, 2016).

Ribaut, J., M. De Vicente, and X. Delannay. 2010. Molecular breeding in developing countries: challenges and perspectives. Current Opinion in Plant Biology 13:213-218.

Ribeiro Trindade, A.P., R.J. Barth Pinto, A.T.d. Amaral Júnior, C.A. Mangolin, S. Machado, and C.A. Scapim. 2010. Genetic diversity of breeding popcorn lines determined by SSR markers. Electronic Journal of Biotechnology 13:4-5.

Rose, S., and F.C. Botha. 2000. Distribution patterns of neutral invertase and sugar contentin sugarcane internodal tissues. Plant Physiology and Biochemistry 38:819-824.

Saha, M.C., H.S. Bhandhari, and J.H. Bouton. 2013. Bioenergy Feedstocks: Breeding and Genetics. John Wiley & Sons, New York.

Schulman, A.H. 2007. Molecular markers to assess genetic diversity. Euphytica 158:313-321.

Shiringani, A.L., M. Frisch, and W. Friedt. 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.

Singh, B., and A.K. Singh. 2015. Marker-assisted plant breeding: principles and practices. Springer, India.

Tarpley, L., and D.M. Vietor. 2007. Compartmentation of sucrose during radial transfer in mature sorghum culm. BMC Plant Biology 7:33-42.

Teka, H.B. 2014. Advance research on Striga control: A review. African Journal of Plant Science 8:492-506.

Timilsina, G.R., and A. Shrestha. 2011. How much hope should we have for biofuels? Energy 36:2055-2069.

Tomar, S.S., S. Sivakumar, and K. Ganesamurthy. 2012. Genetic variability and heritability studies for different quantitative traits in sweet sorghum [*Sorghum bicolor* (L.) Moench] genotypes. Electronic Journal of Plant Breeding 3:806-810.

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.

Vieira, E.A., F.I.F.d. Carvalho, I. Bertan, M.M. Kopp, P.D. Zimmer, G. Benin, et al. 2007. Association between genetic distances in wheat (*Triticum aestivum* L.) as estimated by AFLP and morphological markers. Genetics and Molecular Biology 30:392-399.

Vinutha, K., L. Rayaprolu, K. Yadagiri, A. Umakanth, J. Patil, and P.S. Rao. 2014. Sweet sorghum research and development in India: Status and prospects. Sugar Technology 16:133-143.

Wang, F., and C.-Z. Liu. 2009. Development of an economic refining strategy of sweet sorghum in the Inner Mongolia region of China. Energy & Fuels 23:4137-4142.

Wiley, E.O. 1981. Phylogenetics: theory and practice of phylogenetic systematics. John Wiley, New York.

Zhan, Q., L. Zhou, N. Bi, H. Wu, J. Li, J. Lu, et al. 2012. A comprehensive analysis of genetic diversity and relationships of 13 sweet sorghum varieties as energy sources. Journal of Sustainable Bioenergy Systems 2:86-91.

Zhao, Y.L., A. Dolat, Y. Steinberger, X. Wang, A. Osman, and G.H. Xie. 2009. Biomass yield and changes in chemical composition of sweet sorghum cultivars grown for biofuel. Field Crops Research 111:55-64.

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

2 Assessment of sweet sorghum lines for agronomic performance and genetic diversity using quantitative morphological traits

Abstract

Twenty-five sweet sorghum lines [*Sorghum bicolor* (L.) Moench] collected from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and the African Centre for Crop Improvement (ACCI-South Africa) were evaluated during the 2015/2016 season at Ukulinga Research Farm in Pietermaritzburg and Makhatini Research Station in Jozini. The objectives of this study were; (i) to assess the extent of genetic diversity among sweet sorghum lines for stalk yield and sugar-related traits using quantitative traits, (ii) to estimate the strength of associations between stalk yield and sugar-related traits, and (iii) to determine the best performing sweet sorghum lines with desirable quantitative traits. The 25 sweet sorghum lines were evaluated in a 5 x 5 balanced lattice design with three replications. Seven agronomic traits; fresh biomass yield, fresh stalk yield, grain yield, plant height, stalk diameter, panicle length and days to 50% flowering and six quality traits; fibre, dry matter, ^{o}brix, ^{o}total brix, total fermentable sugars and ethanol were recorded. Analysis of variance per test location revealed highly signficant differences (P≤0.01) among genotypes, indicating the high level of genetic variability among the sweet sorghum genotypes studied. High levels of trait heritability were observed for fresh stalk yield (98%), stalk diameter (93%), fresh biomass yield (81%), panicle length (76%), fiber (73%) and plant height (66%). Fresh stalk yield showed positive and highly significant (P≤0.01) correlation with ^{o}total brix (r = 0.85, P≤0.05), total fermentable sugars (r = 0.85, P≤0.05), plant height (r = 0.75, P≤0.05) and days to 50% flowering (r = 0.53, P≤0.05). There was also positive and significant (P≤0.05) correlation between fresh stalk yield with dry matter (r = 0.15, P≤0.05) and ethanol (r = 0.85, P≤0.05). The first three principal components (PC) from the principal component analysis showed 83% of the total variability among the genotypes. Ethanol, total fermentable sugars, ^{o}total brix, fresh stalk yield and ^{o}brix contributed mainly to PC 1, whereas fresh biomass yield and stalk diameter contributed mainly to PC 2. Cluster analysis for phenotypic traits showed a clear variation between sweet sorghum genotypes. Based on the measured traits, the dendrogram divided the genotypes into two main clusters and three singletons (ICSB 324, ICSB 654 and ICSV 700). Cluster I comprised 54% of the total germplasm and included only one ACCI genotype (SS 27), while cluster II comprised of 33% of the total variation. Genotypes IS 2331, IESV 92008 DL and ICSV 700 exhibited high quality traits and were comparable to the standard checks (AS 244, URJA and SS 27), while genotype IESV 91018 LT showed the highest yield components and stalk diameter.

2.1 Introduction

The rise in industry development and population growth has led to an increase in demand for energy globally. This has further caused a depletion of non-renewable energy and high carbon dioxide emissions which contribute to global warming. To address these concerns, there is need to search for alternative sources of fuel (Archer, 2005; Mutepe et al., 2012). In the energy sector, biofuels have received much attention because of their ability to supplement fossil fuel supplies. The most common biofuel crops include sugarcane, grain sorghum and maize. In South Africa, sugarcane has mostly been used for biofuel production to supply sugar energy requirements. On the other hand, maize has been banned from use for bioenergy due food security threats (Blanchard et al., 2011). Even with the use of the mentioned biofuel crops, there is little bioethanol production in South Africa (Blanchard et al., 2011). Sweet sorghum [*Sorghum bicolor* (L.) Moench; 2n = 2x = 20] has therefore been considered as a promising crop to supplement sugarcane for the production of renewable and sustainable energy (Cartwright, 2007; Tsuchihashi and Goto, 2008). Besides its ability to thrive in marginal environments with low water requirements, the crop also contains higher fermentable sugars in its stalks than sugarcane and is therefore best suited for bioethanol production (Antonopoulou et al., 2008; Reddy et al., 2008; Tsuchihashi and Goto, 2008).

Genotypes are known to respond differently in various environments. Environmental effects on sweet sorghum genotype performance is therefore, an important factor to consider when exploiting existing phenotypic and genetic variability, and the development of improved cultivars (Faisal and Aisha, 2011; Abubakar and Bubuche, 2013). According to Allard and Bradshaw (1964), the environmental effect on the relative performance of genotypes has a significant bearing on the breeding efficiency. Different studies have demonstrated that the environment has a high influence on nearly all phenotypic traits (Murray et al. 2008, Shiringani et al. 2010). Selection by plant breeders has been based mainly on additive main effects, ignoring genotype by environment interaction (Babić et al., 2011). In South Africa, there is distinct variation in terms of biotic and abiotic stresses in different environments. Various factors such as rainfall, temperature and seasonal length can therefore have an impact on qualitative and quantitative traits, relative performance, and influence the choice of selection for high yielding and stable genotypes (Yan and Hunt, 1998; Gurmu et al., 2009). To address the problem, plant breeders use different environments and seasons for proper selection based on the genotype's phenotypic value (Ilker, 2011).

Sweet sorghum breeding programs can be enhanced in South Africa by understanding the relationship between traits through correlation analysis (Iyanar et al., 2010). Correlation is a technique that explains the degree of linear relationships between quantitative traits without regard to cause and effect (Malik et al., 2005). Breeders want to develop new cultivars that will outperform existing ones with respect to a number of specified traits. Stalk yield being a complex quantitative trait, is influenced by a number of environmental and genetic factors (Bocanski et al., 2009). During the selection of this trait, it is therefore vital to confirm relationships between traits that contribute to improved stalk yield in different environments. Different researchers have reported that stalk yield is significantly positively correlated with plant height, days to 50% flowering and sugar content (Murray et al., 2008; Shiringani and Friedt, 2009; Audilakshmi et al., 2010). It is therefore important that selection for stalk yield and sugar content should focus on plant height and days to 50% flowering. Research carried out by Rao et al. (2013b) demonstrated a positive relationship between stalk yield and sugar-related traits by employing correlation selection strategies on °brix content and total soluble sugars. On the other hand, a significant negative correlation has been reported between grain yield and stalk biomass. This has shown that grain yield is generally low in sweet sorghum varieties with high sugar content (Murray et al., 2008; Shiringani and Friedt, 2009; Audilakshmi et al., 2010). Cluster and principal component (PCA) analysis are multivariate techniques that are commonly used to visually observe genetic relationships among genotypes in genetic diversity studies. According to Bertan et al. (2007), the aim of the clustering algorithm is to separate genetic material into homogenous groups, such that the within-group similarities are larger compared to the between-group similarities. On the other hand, the principal components represent the patterns encoding the highest variance in the data set and not to maximize the separation between groups of genotypes directly.

Currently, the sweet sorghum lines being assessed in the study have not been fully exploited under South African conditions. It is, therefore, important to characterize and assess the extent of genetic diversity among sweet sorghum genotypes in order to select superior genotypes for stalk yield and sugar-related traits. This will be important for the development of hybrids needed in the South African biofuel industry. Twenty-five sweet sorghum lines were assessed in this study. These included 22 lines obtained from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Kenya and three lines (standard checks) from the African Centre for Crop Improvement (ACCI) of the University of KwaZulu-Natal, South Africa. The objectives of this study were; (i) to assess the extent of genetic diversity among sweet sorghum lines for stalk yield and sugar-related traits using quantitative traits, (ii) to estimate the strength of

associations between stalk yield and sugar-related traits, and (iii) to determine the best performing sweet sorghum lines with desirable quantitative traits.

2.2 Materials and methods

2.2.1 Sweet sorghum germplasm collection

A total of 25 sweet sorghum lines (Table 2-1) were evaluated during the 2015/2016 cropping season. Twenty-two sweet sorghum lines were sourced from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and multiplied to increase seed at the University of KwaZulu-Natal (UKZN) tunnels before field evaluation. Three standard checks were obtained from the African Centre for Crop Improvement (ACCI) at the University of KwaZulu-Natal, South Africa.

Table 2-1: List of sweet sorghum lines used in the study

No.	Name	Origin
1	AS 244 (local check)	ACCI-South Africa
2	E 36-1	ICRISAT-Kenya
3	Ent # 64DTN	ICRISAT-Kenya
4	ICSB 324	ICRISAT-Kenya
5	ICSB 654	ICRISAT-Kenya
6	ICSR 93034	ICRISAT-Kenya
7	ICSV 700	ICRISAT-Kenya
8	ICSV 93046	ICRISAT-Kenya
9	IESV 91104 DL	ICRISAT-Kenya
10	IESV 91018 LT	ICRISAT-Kenya
11	IESV 92001 DL	ICRISAT-Kenya
12	IESV 92008 DL	ICRISAT-Kenya
13	IESV 92021 DL	ICRISAT-Kenya
14	IESV 92028 DL	ICRISAT-Kenya
15	IESV 92165 DL	ICRISAT- Kenya
16	IESV 94021 DL	ICRISAT-Kenya
17	IS 2331	ICRISAT-Kenya
18	Kari Mtama 1	ICRISAT-Kenya

No.	Name	Origin
19	MR # 22 x IS 8613/1/2/5-2-1	ICRISAT-Kenya
20	MR # 22 x IS 8613/2/3-1-3	ICRISAT-Kenya
21	NTJ 2	ICRISAT-Kenya
22	S 35	ICRISAT-Kenya
23	URJA (local check)	ACCI-South Africa
24	SPV 1411	ICRISAT-Kenya
25	SS 27 (local check)	ACCI-South Africa

2.2.2 Experimental sites

The sweet sorghum lines were planted at Ukulinga Research Farm (URF) and Makhatini Research Station (MRS) as shown in Figure 2-1. The two locations have different environmental conditions as described in Table 2-2.

Figure 2-1: Selected sites across KwaZulu-Natal (KZN) South Africa

Table 2-2: Description of experimental sites

Site	Env. codes	Lat.	Long.	Alt. (m)	Ave. annual RF (mm)	Ave. annual temp (°C)	Soil type	pH
Ukulinga	12UKL	29° 67'S	30° 41'E	809	689	17.9	Westleigh	4.32
Makhatini	10MAK	27° 39'S	32° 17'E	77	569	22.3	fluvisols	5.52

Source: AGROMET (2016)

Data on average monthly rainfall pattern (Rf), average total relative evapotranspiration (ETO), average maximum/minimum relative humidity (RHx, RHn) and average maximum/minimum temperatures (Tx, Tn) for the two sites for 2015/2016 season was obtained from Agricultural Research Council, Pretoria (AGROMET, 2016) as indicated Figure 2-2 and Figure 2-3.

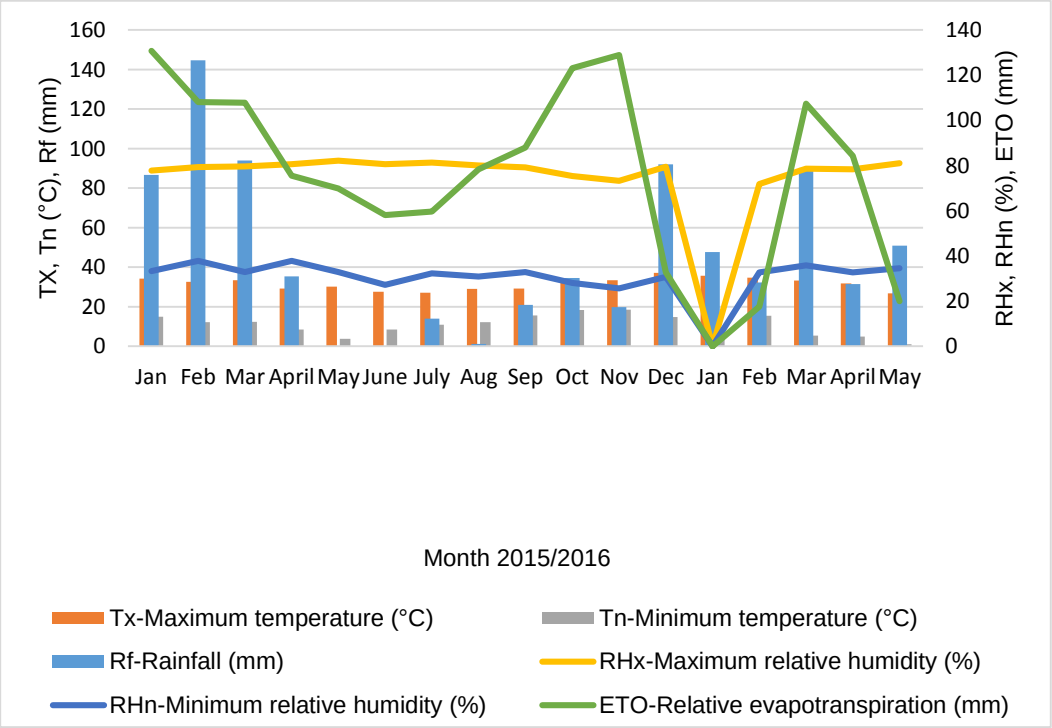


Figure 2-2: The weather variables (average maximum and minimum temperature (°C), average monthly rainfall pattern (Rf), average maximum/minimum relative humidity (RHx, RHn) and average total relative evapotranspiration (ETO)) for 2015/2016 season at Makhatini Research Station

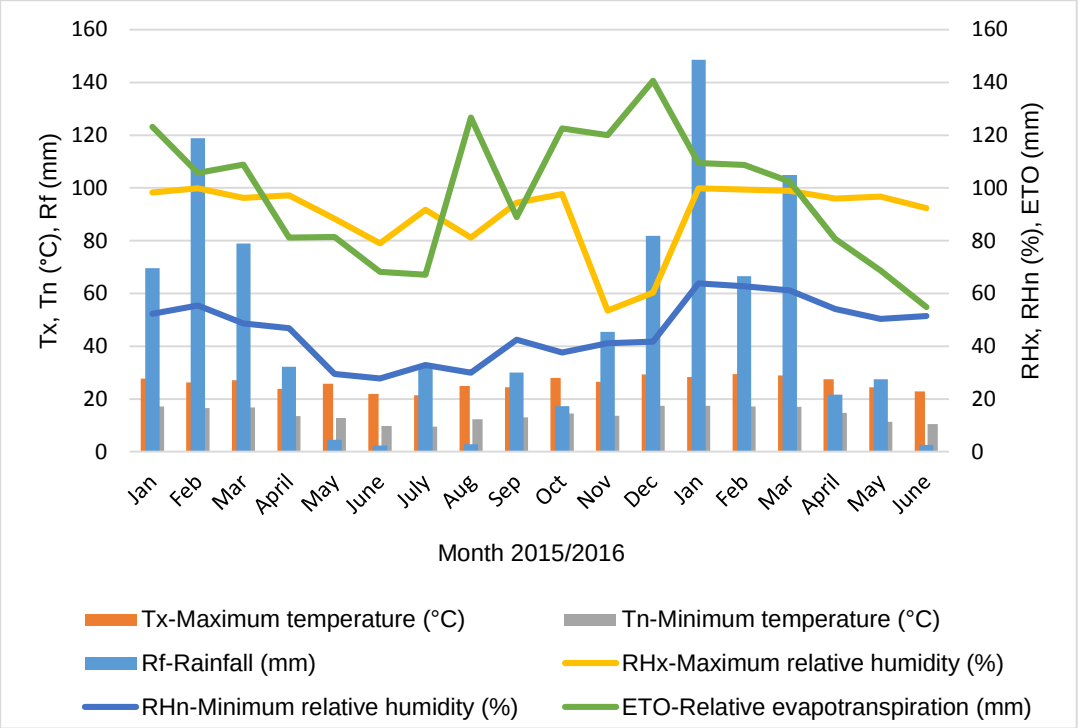


Figure 2-3: The weather variables (average maximum and minimum temperature (°C), average monthly rainfall pattern (Rf), average maximum/minimum relative humidity (RHx, RHn) and average total relative evapotranspiration (ETO)) for 2015/2016 season at Ukulinga Research Farm

2.2.3 Experimental design and management

The trials were planted on the 9th of January 2016 at Ukulinga Research Farm (URF) and the 7th of March 2016 at Makhatini Research Station (MRS). The 25 sweet sorghum lines were laid out in a 5 x 5 balanced lattice design with three replications. The land had relatively uniform slope but blocking was performed to minimize spatial effects and to ensure that the treatment effects were uniform. In both trial sites, each experimental plot had 4 rows of 5 m length with inter-row and intra-row spacing of 0.80 m and 0.3 m (17" plants/row), respectively. The gross plot was 16 m² while the net plot was 8 m². Sowing was done by hand dibbing of seeds with 4 seeds per hill after land preparation. The seedlings were thinned three weeks after planting when the plants were fully established. A basal fertilizer (N:P:K ratio = 2:3:4) was applied at a rate of 250 kg ha^{-1} while Lime Ammonium Nitrate (28% N) was top-dressed at a rate of 200 kg ha^{-1}. The plots were kept weed free by hand hoeing and using gramoxone herbicide. Seedling damage by cutworms or mice was prevented by applying Curater at planting. The stalk borers were controlled by using stalk borer granules (dimethyl-(2, 2, 2-trichloro-1-hydroxyethyl)

phosphonate. Mesh bags were used to cover the sweet sorghum heads at anthesis to prevent seed damage by birds. The trials at both locations were rainfed and supplemented by irrigation.

2.2.4 Data collection

Agronomic characteristics were evaluated based on sorghum descriptors (International Board for Plant Genetic Resources/International Crops Research Institute for the Semi-Arid Tropics - (IBPGR/ICRISAT, 1993). The days to 50% flowering (days from planting to half of the plants in a plot to reach anthesis) was recorded by visual assessment. Stalk diameter was measured using a measuring tape on the three mid-internode sections. Plant height was measured in cm with a graduated measuring stick from the ground to the tip of the panicle of ten plants randomly selected. Ten sweet sorghum plants were harvested at hard dough stage by cutting at ground level using a wire cutter scissors to measure fresh biomass and stalk yield. Fresh biomass was measured by weighing ten whole plants (stalks, leaves and panicles) using a measuring scale. Fresh stalk yield was measured by weighing the ten stalks after stripping off all leaves and panicles. Panicle length was measured from the lower branch to the tip of panicle. Grain yield was measured in grams after drying the grain to a moisture content of 12.5% then converted to t ha^{-1} as shown in Equation 2-1.

$$Y = h \times 1/w \times p \qquad \text{Equation 2-1}$$

Where Y = grain yield (t ha^{-1}); h = average head weight (total weight in sample of heads taken (g)/number of heads); w = unit of weight (1,000,000 tons); p = plant population [10,000 m^2 ha^{-1}/seed spacing in the row (m) x row width (m)].

Laboratory analysis (wet chemistry technique) at hard dough stage was used to measure; % fiber, % dry matter and °brix. Sweet sorghum samples of about 2 kg from each plot were obtained from the chaff cutter, mixed thoroughly and placed in labelled plastic packets for lab analysis. For dry matter analysis, an empty container was weighed after zeroing the scale. Approximately 200 g of sample was added for dry matter analysis and data obtained was recorded. The dry matter sample was placed in the drying oven and reweighed after removal from the oven. For °brix analysis, approximately 1 kg of chopped stalk was obtained from the plastic packet and placed in the cold digester bowl. Water equal to exactly double the weight of stalk was added, placed on the digester and run for 20 minutes. The bowl was carefully removed from the digester and 150 ml of liquid was poured into a beaker through a fine sieve. The beaker was stirred and about 50 ml of digester sample was poured into the funnel containing celite powder. Three °brix readings were recorded using the refractometer when the

sample had cooled to approximately 20°C and the °brix average was recorded. °Total brix ha^{-1}, total fermentable sugars ha^{-1} and ethanol (l ha^{-1}) were calculated from weight of stalks plot^{-1}.

2.3 Data analysis

2.3.1 Analysis of variance (ANOVA)

Genotype analysis was performed on the data obtained using a linear mixed model (where genotypes were fixed, environments, blocks and replications were random). The parameters were estimated using PROC MIXED in SAS version 9.3 (SAS, 2011). The following model for the combined ANOVA was used as shown in equation Equation 2-2.

$$Y_{ijk} = \mu + g_i + e_j + (ge)_{ij} + b(e)_{jk} + \varepsilon_{ijk} \qquad \text{Equation 2-2}$$

Where Y_{ijk}= the phenotype of the *i*th genotype in the *j*th environment and *k*th block, μ = the overall phenotypic mean, g_i = the effect due to *i*th genotype, e_j = the effect due to *j*th environment, $(ge)_{ij}$ = the effect due to the interaction of the *i*th genotype and the *j*th environment, $b(e)_{jk}$ = the effect due to *k*th block nested within the *j*th environment and ε_{ijk} = the random effect associated with the *ijk*th observation.

2.3.2 Variance Components and heritability of traits

2.3.2.1 Variance components

The variance components attributed to G (δ²*g*), E (δ²*e*), and GE (δ²*ge*), were estimated in SAS mixed model using the COVTEST (SAS, 2011).

2.3.2.2 Broad-sense heritability (*H*²)

Broad-sense heritability (*h*²) of each trait was estimated based on random genotypes according to the following equation:

$$h^2 = \sigma_g^2/\sigma_p^2 \qquad \text{Equation 2-3}$$

Where σ_g^2 = variance component of genotype; σ_p^2, phenotypic variance among genotypes grown in r replicates and n environments can be expressed as $\sigma_g^2 + \sigma_{ge}^2/n + \sigma_e^2/nr$; σ_{ge}^2 = the

variance component of the interaction genotype x environment; σ_e^2 = experimental error (Gao, 1986).

2.3.3 Correlation Analysis

Phenotypic correlation analysis was employed to estimate the degree of association among phenotypic traits using PROC CORR in SAS version 9.3 (SAS, 2011).

2.3.4 Principal Component Analysis

To determine the traits that accounted for the most variation between lines, principal component analysis using a correlation matrix based on least square means (LSMEAN) was executed using PROC PRINCOMP in SAS version 9.3 (SAS, 2011).

2.3.5 Cluster Analysis

The measured variables were standardized using the standard deviation of mean by PROC STANDARD and then used for clustering by PLOC CLUSTER with average linkage based on Euclidean distance of the standardized variables (Flury and Riedwyl, 1986). Dendrograms were constructed using PROC TREE in SAS version 9.3 (SAS, 2011).

2.4 Results

2.4.1 Analysis of variance (ANOVA)

Analysis of variance per test location revealed highly signficant differences (P≤0.01) among genotypes with respect to all the measured variables. Therefore, a combined analysis of variance was conducted by combining data from the two locations. The results showed highly signficant (P≤0.01) genotype by environment interactions with respect to grain yield, dry matter, °brix, plant height, panicle length and days to 50% flowering. There was also signficant (P≤0.05) genotype by environment interaction with respect to fresh biomass yield, °total brix, total fermentable sugars and ethanol yield (Table 2-3). However, the combined analysis of variance showed non-significant genotype by environment interaction with respect to fresh stalk yield, fiber and stalk diameter. The effect of environment was significant with respect to all traits except fresh biomass yield and fresh stalk yield (Table 2-3).

2.4.1.1 Yield components

Significant variation for fresh biomass yield was observed using the combined means for both locations (Table 2-3). Fresh biomass yield ranged from 15 to 36 t ha^{-1} for ICSB 654 and IESV 91018 LT respectively and had a mean value of 25 t ha^{-1}. At Ukulinga Research Farm, fresh biomass yield ranged from 18 to 33 t ha^{-1} for ICSB 324 and IESV 91018 LT with a mean value of 26 t ha^{-1}. Makhatini Research Station exhibited a mean value of 25 t ha^{-1} with a range between 12 to 38 t ha^{-1} for ICSB 654 and IESV 91018 LT (Table 2-4). The genotypes IESV 91018 LT and ICSV 93046 showed consistency in both locations with highest fresh stalk yield and fresh biomass yield. The combined means for both locations showed that fresh stalk yield ranged between 9 to 25 t ha^{-1} for ICSB 654 and IESV 91018 LT respectively with a mean value of 16 t ha^{-1}. Fresh stalk yield at Ukulinga Research Farm had a mean value of 17 t ha^{-1}, with a range between 11 to 24 t ha^{-1} for MR # 22xIS 8613/2/3-1-3 and IESV 91018 LT respectively. Makhatini Research Station exhibited a mean value of 16 t ha^{-1} and a range between 7 to 26 t ha^{-1} for ICSB 654 and IESV 91018 LT (Table 2-4). The combined means for both locations exhibited grain yield ranging from 0.3 to 2.2 t ha^{-1} for ICSB 654 and IESV 92165 DL with a mean value of 1.0 t ha^{-1} (Table 2-3). Makhatini Research Station had significant higher grain yield ranging between 0.5 to 4.0 t ha^{-1} for ICSV 93046 and IESV 92165 LT, with a mean value 2.0 t ha^{-1} than Ukulinga Research Farm which had grain yield ranging between 0.1 to 0.9 t ha^{-1} for ICSB 324 and IESV 91018 LT with a mean value 0.4 (Table 2-4).

2.4.1.2 Quality traits

Varieties with the highest °brix also showed high fibre, dry matter, °total brix, total fermentable sugars and ethanol. The same trend was exhibited among varieties with low °brix. Among the genotypes, IESV 91018 LT, MR # 22 x IS 8613/2/3-1-3 and ICSB 654 were observed to have the least °brix, fibre, dry matter, °total brix, total fermentable sugars and ethanol. On the other hand URJA, SS 27 and ICSV 700 had the highest °brix and its related traits. For both locations, the means for °brix ranged between 8 to 14% and had a mean value of 11%. At Ukulinga Research Farm and Makhatini Research Station, °brix ranged between 9 to 15% and 6 to 14% with mean values of 13% and 9% respectively (Table 2-4). The means for each location and combined locations for fibre ranged between 9 to 15%, while the overall mean values were 12%, 12% and 11% for the combined locations, Ukulinga Research Farm and Makhatini Research Station, respectively. For combined locations, the means for dry matter ranged between 17 to 30% with a mean value of 23%. At each location, Ukulinga Research Farm showed dry matter ranging between 18 to 29% while MRS had a range between 15 to 29% with a mean value of 26% and 20% respectively (Table 2-4).

Table 2-3: Mean, range, standard error (SE) and mean squares (MS) on 13 quantitative traits of 25 sweet sorghum lines evaluated at Ukulinga Research Farm (URF) and Makhathini Research Station (MRS)

	Traits												
Genotype	**FBY**	**FSY**	**GY**	**FBR**	**DM**	**°B**	**°TB**	**TFS**	**ETH**	**PH**	**DMT**	**PL**	**FL**
AS 244	25.27	15.19	1.26	10.16	22.26	11.72	1.72	1.46	850.61	181.14	6.08	24.26	81.12
E 36-1	20.92	12.18	1.29	11.22	21.66	10.07	1.24	1.05	613.38	155.14	5.89	22.59	75.64
Ent # 64DTN	21.37	12.25	1.43	12.41	22.80	9.88	1.22	1.04	606.18	153.78	6.72	21.49	77.30
ICSB 324	21.65	13.31	1.21	14.04	25.29	10.87	1.39	1.19	689.58	142.31	5.77	30.97	85.62
ICSB 654	15.37	9.24	0.26	12.96	23.44	9.90	0.99	0.84	488.12	150.04	4.98	24.45	72.28
ICSR 93034	23.69	14.28	0.98	12.69	24.07	10.99	1.57	1.34	777.04	166.91	6.73	21.23	80.30
ICSV 700	24.86	17.43	0.29	11.48	23.85	12.00	2.05	1.74	1014.26	259.92	6.88	22.06	88.31
ICSV 93046	31.76	22.96	0.58	12.50	24.44	11.42	2.56	2.18	1266.73	224.16	6.62	18.78	88.18
IESV 91104 DL	27.99	17.28	1.91	13.78	25.50	11.07	1.95	1.66	966.18	174.47	6.46	21.52	78.54
IESV 91018 LT	35.79	25.20	1.92	9.05	17.08	7.55	2.05	1.74	1014.24	219.89	8.03	24.74	85.75
IESV 92001 DL	28.16	18.57	1.75	12.25	23.45	10.56	2.03	1.73	1004.69	174.58	6.80	21.76	78.39
IESV 92008 DL	27.19	17.59	1.58	13.25	25.33	11.55	2.09	1.77	1032.42	171.42	6.89	21.68	79.50
IESV 92021 DL	23.40	14.59	1.26	9.88	20.71	10.33	1.61	1.37	795.37	154.62	6.70	21.30	79.55
IESV 92028 DL	27.85	17.11	1.45	10.73	22.28	11.01	1.94	1.65	961.10	162.2	6.31	23.51	85.97
IESV 92165 DL	28.59	17.92	2.20	11.15	21.86	10.22	1.83	1.56	907.14	161.89	6.61	22.17	74.03
IESV 94021 DL	24.55	12.82	1.62	10.20	20.62	10.12	1.25	1.07	619.21	136.26	5.95	21.70	77.73
IS 2331	27.39	19.84	1.37	12.02	25.40	12.82	2.56	2.18	1265.88	238.67	5.85	20.58	85.75

	Traits												
Genotype	**FBY**	**FSY**	**GY**	**FBR**	**DM**	**°B**	**°TB**	**TFS**	**ETH**	**PH**	**DMT**	**PL**	**FL**
Kari Mtama 1	26.62	15.89	1.92	10.95	20.96	9.57	1.53	1.30	756.68	150.51	7.86	25.28	76.01
MR # 22xIS 8613/1/2/5-2-1	20.72	12.82	0.98	14.35	25.08	10.12	1.41	1.19	694.41	160.01	6.58	19.89	82.63
MR # 22xIS 8613/2/3-1-3	21.34	11.19	0.94	9.26	18.37	8.62	0.92	0.78	452.33	111.56	7.66	20.96	82.62
NTJ 2	25.58	16.38	1.26	11.26	20.77	9.26	1.55	1.32	766.06	189.15	6.14	22.85	80.12
S35	-	-	-	-	-	-	-	-	-	-	-	-	-
URJA	24.88	19.56	0.31	12.57	27.03	13.91	2.79	2.37	1380.17	247.02	6.57	21.70	87.02
SPV 1411	23.82	17.88	0.50	12.61	23.03	9.84	1.83	1.56	905.95	220.45	6.59	17.90	85.21
SS 27	25.03	20.48	0.32	14.90	29.69	14.08	2.89	2.45	1428.44	242.31	5.74	16.63	80.79
Mean	25.17	16.33	1.19	11.90	23.13	10.73	1.79	1.52	885.6	181.17	6.5	22.08	81.17
Range	15-36	9-25	0.26-2.20	9-15	17-30	8-15	1-3	1-3	452-1428	112-260	5-8	17-31	72-89
SE	1.34	0.97	0.16	0.21	0.30	0.15	0.09	0.08	45.29	2.46	0.04	0.12	0.58
Genotype MS	82.90**	68.68**	1.73**	13.25**	39.72**	12.73**	1.45**	1.05**	355149.6**	1819.75**	2.14**	38.45**	101.25**
Environment MS	53.61	33.44	86.91**	38.74**	1073.05**	702.25**	24.78**	17.90**	6081131**	172863.77**	1.64**	42.61**	793.36**
G X E MS	34.87*	13.48	0.86**	2.86	9.58**	4.92**	0.40*	0.29*	98303.63*	647.51**	0.18	7.18**	37.18**

**,* significant at P≤0.01 and P≤0.05, respectively; FBY, fresh biomass yield (t ha^{-1}); FSY, fresh stalk yield (t ha^{-1}); GY, grain yield (t ha^{-1}); FBR, fibre (%); DM, dry matter (%); °B, brix (%); °TB, total brix (t ha^{-1}); TFS, total fermentable sugars (% ha^{-1}); ETH, ethanol (l ha^{-1}); PH, plant height (cm); DMT, stalk diameter (cm); PL, panicle length (cm); FL, days to 50% flowering; -, missing value

The combined mean value for °total brix was 1.8 t ha^{-1} and had a range between 1.0 to 3.0 t ha^{-1}. At each location, Ukulinga Research Farm showed °total brix ranging between 1.0 to 3.0 t ha^{-1} with a mean value of 2.2 t ha^{-1}, while Makhatini Research Station had a range between 0.5 to 3.0 t ha^{-1} with a mean value of 1.4 t ha^{-1}. Total fermentable sugars ha^{-1} ranged between 0.8 to 2.5 t ha^{-1} and had a mean value of 1.5 t ha^{-1} using the combined means for both locations. At Ukulinga Research Farm, total fermentable sugars ha^{-1} ranged between 1.0 to 3.0 t ha^{-1} with a mean value of 1.9 t ha^{-1}, while Makhatini Research Station showed a range between 0.4 to 2.0 t ha^{-1} with a mean value of 1.2 t ha^{-1}. Ethanol ranged between 452 to 1428 l ha^{-1} and had a mean value of 885.6 l ha^{-1} using the combined means of both locations. At Ukulinga Research Farm, ethanol ranged between 535 to 1540 l ha^{-1} and had a mean value of 1091 l ha^{-1}, while Makhatini Research Station showed a range from 258 to 1433 l ha^{-1} with a mean value of 679 l ha^{-1} (Table 2-5).

2.4.1.3 Flowering, plant height, stalk diameter and panicle length

Plant height ranged from 112 to 260 cm for MR # 22 x IS 8613/2/3-1-3 and ICSV 700, with a mean value of 181 cm using the combined means for both locations (Table 2-3). The shortest and tallest varieties were observed to be MR # 22 x IS 8613/2/3-1-3 and ICSV 700, with a height ranging from 119 to 273 cm and had a mean value of 197 cm at Ukulinga Research Farm. Makhatini Research Station showed that plant height ranged from 104 to 247 cm with a mean value of 166 cm (Table 2-5). Stalk diameter is another trait that contributes to sugar content. Stalk diameter showed a consistent range from 5 to 8 cm, with a mean value of 7 cm using the combined means for both locations and in the individual analysis per test location. The genotype with the thinnest stalk was ICSB 654, while 91018 LT had the thickest stalk (Table 2-5). The combined means for both locations showed that panicle length ranged between 17 to 31 cm for SS 27 and ICSB 324 with the mean value of 22 cm. The mean values of panicle length of the genotypes were significantly higher at Ukulinga Research Farm (23 cm) than at Makhatini Research Station (22 cm). Although the panicle length among genotypes was kept mostly constant at the two locations, Ukulinga Research Farm showed a range from 17 to 30 cm for SS 27 and ICSB 324, and Makhatini Research Farm showed a range between 16 to 32 cm for SS 27 and ICSB 324 (Table 2-5). The combined means for both locations showed days to 50% flowering ranging from 72 to 88 days for ICSB 654 and ISCV 700 with a mean value of 81 days (Table 2-3). At Ukulinga Research Farm, days to 50% flowering ranged from 76 and 102 days with a mean value of 84 days (Table 2-5). However, at Makhatini Research Station, days to 50% flowering ranged from 69 and 91 days with a mean value of 79 days.

Table 2-4: Mean, range and standard error (SE) on 13 quantitative traits of 25 sweet sorghum lines evaluated at Ukulinga Research Farm (URF) and Makhathini Research Station (MRS)

	FBY		FSY		GY		FBR		DM		°B		°TB	
Genotypes	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**
Mean	24.56**	25.76**	16.82**	15.84**	0.42**	1.97**	12.42**	11.38**	25.86**	20.40**	12.9**	8.52**	2.20**	1.37**
Range	18-33	12-38	11-24	7-26	0.05-0.86	0.50-3.75	9-15	9-15	18-29	15-29	9-15	6-14	1-3	0.5-3
SE	1.88	0.51	1.33	0.41	0.02	0.22	0.23	0.17	0.37	0.23	0.18	0.14	0.13	0.04

**,* significant at P≤0.01 and P≤0.05, respectively; FBY, fresh biomass yield (t ha^{-1}); FSY, fresh stalk yield (t ha^{-1}); GY, grain yield (t ha^{-1}); FBR, fibre (%); DM, dry matter (%); °B ,brix (%); °TB, total brix (t ha^{-1}); -, missing value.

Table 2-5: Mean, range and standard error (SE) on 13 quantitative traits of 25 sweet sorghum lines evaluated at Ukulinga Research Farm (URF) and Makhathini Research Station (MRS)

	TFS		ETH		PH		DMT		PL		FL	
Genotypes	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**	**URF**	**MRS**
Mean	1.87**	1.17**	1091.17**	679.37**	196.68**	165.68**	6.41**	6.51**	22.63**	21.53**	84.36**	78.82**
Range	1-3	0.40-2	535-1540	258-1433	119-273	104-247	5-8	5-8	17-30	16-32	76-102	69-91
SED	0.11	0.03	62.92	19.60	3.45	0.67	0.05	0.04	0.13	0.15	0.78	0.25

**,* significant at P≤0.01 and P≤0.05, respectively; TFS, total fermentable sugars (t ha^{-1}); ETH, ethanol (l ha^{-1}); PH, plant height (cm); DMT, stalk diameter (cm); PL, panicle length (cm); FL, days to 50% flowering; -, missing value.

2.4.2 Variance components and heritability of traits

The effects of genotype (G), environment (E), and genotype by environment (GXE) interaction caused by variation from calculated variance components over two locations of 13 traits are summarised in Table 2-6. The values of genotypic and phenotypic variances, broad sense heritability (H^2), G X E interaction, and environment variances to phenotypic variance are presented in Table 2-6. Analysis across two locations indicated that traits; fresh biomass yield, fresh stalk yield, fibre, plant height, stalk diameter and panicle length had the highest genotypic variance components estimated at; 10.3, 11.6, 2.04, 1512.3, 0.5 and 6.3 respectively, while °brix and total fermentable sugars showed the lowest (Table 2-6). Phenotypic traits such as fresh stalk yield and stalk diameter had the highest H^2 at 98% and 93% respectively, followed by fresh biomass yield, panicle length and fibre. Characters such as °brix (8%) and total fermentable sugars (8%) exhibited the lowest heritability (Table 2-6). The proportion of G X E interaction was high in fresh biomass yield (19%), followed by total fermentable sugars and days to 50% flowering (17%). The environment had no effect (0%) on fresh biomass yield and fresh stalk yield. The results indicated that for fresh biomass yield and fresh stalk yield, 19% and 2% of the phenotypic variance was attributed to only G X E interaction respectively.

Table 2-6: Variance components and broad sense heritability values on 13 traits of sweet sorghum lines evaluated at Ukulinga Research Farm (URF) and Makhathini Research Station (MRS)

	combined over locations						
Traits	**$\sigma^2 p$**	**$\sigma^2 g$**	**$\sigma^2 e$**	**$\sigma^2 ge$**	**$\sigma^2 g / \sigma^2 p$**	**$(\sigma^2 ge/n) / \sigma^2 p$**	**$(\sigma^2 e/nr) / \sigma^2 p$**
FBY	12.77	10.28	0.00	4.98	0.81	0.19	0.00
FSY	11.90	11.63	0.00	0.54	0.98	0.02	0.00
GY	2.07	0.18	1.17	0.26	0.09	0.06	0.09
FBR	2.81	2.04	0.51	0.00	0.73	0.00	0.03
DM	28.79	5.70	14.76	1.89	0.20	0.03	0.09
°B	16.52	1.31	9.68	1.37	0.08	0.04	0.10
°TB	0.75	0.22	0.33	0.06	0.30	0.04	0.07
TFS	0.48	0.04	0.24	0.16	0.08	0.17	0.08
ETH	183432.00	53553.00	81800.00	14358.00	0.29	0.04	0.07
PH	2275.28	1512.30	470.17	115.44	0.66	0.03	0.03
DMT	0.49	0.45	0.02	0.01	0.93	0.01	0.01
PL	8.30	6.31	0.47	2.56	0.76	0.15	0.01
FL	34.44	13.27	10.15	11.89	0.39	0.17	0.05

2.4.3 Correlation analysis

Fresh biomass yield showed a positive and highly significant correlation with fresh stalk yield ($r = 0.89$, $p < 0.01$), °total brix ($r = 0.61$, $p<0.01$), total fermentable sugars ($r = 0.61$, $p<0.01$) and ethanol ($r = 0.61$, $p<0.01$) (Table 2-7). It also showed positive and significant correlation with grain yield ($r = 0.46$, $p < 0.05$), plant height ($r = 0.43$, $p<0.05$), and stalk diameter ($r = 0.50$, $p < 0.05$). Fresh stalk yield also showed positive and highly significant ($P \leq 0.01$) correlation with °total brix ($r = 0.85$, $p < 0.01$), total fermentable sugars ($r = 0.85$, $p < 0.01$), ethanol ($r = 0.46$, $p < 0.01$), plant height ($r = 0.75$, $p < 0.01$) and days to 50% flowering ($r = 0.53$, $p < 0.01$). Fresh stalk yield exhibited positive and significant ($P \leq 0.05$) correlation with stalk diameter and ethanol (Table 2-7). Days to 50% flowering showed to be negatively and significantly correlated with grain yield ($r = -0.39$, $p<0.05$) and was also negatively correlated with panicle length even though it was not significant. It was observed that the correlation coefficients between grain yield with °brix, dry matter, fiber and plant height were negative and significant ($P \leq 0.05$). There was a negative and significant correlation between °brix and stalk diameter ($r = -0.43$, $P \leq 0.05$).There was a strong association between ethanol yield and its related traits (Table 2-7). Ethanol showed a positive and highly significant ($P \leq 0.01$) correlation with total fermentable sugars ($r = 1.00$, $p < 0.01$), °total brix ($r = 1.00$, $p < 0.01$), °brix ($r = 0.72$, $p < 0.01$) and fresh biomass yield ($r = 0.61$, $p<0.01$). There was a positive and significant correlation between ethanol with dry matter ($r = 0.58$, $p < 0.01$) and fresh stalk yield ($r = 0.85$, $p<0.01$).

Table 2-7: Correlation coefficients on 13 traits of sweet sorghum lines evaluated at Ukulinga Research Farm (URF) and Makhathini Research Station (MRS)

	FL	PL	DMT	PH	ETH	TFS	°TB	°B	DM	FBR	GY	FSY	FBY
FL	1.00												
PL	-0.09	1.00											
DMT	0.21	-0.03	1.00										
PH	0.60**	-0.39	-0.02	1.00									
ETH	0.52**	-0.40	-0.02	0.84**	1.00								
TFS	0.52**	-0.40	-0.02	0.84**	1.00**	1.00							
°TB	0.52**	-0.41	-0.02	0.84**	1.00**	1.00**	1.00						
°B	0.30	-0.30	-0.43*	0.57**	0.72**	0.72**	0.72**	1.00					
DM	0.17	-0.30	-0.49*	0.45*	0.58*	0.58**	0.58**	-0.43*	1.00				
FBR	0.01	-0.20	-0.44*	0.22	0.29	0.29	0.09	0.52**	0.88**	1.00			
GY	-0.39*	0.34	0.35	-0.44*	-0.19	-0.19	-0.19	-0.44*	-0.45*	-0.35	1.00		
FSY	0.53**	-0.26	0.34	0.75**	0.85*	0.85**	0.85**	0.27	0.15*	-0.01	0.09	1.00	
FBY	0.38	-0.08	0.50*	0.43*	0.61**	0.61**	0.61**	-0.01	-0.16	-0.27	0.46*	0.89**	1.00

**,* significant at P≤0.01 and P≤0.05, respectively.

2.4.4 Principal component analysis

In the current study, the PCA analysis revealed only the first three eigenvectors had eigenvalues larger than one. These cumulatively explained 83% of the total variation among the genotypes considering all the quantitative traits (Table 2-8). The first principal component (PC 1) alone accounted for 49% total variation, mainly due to ethanol (ETH), total fermentable sugars (TFS), total brix (°TB), fresh stalk yield (FSY) and °brix (°B). Principal component two (PC2) explained 25% of the total variation with most of the variation being attributed to fresh biomass yield (FBY) and stalk diameter (DMT). The traits days to 50% flowering (FL) and grain yield (GY) explained 9% of most of the variation to PC3.

Table 2-8: Principal component analysis of 13 quantitative traits in sweet sorghum lines showing eigenvectors, eigenvalues, individual and cumulative percentage of variation explained by the first three principal components (PC) axes

Eigenvectors			
Trait	**PC1**	**PC2**	**PC3**
FL	0.23	0.12	-0.50
PL	-0.18	0.07	0.23
DMT	-0.02	0.43	-0.17
PH	0.35	0.03	-0.23
ETH	0.39	0.07	0.10
TFS	0.39	0.07	0.10
°TB	0.39	0.07	0.10
°B	0.30	-0.27	0.12
DM	0.26	-0.38	0.23
FBR	0.15	-0.39	0.27
GY	-0.13	0.33	0.63
FSY	0.32	0.31	0.07
FBY	0.20	0.45	0.21
Eigenvalue	6.39	3.2	1.17
Individual %	49.16	24.64	9.01
Cumulative %	49.16	73.79	82.81

The existence of wider phenotypic variability among the sweet sorghum genotypes studied was further explained by the principal component analysis (PCA) biplot (Figure 3-2). The PCA biplot provided an overview of the similarities and differences among genotypes and of the interrelationships between the measured variables. The biplot demarcated the genotypes with characteristics explained by the first two dimensions. The PCA grouped the genotypes into groups and remained scattered in all four quadrants based on the phenotypic traits. The genotypes IESV 91018 LT and ICSB 654 were placed at extreme positions from the origin in the PCA biplot whereas the genotypes IESV 94021 DL, AS 244 and IESV 92008 DL were concentrated around the origin on PC2.

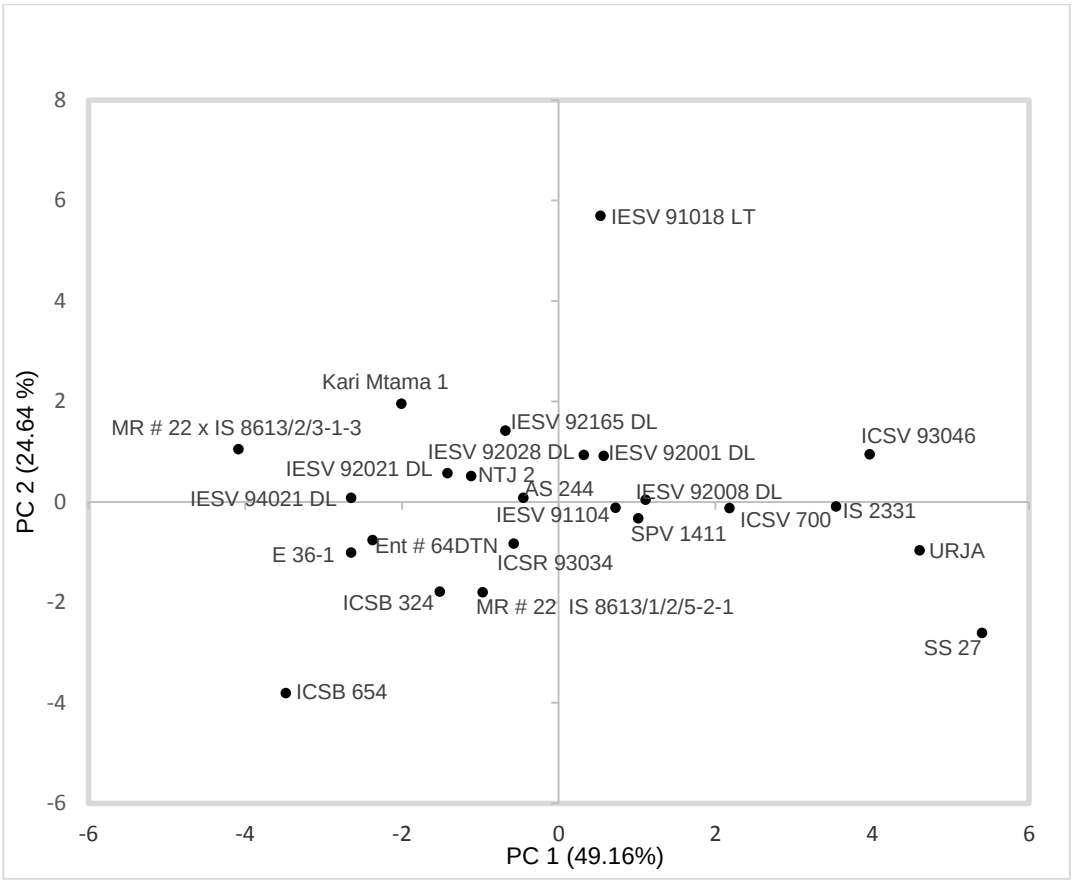


Figure 2-4: Principal component score plot of PC 1 and PC 2 describing the overall variation among 24 sweet sorghum genotypes estimated using phenotypic characteristics

2.4.5 Hierarchical clustering

Cluster analysis for phenotypic traits showed a clear variation between sweet sorghum genotypes (Figure 2-5). Based on these traits, the dendrogram divided the genotypes into two main clusters and three singletons (ICSB 324, ICSB 654 and ICSV 700) (Figure 2-5). Cluster I comprised 54% of the total germplasm and included only one ACCI genotype (SS 27). Thirteen genotypes in cluster I were namely; IESV 94021 DL, SPV 1411, IESV 92165 DL, IESV 92021 DL, IESV 92028 DL, MR # 22 x IS 8613/2/3-1-3, IESV 92001 DL, IESV 92008 DL, IESV 91104 DL, URJA, IS 2331, IESV 91018 LT and ICSV 93046. This largest group was further sub-divided into cluster A and Cluster B. The dendrogram further demarcated cluster A into two sub-clusters. Four genotypes were grouped in one of the sub-clusters with IESV 92021 DL and IESV 92165 DL being very closely related. On the other hand, IESV 92001 DL and MR # 22 x IS 8613/2/3-1-3, and IESV 91104 DL and IESV 92008 DL were observed to be closely related. Eight genotypes were clustered together in Cluster II whereby two of them (URJA and

AS 244) being ACCI materials and six; NTJ 2, Ent # 64DTN, E 36-1, MR # 22 x IS 8613/1/2/5-2-1, Kari Mtama1, ICSR 93034, SS 27 and AS 244. This group further formed 2 sub-clusters which showed Ent # 64DTN and NTJ 2, and URJA and AS 244 to be the closest related genotypes.

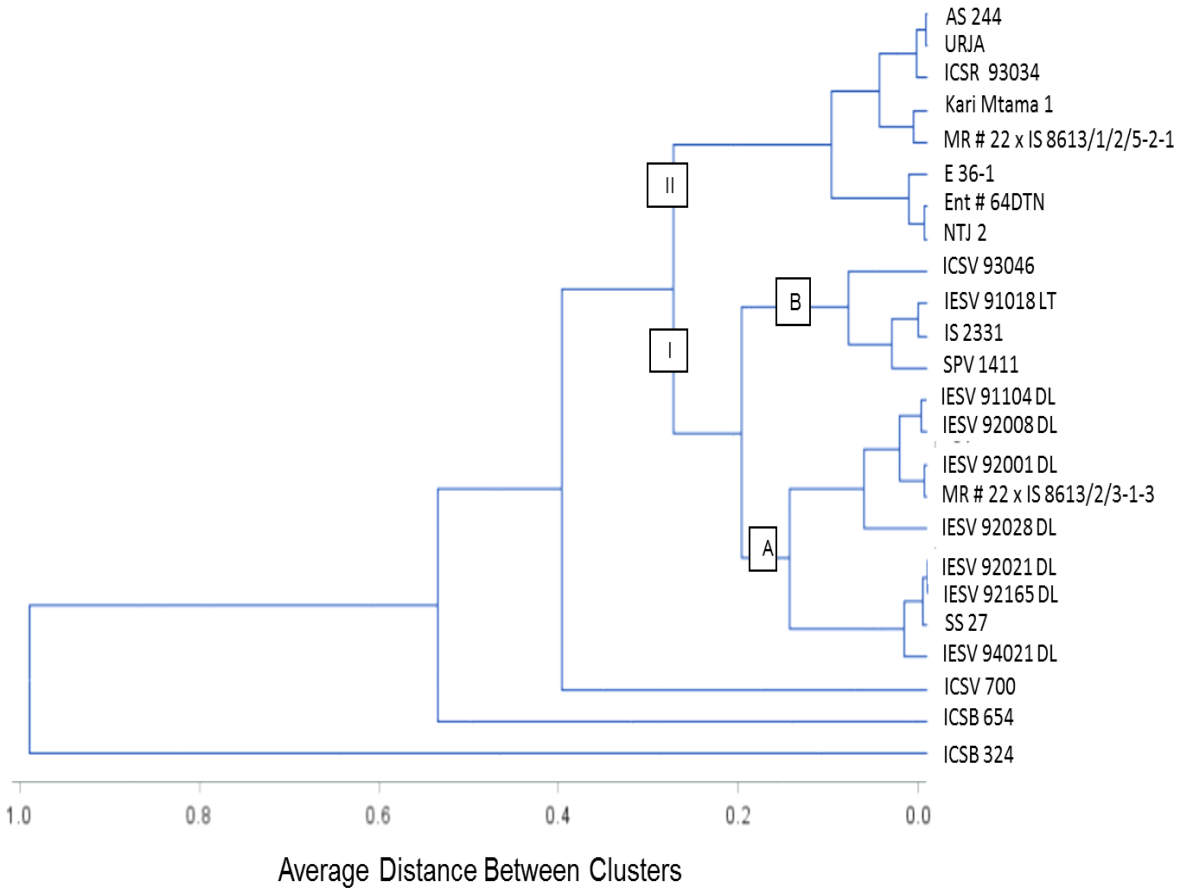


Figure 2-5: Dendrogram of 24 sweet sorghum lines generated by average distance based on 13 quantitative traits

2.5 Discussion

The highly signficant differences (P≤0.01) among genotypes with respect to all the measured variables suggests that there is considerable variation among genotypes that could be used in the development of new cultivars for bioethanol production. On the other hand, the significant genotype by environment interactions with respect to ten traits obtained from a combined analysis of variance showed that the environmental conditions in the two locations influenced the relative performance of genotypes. These results are consistent with the findings of other researchers (Zou et al., 2011; Elangovan et al., 2014). The effect of environment, also being significant for all traits except for fresh biomass yield and fresh stalk yield indicated that the

environmental conditions in the two locations were not similar and therefore genotypes showed differences in performance at the two locations. This observation is supported by Makanda (2009) who reported that quantitative traits are highly influenced by environment factors. Therefore, stalk sugar accumulation is complicated because the effects of the environment, genotype by environment and the genetic background all play a role.

Rainfall and temperature are the most important environmental factors that determine sweet sorghum performance (Shinde et al., 2013). Ukulinga Research Farm and Makhatini Research Station were both supplemented by irrigation because the field trials were planted later than the recommended time. On the other hand, Makhatini Research Station has generally been considered to have little rainfall. The delayed planting at this site may have caused a reduction in the mean values of days to 50% flowering, plant height and all traits related to plant growth, development and sugar production (Makanda et al., 2009). The lower performance observed among the genotypes could have been due to the decline in photoperiod and accumulated thermal time from planting to physiological maturity. These findings are also supported by Almodares and Darany (2006) who reported that plant height in sweet sorghum is increased with earlier planting. A study carried out by Balole (2001) at the University of Pretoria also found that late summer plantings of sweet sorghum resulted in lower stalk and sugar yields. Early planting, therefore, could result in higher plant height, a greater number of tillers and higher stalk yield.

Although there was a decline in the performance of the genotypes at Ukulinga Research Farm compared to previous studies, the results showed that it was a better location than Makhatini Research Station for the selection of genotypes with higher stalk yield and sugar-related traits. Between the two sites, Ukulinga Research Farm is located in the mid-latitude region with lower temperatures (14-26°C) and may have caused varieties to mature later than those at Makhatini Research Station, which is located in the low altitude region with higher temperatures (11-32°C). Rao et al. (2013a) also emphasised that certain mechanisms such as escape by early flowering, avoidance and tolerance by delaying the reproductive development, enable the varieties to achieve economic yields. The environmental conditions at Ukulinga Research Farm did not only have a positive impact on plant height, but higher mean values were generally obtained for all the measured six quality traits (fibre, dry matter, °total brix, total fermentable sugars and ethanol) and fresh stalk yield in this study. This could have been due to genotypes taking a longer period of time to reach maturity and caused the stalks to accumulate more carbohydrates and higher fresh stalk weight as reported in previous studies (Prakasham et al.,

2014; Tovignan et al., 2016). Similar findings were also reported by Turki and Al Jamali (2011). Plant height being an indicator of stalk yield as documented by Calviño and Messing (2012), could have also caused higher fresh stalk yield due to varieties being taller than those at Makhatini Research Station. The character also gives a good foundation for dry matter yield which directly increases the °brix content as observed in the study (Qu et al., 2014). It was also observed that genotypes accumulating a large quantity of sugar in the stalks were generally characterized by the longest cycle and the highest stalk dry weight due to the influence of climate associated with the higher altitude. Similar findings were reported by Tovignan et al. (2016).

Shorter plants at Makhatini Research Station seemed to have been caused by higher temperatures which could have accelerated the growth and developmental stages and therefore caused early flowering among genotypes. The negative effect of early flowering on plant height was also emphasised by Disasa et al. (2016), Zou et al. (2011) and Amelework et al. (2015) who reported that early flowering is not desirable for sugar accumulation. The slightly higher mean value for fresh biomass yield and grain weight observed at Makhatini Research Farm could have been due to more dry matter being converted into grain among the genotypes as compared to Ukulinga Research Farm. These observations are consistent with the work done Nguyen and Blum (2004). It was also observed that genotypes having highest and lowest fresh biomass yield and fresh stalk yield were similar in both locations. IESV 91018 LT and ICSV 93046 showed consistency in both locations with highest fresh stalk yield and fresh biomass yield.

Even though thicker stalks have been reported to increase sugar content, genotypes like IS 2331 with high °brix content were characterized by thinner stalks, while those with low °brix content had thicker stalks like 91018 LT. Similarly, Murray et al. (2008) and Disasa et al. (2016) reported that sweet sorghum stalks having higher °brix value had much thinner stalks. This creates difficulty in undertaking direct selection for sugar-related traits based on the thickness of the stalks. On the other hand, panicle length was observed to be consistent in both locations. A similar observation was found by Zou et al. (2011) who reported on sweet sorghum lines having similar panicle length in different locations.

A large proportion of genotypic variance contained in the phenotypic variance (high H^2), showed that genetic effects explained a large proportion of total phenotypic variance of these

traits, while environmental effects and G X E interaction effects accounted for a smaller proportion of the phenotypic variation as documented by different literature (Zou et al., 2011; Elangovan et al., 2014). Therefore, improvement of a trait with high H^2 could be brought into appropriate selection programs. On the other hand, grain yield, °brix and total fermentable sugars had the lowest H^2 respectively, indicating that these traits were more easily affected by environment and G X E interaction than traits such as stalk diameter and fresh stalk yield. Similar findings were reported by Zou et al. (2011) and Amelework et al. (2015) who emphasised that selection of traits with low H^2 would be ineffective because the environment variation and G X E interaction variation was so great that it may have masked the genetic variation.

The significant correlation between fresh stalk yield with °total brix, total fermentable sugars, plant height, days to 50% flowering and stalk diameter indicated that the specified plant height and stalk diameter are major components of fresh stalk yield, even though fresh stalk yield showed positive but non-significant correlation with stalk diameter. These findings are consistent with earlier reports (Rao et al., 2009; Shiringani and Friedt, 2009; Audilakshmi et al., 2010; Atokple et al., 2014). This observation suggests that the longer the growing cycle, the higher the stalk yield, which in turn increases the stalk sugar content and all the quality traits. On the other hand, the negative and significant correlation between days to 50% flowering and grain yield shows that delayed flowering provides more time for a plant to grow and store photosynthetic assimilates in its stalk which is a major sink to carbohydrate accumulation (Turki and Al Jamali, 2011). The positive correlation between grain yield with fresh biomass yield indicated that there might be less effect on seed production if cultivars for stalk sugar production are to be sustained.

The negative and significant ($P \leq 0.05$) correlation observed between grain yield with °brix, dry matter, fiber and plant height could be due to the fact that sugar accumulated in the sweet sorghum stalks at flowering stage could have later been transported to the grain until physiological maturity (Bihmidine et al., 2013; Atokple et al., 2014). These results are also consistent with the findings of different researchers (Murray et al., 2008; Guan et al., 2011; Bihmidine et al., 2013). Therefore high amounts of sugar in the grain could have lowered the amount stored in the stalk juice at hard dough stage. The observed negative and significant correlation between °brix and stalk diameter suggests that thicker stalks may accumulate less concentration of sugar as compared to thinner stalks. Therefore, selection for genotypes with thicker stalks may not improve the sugar concentration of the crop. Similar findings were also

reported by Makanda et al. (2009) who found a similar relationship between stalk diameter and sugar concentration. The positive and highly significant correlation (P≤0.01) between ethanol yield with total fermentable sugars, °total brix, °brix, dry matter, fresh stalk yield and fresh biomass yield showed that intensive selection based on these traits will be effective in improving ethanol yields. Similar findings were reported by Han et al. (2013), and Rani and Umakanth (2012) in their publications.

The information obtained from the principal component analysis (PCA) assists breeders in identifying phenotypic traits that contribute to great variation among genotypes for selection of potential parents for crossing blocks for the trait of interest. The total variation explained by the first principal component (PC 1) was mainly due the five traits (ethanol, total fermentable sugars, °total brix, fresh stalk yield and °brix) which had a positive correlation between each other. This suggests that increasing one trait will directly increase the other traits. Principal component two (PC2) which had the most variation being contributed by fresh biomass yield and stalk diameter indicates that increasing stalk thickness, fresh biomass yield and GY will cause a reduction in dry matter, fibre and °brix. The total variation mainly due to days to 50% flowering and grain yield observed under the third principal component indicate that a reduction in days to 50% flowering increases the grain yield. Overall, PC 1 and PC2 explained the most variation among the genotypes, exhibiting a high degree of association and interrelationships among the traits studied. The scattered genotypes generated in all four quadrants by the principal component score biplot indicates the wide genetic variability for the traits studied. The genotypes placed at extreme positions from the origin suggests that the lines were genetically distinct, whereas, the genotypes concentrated around the origin on PC2, shows that the lines were genetically similar for the studied traits. It is therefore important to include new genetically unrelated sweet sorghum genotypes in the breeding program in order to widen the genetic base of the sweet sorghum materials.

2.6 Conclusion

This preliminary study confirms previous reports showing that there is high genetic variation among sweet sorghum varieties. The nature of genetic variation with respect to various traits is important for planning a successful breeding program. The study of relationships among quantitative traits is important for assessing the feasibility of joint selection of two or more traits and hence evaluating the effect of selection for secondary traits on genetic gain for the primary trait under consideration. The study identified good association between traits (fresh biomass

yield, fresh stalk, stalk diameter, °total brix, plant height and days to 50% flowering) that were important to gain the highest yields for estimated sugar and theoretical ethanol productivity. The first three principal components showed 83% of the total variability among the genotypes. Ethanol, total fermentable sugars, °total brix, fresh stalk yield and °brix contributed mainly to PC 1, whereas fresh biomass yield and stalk diameter contributed mainly to PC 2. Cluster analysis for phenotypic traits showed a clear variation between sweet sorghum genotypes. Based on the measured traits, the dendrogram divided the genotypes into two main clusters and three singletons (ICSB 324, ICSB 654 and ICSV 700). Cluster I comprised 54% of the total germplasm and included only one ACCI genotype (SS 27), while cluster II comprised of 33% of the total variation. Genotypes IS 2331, IESV 92008 DL and ICSV 700 exhibited high quality traits and were comparable to the standard checks (AS 244, URJA and SS 27), while genotype IESV 91018 LT showed the highest yield components and stalk diameter. Therefore, these genotypes can be used as parents in sweet sorghum breeding programs to develop superior sugar-rich cultivars for bioethanol production in South Africa.

2.7 References

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

AGROMET. 2016. Agromet database, ARC-Institute for Soil, Climate and Water, Pretoria, South Africa.

Allard, R.W., and A. Bradshaw. 1964. Implications of genotype-environmental interactions in applied plant breeding. Crop Science 4:503-508.

Almodares, A., and S.M. Darany. 2006. Effects of planting date and time of nitrogen application on yield and sugar content of sweet sorghum. Journal of Environmental Biology 27:601-605.

Amelework, B., H. Shimelis, P. Tongoona, M. Laing, and F. Mengistu. 2015. Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. Plant Genetic Resources 13:131-141.

Antonopoulou, G., H.N. Gavala, I.V. Skiadas, K. Angelopoulos, and G. Lyberatos. 2008. Biofuels generation from sweet sorghum: fermentative hydrogen production and anaerobic digestion of the remaining biomass. Bioresource Technology 99:110-119.

Archer, D. 2005. Fate of fossil fuel CO_2 in geologic time. Journal Geophysical Research, 110, C09S05. doi:10.1029/2004JC002625.

Atokple, I., G. Oppong, and S. Chikpah. 2014. Evaluation of Grain and Sugar Yields of Improved Sweet Sorghum (*Sorghum bicolor*) Varieties in the Guinea Savanna Zone of Ghana. Pinnacle Agricultural Research & Management 2:374-382.

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

Babić, V.B., M.M. Babić, M.R. Ivanović, and M.R. Filipović. 2011. Pattern in interaction in the maize yield trial. Journal of Agricultural Sciences, 56:101-110.

Balole, T.V. 2001. Strategies to improve yield and quality of sweet sorghum as a cash crop for small scale farmers in Botswana. PhD dissertation, University of Pretoria, South Africa.

Bertan, I., F. Carvalho, and A.d. Oliveira. 2007. Parental selection strategies in plant breeding programs. Journal of Crop Science and Biotechnology 10:211-222.

Bihmidine, S., C.T. Hunter III, C.E. Johns, K.E. Koch, and D.M. Braun. 2013. Regulation of assimilate import into sink organs: update on molecular drivers of sink strength. Frontiers in Plant Science 4:177. doi: 10.3389/fpls.2013.00177.

Blanchard, R., D.M. Richardson, P.J. O'Farrell, and G.P. Von Maltitz. 2011. Biofuels and biodiversity in South Africa. South African Journal of Science 107:19-26.

Calviño, M., and J. Messing. 2012. Sweet sorghum as a model system for bioenergy crops. Current Opinion in Biotechnology 23:323-329.

Cartwright, A. 2007. Biofuels trade and sustainable development: An analysis of South African bioethanol. International Institute for Sustainable Development, Winnipeg, Canada.

Disasa, T., T. Feyissa, and B. Admassu. 2016. Characterization of Ethiopian Sweet Sorghum Accessions for ^{o}Brix, Morphological and Grain Yield Traits. Sugar Technology:1-11.

Elangovan, M., N. Seetharama, and J. Patil. 2014. Genetic diversity and heritability characters associated in sweet sorghum [*Sorghum bicolor* (L.) Moench]. Sugar Technology 16:200-210.

Faisal, E., and O. Aisha. 2011. Genotype x seed production environment interaction on the performance of sorghum (*Sorghum bicolor* [L.] Moench) under irrigation. Agriculture and Biology Journal for North America 2:2151-7517.

Flury, B.K., and H. Riedwyl. 1986. Standard distance in univariate and multivariate analysis. The American Statistician 40:249-251.

Gao, Z. 1986. Quantitative genetics. Sichang University, Chongqing, China.

Guan, Y.-a., H.-l. Wang, L. Qin, H.-w. Zhang, Y.-b. Yang, F.-j. Gao, et al. 2011. QTL mapping of bio-energy related traits in Sorghum. Euphytica 182:431-440.

Gurmu, F., H. Mohammed, and G. Alemaw. 2009. Genotype x Environment interactions and stability of soybean for grain yield and nutrition quality. African Crop Science Journal 17:87-99.

Han, K.J., W.D. Pitman, M. Kim, D.F. Day, M.W. Alison, M.E. McCormick, et al. 2013. Ethanol production potential of sweet sorghum assessed using forage fiber analysis procedures. GCB Bioenergy 5:358-366.

IBPGR/ICRISAT. 1993. Descriptors for Sorghum [*Sorghum bicolor* (L) Moench]. International Board of Plant Genetic Resources. Rome, Italy/International Crop Research Institute for Semi-Arid Tropics, Patancheru, India.

Iyanar, K., G. Vijayakumar, and A. Khan. 2010. Correlation and path analysis in multicut fodder sorghum. Electronic Journal of Plant Breeding 1:1006-1009.

Makanda, I. 2009. Combining ability and heterosis for stem sugar traits and grain yield components in dual-purpose sorghum (*Sorghum bicolor* L. Moench) germplasm. PhD dissertation, University of KwaZulu-Natal, South Africa.

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.

Manggoel, W., M. Uguru, O. Ndam, and M. Dasbak. 2012. Genetic variability, correlation and path coefficient analysis of some yield components of ten cowpea [*Vigna unguiculata* (L.) Walp] accessions. Journal of Plant Breeding and Crop Science 4:80-86.

Murray, S.C., W.L. Rooney, S.E. Mitchell, A. Sharma, P.E. Klein, J.E. Mullet, et al. 2008. Genetic improvement of sorghum as a biofuel feedstock: II. QTL for stem and leaf structural carbohydrates. Crop Science 48:2180-2193.

Mutepe, R.D., S. Marx, and P. Van der Gryp. 2012. Ethanol production from sweet sorghum. MSc thesis, North-West University, Potchefstroom Campus, South Africa.

Nguyen, H.T., and A. Blum. 2004. Physiology and biotechnology integration for plant breeding. CRC Press.

Prakasham, R.S., D. Nagaiah, K.S. Vinutha, A. Uma, T. Chiranjeevi, A.V. Umakanth, et al. 2014. Sorghum biomass: a novel renewable carbon source for industrial bioproducts. Biofuels 5:159-174.

Qu, H., X.B. Liu, C.F. Dong, X.Y. Lu, and Y.X. Shen. 2014. Field performance and nutritive value of sweet sorghum in eastern China. Field Crops Research 157:84-88.

Rani, C., and A. Umakanth. 2012. Genetic variation and trait inter-relationship in F1 hybrids of sweet sorghum (*Sorghum bicolor* (L.) Moench). Journal of Tropical Agriculture 50:80-83.

Rao, S., J. Patil, P. Prasad, D. Reddy, J. Mishra, A. Umakanth, et al. 2013a. Sweet sorghum planting effects on stalk yield and sugar quality in semi-arid tropical environment. Agronomy Journal 105:1458-1465.

Rao, S., J. Patil, A. Umakanth, J. Mishra, C. Ratnavathi, G.S. Prasad, et al. 2013b. Comparative performance of sweet sorghum hybrids and open pollinated varieties for millable stalk yield, biomass, sugar quality traits, grain yield and bioethanol production in tropical Indian condition. Sugar Technology 15:250-257.

Rao, S., S. Rao, N. Seetharama, A. Umakath, P.S. Reddy, B. Reddy, et al. 2009. Sweet sorghum for biofuel and strategies for its improvement. Information Bulletin No. 77. International Crops Research Institute for the Semi-Arid Tropics, Patancheru - 502 324, A.P, India.

Reddy, B.V., S. Ramesh, A.A. Kumar, S. Wani, R. Ortiz, H. Ceballos, et al. 2008. Bio-fuel crops research for energy security and rural development in developing countries. Bioenergy Research 1:248-258.

SAS, S. 2011. Windows Version 9.3. SAS Institute Inc., Cary, NC, USA.

Shinde, M., B. Mutkule, A. Gaikwad, U. Dalvi, and S. Gadakh. 2013. Photoperiod sensitivity studies in sweet sorghum. Journal of Academia and Industrial Research. Maharashtra, IN 1.

Shiringani, A., and W. Friedt. 2009. Genotype-environmental analysis of RIL population segregating for sugar-related traits in Sorghum bicolor L. Moench. Identification of genomic regions of *Sorghum bicolor* (L.) Moench linked to biofuel-related traits in grain x sweet sorghum recombinant inbred lines. PhD dissertation, Justus-Liebig-University, Germany.

Tovignan, T.K., D. Fonceka, I. Ndoye, N. Cisse, and D. Luquet. 2016. The sowing date and post-flowering water status affect the sugar and grain production of photoperiodic,

sweet sorghum through the regulation of sink size and leaf area dynamics. Field Crops Research 192:67-77.

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.

Turki, D.A., and M. Al Jamali. 2011. Relationship between shoot and sugar-related characters of sweet sorghum (*Sorghum bicolor* L.) under contrasting moisture regimes. American-Eurasian Journal of Sustainable Agriculture 5:252-258.

Yan, W., and L. Hunt. 1998. Genotype by environment interaction and crop yield. Plant Breeding Reviews 16:135-178.

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

3 Assessment of sweet sorghum lines for genetic diversity and interrelationships using simple sequence repeat (SSR) markers

Abstract

South Africa has the potential to contribute to a sustainable renewable bioethanol production by engaging in improved sweet sorghum production. In order for plant breeders to effectively select superior germplasm and eliminate undesirable characteristics for hybridization, there is need to have proper knowledge of genetic diversity and genetic variation. The objectives of the study were to assess the genetic variability present at molecular level using simple sequence repeat (SSR) markers and to identify potential heterotic groups for hybrid development of sweet sorghum lines. Twenty-four sweet sorghum lines obtained from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and the African Centre for Crop Improvement (ACCI-South Africa) were genotyped with 10 SSR markers and distance-based method was used to analyze the data. Variation was observed for all the markers with allelic size ranging from 1 to 36 bp. A total of 61 alleles were generated with an average of 6.1 alleles per locus. The PIC values ranged from 0.32 to 0.86 with an overall mean value of 0.62, showing a high discriminating ability of the markers used. The genetic distance ranged from 0.50 to 1.9 with AS 244 (GD = 1.9) showing the largest genetic distance, while IESV 92001 DL and IESV 92008 had the smallest genetic distance (GD = 0.50). On the basis of cluster analysis using SSR markers, the 24 sweet sorghum lines were classified into two major clusters. Cluster I comprised of 13% of the total genotypes which included URJA, SS 27 and ICSB 654. It was observed that all ACCI genotypes apart from AS 244 were grouped in Cluster I, with URJA and SS 27 being very closely related. Cluster II was observed to be the largest (87%) with 21 genotypes, which further formed 3 sub-clusters (A, B and C) and a singleton (AS 244). The same results were obtained using PCA analysis which grouped genotypes into four clusters with same type of genotypes in each group. The results showed that the sweet sorghum lines within clusters were closely related than between clusters. It is concluded that genetic diversity exists in the germplasm. Therefore, the information obtained from this study can be used by plant breeders to select parents for hybridization towards development of improved cultivars from superior progenies and in turn contribute to the production and promotion of bioethanol in South Africa.

3.1 Introduction

Sweet sorghum [*Sorghum bicolor* (L.) Moench; 2n = 2x = 20] is a biofuel crop that is receiving interest globally because of the high sugar content in its stalk juice (Antonopoulou et al., 2008; Tsuchihashi and Goto, 2008; Khalil et al., 2015). There has been a significant breakthrough in developing sweet sorghum hybrids and varieties for commercial cultivation in India and elsewhere (Kumar et al., 2011). South Africa has the potential to contribute to a sustainable renewable bioethanol production by engaging in improved sweet sorghum production. In order for plant breeders to effectively select superior germplasm and eliminate undesirable characteristics for hybridization, there is need to have proper knowledge of the genetic diversity and genetic variation in their germplasm (Govindaraj et al., 2015). Genetic diversity refers to the total number of genetic characteristics in the genetic makeup of species while genetic variation occurs within alleles both within and among species. Selection of desirable traits in a breeding program is determined by genetic variation which further improves the genetic diversity of plant genetic resources (Cholastova et al., 2013).

Sweet sorghum genotypes are particularly known to have substantial variations for sugar content and sugar-related traits (Rao et al., 2009). Knowledge of phenotypic and genotypic variation and association between the components is important for various reasons. These include; (i) identification and selection of desirable lines for hybrid development and conservation (Tomar et al., 2012; Singh and Singh, 2015), (ii) characterization of individuals in determining duplications for germplasm collections (Singh and Singh, 2015) and (iii) selecting parents from heterotic groups formed from cluster analysis (Tomar et al., 2012; Singh and Singh, 2015).

Heterotic groups are achieved based on various data such as pedigree, geographical origin, morphological and molecular markers. Morphological markers are mostly used by scientists because of the affordability of the technology based on visual trait assessment. Phenotypic selection allows the plants to grow to full maturity for the traits to be significantly distinct prior to identification and differentiates matured plants in the field from their genetic contamination. However, morphological data does not exhibit a true reflection of molecular data because of the environmental influence and the largely unknown genetic control of poly genetically inherited morphological and agronomic traits (Assar et al., 2005; Geleta et al., 2005). The desirable quantitative traits controlled by many genes in sweet sorghum can be efficiently detected using molecular markers (Aremu, 2012; Elangovan et al., 2014). Molecular data is known to provide more accuracy in determining the similarities and differences among

germplasm materials because of lack of environmental influences. Detection of molecular data can also be carried out at any growth stage of the plant (Soriano et al., 2005; Turki et al., 2011).

Various types of DNA based markers can be effectively used in genetic diversity studies, namely; (i) random amplified polymorphic DNA (RAPD), (ii) restriction fragment length polymorphism (RFLP), (iii) amplified fragment length polymorphism (AFLP) and (iv) simple sequence repeats (SSRs), also known as microsatellites (Avise, 2012; Karp et al., 2012). In sweet sorghum, SSRs have mostly been used due to their high levels of polymorphism and abundance (Geleta et al., 2006; Ali et al., 2008; Shehzad et al., 2009). Simple sequence repeats are mostly used to discriminate among closely related individuals and effectively identify heterotic groups (Assar et al., 2005; Geleta et al., 2006; Lekgari and Dweikat, 2014).

In South Africa, there is limited information on the availability of improved sweet sorghum cultivars due to various reasons, such as; (i) poor gene pool creation and characterization of sweet sorghum lines and (ii) poor selection of appropriate parental lines for hybrid combinations. The study was therefore undertaken to assess genetic diversity among sweet sorghum lines obtained from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and the African Centre for Crop Improvement (ACCI-South Africa) using SSR markers. The aim was to identify specific genotypes exhibiting high levels of biomass yield and sugar-related traits from various heterotic groups that will improve the genetic base and characterization of the current cultivated sweet sorghum (Singh and Singh, 2015). As a result, there will be a development of hybrids for ethanol production that will promote the biofuel industry in South Africa. The small holder farmers, in turn, would benefit and have better livelihoods by producing high yielding sweet sorghum varieties. The study was therefore undertaken to; (i) to assess the genetic variability present at molecular level using simple sequence repeat (SSR) markers and (ii) to identify potential heterotic groups for hybrid development of sweet sorghum lines.

3.2 Materials and methods

3.2.1 Germplasm collection

The twenty-four sweet sorghum genotypes used in the study were obtained from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Kenya) and the African Centre for Crop Improvement (ACCI-South Africa) as shown Table 3-1.

Table 3-1: A list of 24 sweet sorghum lines used in the study and country of origin

No.	Name	Origin
1	AS 244	ACCI-South Africa
2	E 36-1	ICRISAT-Kenya
3	Ent # 64DTN	ICRISAT-Kenya
4	ICSB 324	ICRISAT-Kenya
5	ICSB 654	ICRISAT-Kenya
6	ICSR 93034	ICRISAT-Kenya
7	ICSV 700	ICRISAT-Kenya
8	ICSV 93046	ICRISAT-Kenya
9	IESV 91104 DL	ICRISAT-Kenya
10	IESV 91018 LT	ICRISAT-Kenya
11	IESV 92001 DL	ICRISAT-Kenya
12	IESV 92008 DL	ICRISAT-Kenya
13	IESV 92021 DL	ICRISAT-Kenya
14	IESV 92028 DL	ICRISAT-Kenya
15	IESV 92165 DL	ICRISAT-Kenya
16	IESV 94021 DL	ICRISAT-Kenya
17	IS 2331	ICRISAT-Kenya
18	Kari Mtama 1	ICRISAT-Kenya
19	MR # 22 x IS 8613/1/2/5-2-1	ICRISAT-Kenya
20	MR # 22 x IS 8613/2/3-1-3	ICRISAT-Kenya
21	NTJ 2	ICRISAT-Kenya
22	URJA	ACCI-South Africa
23	SPV 1411	ICRISAT-Kenya
24	SS 27	ACCI-South Africa

3.2.2 Sweet sorghum Samples

The 24 sweet sorghum lines were grown on polystyrene transplant trays using pine bark medium in the ACCI tunnel at the University of KwaZulu-Natal. Leaf tissues were harvested from 10 plants of each genotype four weeks after planting and were stored at -20°C in brown paper bags. The leaf samples were sent to SciCorp laboratories (Incotech South Africa Pty Ltd, Mkondeni, Pietermaritzburg, South Africa) for SSR analysis.

3.2.3 DNA extraction

Deoxyribonucleic acid (DNA) was extracted from all leaf samples using the standard cetyltrimethylammomium bromide (CTAB) based method (Mace et al., 2003). Leaf samples and two steel beads were put in each of the wells of a Geno Grinder 2000 (Spex CertiPrep, USA) plate and then placed in liquid nitrogen to make the leaf samples easy to grind. 500 µL preheated (65°C) extraction buffer (3% (w/v) CTAB, 1.4 M NaCl, 0.2% (v/v) β-Mercapto-ethanol and 20 mM EDTA) was added and ground using the Genogrinder. The ground substance was then transferred to fresh microfuge tubes and incubated at 65°C for 15 minutes. Solvent extraction was followed by adding 450 µl chloroform: isoamyalcohol (24:1) to each sample and thoroughly mixed by inversion. The tubes were centrifuged at 12000 rpm at 24°C for 10 minutes and the upper portion transferred into fresh tubes (about 400 µl). Iso-propanol (0.7 Volumes) was added and thoroughly mixed by inversion after 20-30 minutes, the tubes were centrifuged at 12000 rpm for 15 minutes. 200µl Low salt TE buffer (1 mM Tris and 0.1 mM EDTA [PH 8]) with 3 µl RNase A (10 mg/ml) was added to each sample and incubated overnight at room temperature. The volume was then transferred to fresh tubes and chloroform: isoamyalcohol (24:1) was added to each tube and inverted twice to mix and centrifuged. The aqueous layer was then later transferred into fresh tubes and DNA purification was followed by adding 315 µl ethanol and 1/10 volume of 3 M sodium acetate solution (PH 5.2) to each sample and then placed in -20°C for 5 minutes. The tubes were then centrifuged at 12000 rpm for 5 minutes and the supernatant decanted. The quantity and quality of the DNA was examined using a nano-drop spectrophotometer.

3.2.4 SSR primers and PCR amplification

Ten SSR primers were used to assess genetic diversity of sweet sorghum lines (Table 3-2). The selection of primers was based on their genome position, repeat size and the number of previously reported alleles. The SSR reaction contained 1 µl pooled DNA, 0.5 µl LIZ size standard and 8.5 µl Hi-Di Formamide. For initial denaturation, the PCR reaction conditions

consisted of 2 minutes at 94°C, followed by 26 cycles of polymerization reaction, each comprising of denaturation step of 30 seconds at 94°C. Annealing was then carried out at 60°C for 30 seconds and a polymerization step of 45 seconds at 72°C. The PCR reaction was repeated for 26 cycles of denaturation of 30 seconds at 94°C, an annealing step of 30 seconds at 52°C, and a polymerization step of at 72°C for 45 seconds. The final stage was polymerization of 5 minutes at 72°C based on their expected amplicon size and dye. PCR products were fluorescently labelled and separated by capillary electrophoresis on an ABI 3130xl automatic sequencer (Applied Biosystems, Johannesburg, South Africa).

Table 3-2: SSR primer sequences used in the study

Marker	Motif	Chr.	Forward Primer	Reverse Primer
mSbCIR223	$(AC)_6$	2	CGTTCCAATGACTTTTCTTC	GCCAATGTGGTGTGATAAAT
mSbCIR240	$(TG)_9$	8	GTTCTTGGCCCTACTGAAT	TCACCTGTAACCCTGTCTTC
mSbCIR246	$(CA)_7$ $(GA)_5$	7	TTTTGTTGCACTTTTGAGC	GATGATAGCGACCACAAATC
mSbCIR248	$(GT)_7$	5	GTTGGTCAGTGGTGGATAAA	ACTCCCATGTGCTGAATCT
mSbCIR276	$(AC)_9$	3	CCCCAATCTAACTATTTGGT	GAGGCTGAGATGCTCTGT
mSbCIR283	$(CT)_8$ $(GT)_8$	10	TCCCTTCTGAGCTTGTAAAT	CAAGTCACTACCAAATGCAC
mSbCIR286	$(AC)_9$	1	GCTTCTATACTCCCCTCCAC	TTTATGGTAGGATGCTCTGC
Xcup02	$(GCA)_6$	9	GACGCAGCTTTGCTCCTATC	GTCCAACCAACCCACGTATC
Xgap72	$(AG)_{16}$	6	TGCCACCACTCTGGAAAAGGCTA	CTGAGGACTGCCCCAAATGTAGG
Xtxp012	$(CT)_{22}$	4	AGATCTGGCGGCAACG	AGTCACCCATCGATCATC

Chr. = Chromosome

3.3 Data analysis

3.3.1 Polymorphism and allelic diversity of SSR markers

The allelic data obtained from Genemapper 4.1 software was used to determine genetic diversity parameters including; the total number of alleles (N_a), the number of effective alleles per locus (N_e), observed gene diversity within genotypes (H_o), average gene diversity within genotype (H_e), major allele frequency (A) and polymorphism information content (PIC) using the protocol of Nei (1978). FSTAT version 2.9.3 (Goudet, 2001) and GENALEX version 6.5 (Peakall and Smouse, 2012) were used to calculate fixation index (F).

3.3.2 Genetic distance and dissimilarity analysis

The program GGT 2.0 (van Berloo, 2008) was used to calculate the Euclidean distances between genotypes.

3.3.3 Cluster and PCA analysis

Cluster analysis was performed based on Euclidean dissimilarity matrix and a dendrogram was constructed using the unweighted pair group method with arithmetic mean (UPGMA). Principal component analysis based on Euclidean dissimilarity matrix was employed using PROC PRINCOMP in SAS version 9.3 (SAS, 2011).

3.4 Results

3.4.1 Polymorphism and allelic diversity of SSR markers

Polymorphism within and among the 24 sweet sorghum genotypes was successfully assessed using all 10 SSR markers. The total number of alleles amplified per locus (N_a) was 61, with a mean value of 6.1 (Table 3-3). This study revealed mSbCIR240 and mSbCIR286 loci to have the highest alleles (11), while SbCIR276 was observed to have the least (2) per locus. On the other hand, the number of effective alleles per locus (N_e) was observed to range from 1.40 to 7.02 for mSbCIR248 and mSbCIR286 respectively, with an overall mean value of 3.33. Additionally, only two (mSbCIR240 and mSbCIR286) SSR loci had above average number of effective alleles per locus. The results showed 80% of the primers exhibiting PIC values more than 0.5. The PIC values for SSR loci ranged from 0.32 to 0.86 for mSbCIR248 and mSbCIR286 respectively, with an overall mean value of 0.62. Additionally, the amplified fragment sizes (difference between the shortest and longest fragments) ranged from 1 to 36 bp for locus mSbCIR276 and Locus mSbCIR283. The average gene diversity within genotypes (H_e) across all loci ranged from 0.29 (mSbCIR248) to 0.88 (mSbCIR286), with an overall mean value of 0.63 (Table 3-3). On the other hand, the observed gene diversity within genotypes (H_o) ranged from 0 to 0.83, with an overall mean value of 0.33. Locus mSbCIR276 was observed to be homozygous (Table 3-3) while four SSR loci had above average H_o. The assessment of SSR classes based on the number of repeats and motifs exhibited an average of 11 per locus with mSbCIR286 and mSbCIR240 having higher allele numbers. However, AC repeats with mean alleles of 11 per locus and PIC value of 0.86 was more informative as compared to TG repeats with mean alleles of 11 per locus and PIC value of 0.84. The markers with the least allele numbers included mSbCIR276, mSbCIR248 and mSbCIR246 with an

average of 2.6 per locus. On the other hand, the average for PIC values among the three markers was 0.38 per marker (Table 3-3).

Table 3-3: Genetic diversity within and among 24 sweet sorghum genotypes based on 10 SSR markers

Marker	Allele size range (bp)	Genetic parameters							
		N_a	N_e	I	H_o	H_e	F	A	PIC
mSbCIR223	114 - 133	6	3.06	1.37	0.71	0.69	-0.05	0.50	0.67
mSbCIR240	116 - 131	11	6.10	2.06	0.83	0.85	0.00	0.29	0.84
mSbCIR246	116 - 122	3	2.13	0.88	0.04	0.54	0.92	0.60	0.53
mSbCIR248	110 - 120	3	1.40	0.54	0.08	0.29	0.71	0.83	0.29
mSbCIR276	247 - 248	2	1.47	0.50	0.00	0.33	1.00	0.80	0.32
mSbCIR283	130 - 166	7	3.54	1.50	0.26	0.73	0.64	0.43	0.72
mSbCIR286	118 - 150	11	7.02	2.11	0.83	0.88	0.03	0.23	0.86
Xcup02	210 - 224	5	2.61	1.20	0.08	0.63	0.86	0.56	0.62
Xgap72	202 - 213	6	2.76	1.34	0.38	0.65	0.41	0.56	0.64
Xtxp012	194 - 225	7	3.23	1.46	0.04	0.70	0.94	0.48	0.69
Overall mean	155.7 – 173.2	6.1	3.33	1.29	0.33	0.63	0.55	0.53	0.62
SE		1.0	0.59	0.17	0.11	0.06	0.13	0.06	0.06

N_a= total number of alleles per locus; N_e= number of effective alleles per locus; H_o= observed gene diversity within genotypes; H_e= average gene diversity within genotypes; PIC= polymorphic information content; A= major allele frequency; SE, Standard error.

3.4.2 Genetic distance and dissimilarity analysis

The Euclidean dissimilarity matrix estimates as measures of genetic distances ranged from 1.0 to 5.5 with an average value of 3.5 (Table 3-4). Among the 24 genotypes studied, IESV 92001 DL and IESV 92008 DL had the lowest dissimilarity index (1.0), whereas Kari Mtama 1 and SS 27 exhibited the highest dissimilarity (5.5).The genotypes sourced from ACCI (AS 244, URJA and SS 27) were observed to be distantly related to most genotypes tested. Among the ICRISAT genotypes, ICSB 654 was the most dissimilar. On the other hand, genotypes such as Kari Mtama 1, MR # 22 x IS 8613/1/2/5-2-1, MR # 22 x IS 8613/2/3-1-3, NTJ 2, URJA, SPV 1411 and SS 27 displayed high genetic dissimilarity (Table 3-4).

Table 3-4: The Euclidean dissimilarity matrix for all pair-wise comparison of 24 sweet sorghum lines

No.	Genotype	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	AS 244	1.0																						
2	E 36-1	3.8	1.0																					
3	Ent # 64DTN	4.4	3.2	1.0																				
4	ICSB 324	3.5	1.8	3.1	1.0																			
5	ICSB 654	3.5	3.2	4.6	2.6	1.0																		
6	ICSR 93034	4.0	3.3	4.5	3.0	3.7	1.0																	
7	ICSV 700	3.3	3.1	3.8	2.5	3.1	3.0	1.0																
8	ICSV 93046	3.9	4.2	3.9	3.0	3.9	3.5	2.3	1.0															
9	IESV 91104 DL	3.3	4.2	4.7	3.4	4.1	2.4	2.6	3.2	1.0														
10	IESV 91018 LT	3.6	3.3	2.7	2.7	4.5	4.4	3.3	3.0	4.3	1.0													
11	IESV 92001 DL	3.6	3.4	3.4	3.7	4.6	3.2	2.4	3.4	3.5	3.4	1.0												
12	IESV 92008 DL	3.3	3.4	3.5	3.5	4.3	3.0	2.2	3.2	3.0	3.5	1.0	1.0											
13	IESV 92021 DL	4.5	3.4	3.0	3.0	4.0	3.9	3.1	3.5	4.2	2.9	3.5	3.6	1.0										
14	IESV 92028 DL	3.9	2.5	2.8	3.1	4.1	4.3	3.0	4.1	4.5	2.7	2.7	2.9	2.2	1.0									
15	IESV 92165 DL	4.3	3.2	4.1	3.6	4.2	2.7	3.3	3.9	4.2	4.5	2.5	2.7	4.7	4.1	1.0								
16	IESV 94021 DL	3.7	4.1	3.3	3.6	4.5	2.7	2.8	2.6	2.7	3.8	2.1	1.8	3.7	3.9	2.8	1.0							
17	IS 2331	4.6	3.6	2.2	2.6	4.0	4.0	3.4	3.1	4.1	2.9	3.9	4.0	1.9	3.3	4.6	3.4	1.0						
18	Kari Mtama 1	4.2	3.0	3.8	3.5	4.4	2.5	3.5	4.2	4.0	4.3	2.2	2.4	4.4	3.8	1.1	2.7	4.4	1.0					
19	MR # 22 x IS 8613/1/2/5-2-1	3.5	2.9	1.9	2.6	4.2	4.1	3.0	3.0	4.2	1.2	2.6	2.8	2.4	1.9	3.8	3.1	2.5	3.6	1.0				
20	MR # 22 x IS 8613/2/3-1-3	4.0	3.3	2.7	2.7	3.8	4.1	3.1	3.3	3.8	3.0	3.7	3.3	2.3	2.7	4.7	3.4	2.6	4.4	2.5	1.0			
21	NTJ 2	3.8	3.1	3.4	2.6	3.2	4.4	2.8	3.8	4.0	3.9	4.1	3.7	4.2	3.6	4.3	3.9	3.6	4.4	3.5	3.0	1.0		
22	URJA	4.2	3.8	4.2	3.0	2.8	4.6	3.7	3.4	5.0	3.9	4.7	4.4	4.1	4.2	4.6	4.4	4.1	4.8	3.7	3.2	3.6	1.0	
23	SPV 1411	3.0	3.1	3.3	2.4	3.3	3.7	2.1	2.4	3.5	2.7	2.9	2.9	3.3	3.0	3.3	3.0	3.2	3.6	2.3	3.3	2.8	3.5	1.0
24	SS 27	4.1	3.7	4.6	2.6	2.6	4.7	3.8	3.7	4.7	3.8	5.4	5.1	4.3	4.6	5.3	5.1	4.0	5.5	4.2	3.6	3.6	2.3	3.9

3.4.3 Cluster analysis

The Euclidean dissimilarity matrix was used to cluster genotypes using the UPGMA algorithm (Table 3-4). The dendrogram demarcated the genotypes into two main clusters; cluster I with 0.69 and cluster II with 0.14 Euclidean distances (Figure 3-1). Cluster I comprised of 13% of the total genotypes which included URJA, SS 27 and ICSB 654. It was observed that all ACCI genotypes apart from AS 244 were grouped in Cluster I, with URJA and SS 27 being very closely related. Cluster II was observed to be the largest (88%) with 21 genotypes, which further formed 3 sub-clusters (A, B and C) and a singleton (AS 244) which was obtained from ACCI. Sub-cluster A included IESV 91104 DL, ICSR 93034, IESV 94021 DL, IESV 92008 DL, IESV 92001 DL, Kari Mtama 1 and IESV 92165 DL. On the other hand, Cluster B included IS 2331, IESV 92021 DL, MR # 22 x IS 8613/2/3-1-3, IESV 92028 DL, MR # 22 x IS 8613/1/2/5-2-1, IESV 91018 LT and Ent #64DTN. Four genotypes that were grouped in Cluster C included ICSV 93046, SPV 1411 ICSV 700, NTJ, ICSB 324 and E 36-1 as shown in Figure 3-1.

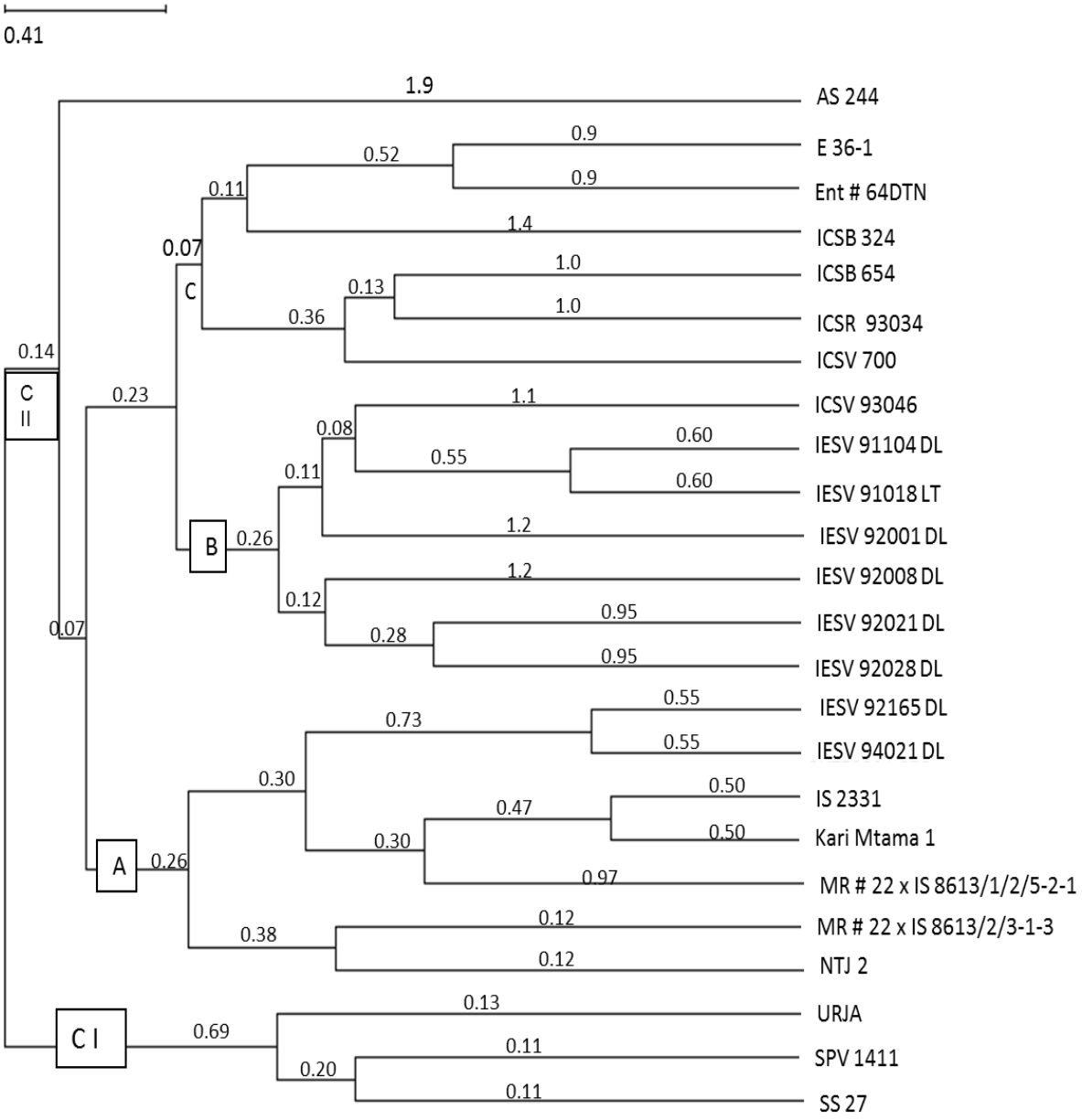
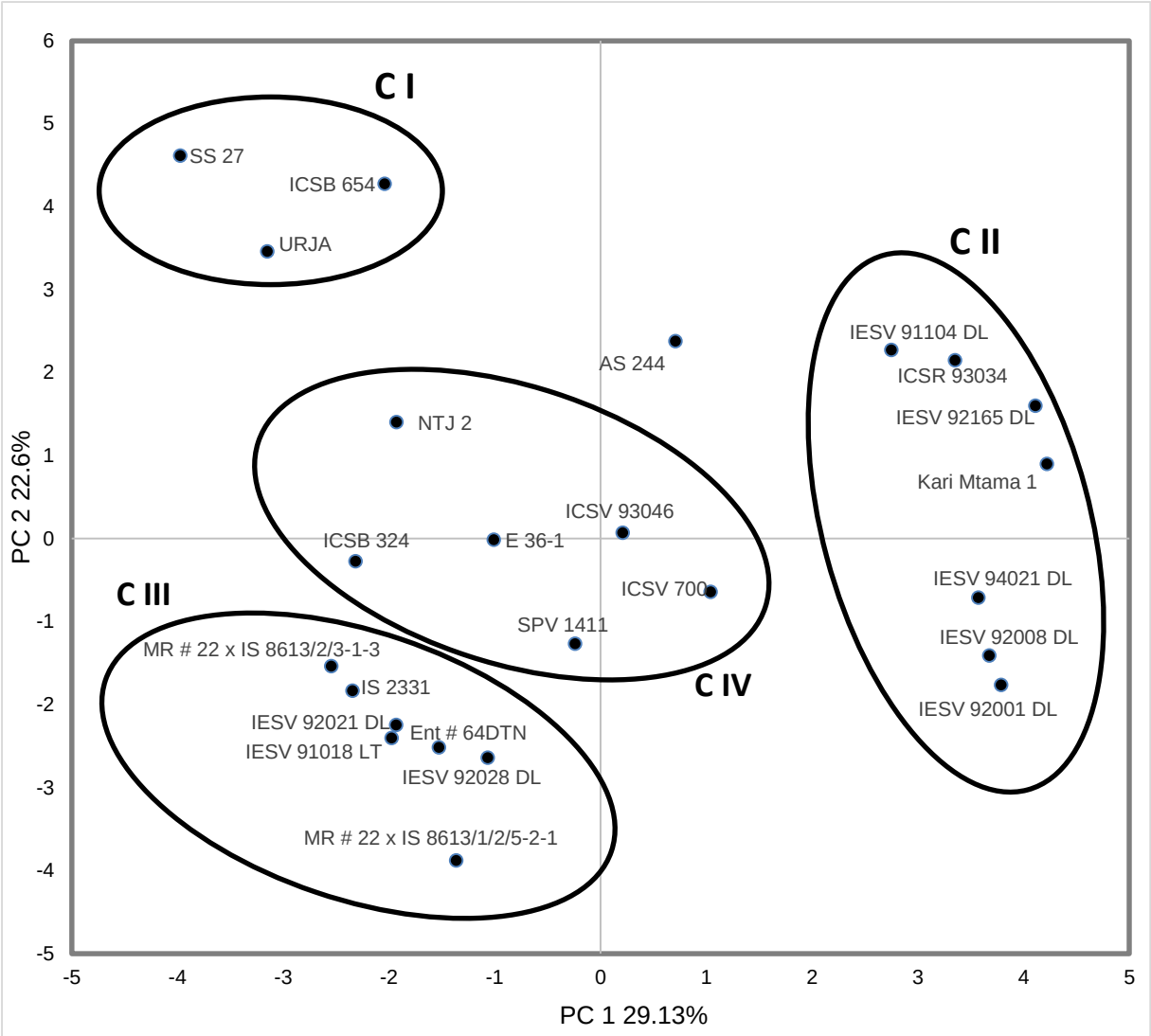


Figure 3-1: Dendrogram showing genetic relationships among 24 sweet sorghum lines based on SSR analysis of Euclidean similarity coefficients with UPGMA clustering

3.4.4 Principal component analysis (PCA)

Principal component analysis (PCA) was performed on the basis of Euclidean dissimilarity index by simple matching of 24 genotypes. The generated PCA biplot showed four (CI, CII, CIII, and CIV) interrelationship groupings among the sweet sorghum lines (Figure 3-2). Cluster I and II had three and seven genotypes, respectively. On the other hand, cluster III and IV exhibited 7 genotypes in each group. The clusters generated from PCA based on different algorithms using eigenvectors were exactly the same as that of the dendrogram (Figure 3-1).



.

Figure 3-2: Principal component score plot of PC 1 and PC 2 describing the overall variation among 24 sweet sorghum genotypes estimated using 10 SSR markers

3.5 Discussion

Simple sequence repeat (SSR) markers are popularly utilized in sweet sorghum genetic studies due to their high degree of polymorphism, co-dominance and multiple allele properties (Geleta et al., 2006). In this study, all the 10 SSR markers were able to discriminate between the genotypes studied, confirming their importance in genetic diversity studies. The total number of 61 alleles per locus observed in this study is higher than earlier findings of Ng'uni et al. (2011) and Motlhaodi (2016) who recorded a total number of 44 and 53 alleles per locus using 10 SSR markers. Nonetheless, the total number of alleles per locus was lower than previous findings by other researchers. For example, Olweny et al. (2014) reported a total of 88 alleles with 11 microsatellites in 86 sweet sorghum accessions from five different countries, while Folkertsma et al. (2005) detected 123 alleles using 21 microsatellites in 100 sorghum accessions from 10 African countries and India. These studies, however, contained sorghum accessions from different geographic areas and the number of microsatellites used was higher when compared with the current study. The overall mean value of 6.1 for the total number of alleles per locus (N_a) is similar to 6.5 and 6.4 alleles per locus reported by Reedy and Khanna (2010) and Mofokeng et al. (2014), but higher than 4.9 which was reported by (Pei et al., 2010). On the other hand, Wang et al. (2009) and Menz et al. (2004) reported higher overall mean values of 7.6 and 8.7 respectively. The low overall mean observed in this study could have been due to the small population size of the genotypes. According to Mofokeng et al. (2014), the greater number of alleles created by SSR markers gives a good indication of genetic strength for subsequent selection strategies. On the other hand, dissimilarities observed between the overall mean values for the total number of alleles per locus (N_a) and number of effective alleles per locus (N_e) could have been due to the variation in the major allele frequencies among the genotypes.

The SSR markers in this study were polymorphic. The overall mean value (0.62) of polymorphic information content (PIC) obtained in this study is similar with 0.65 and 0.60 published by Geleta et al. (2006) and Amelework et al. (2015). The level of diversity exhibited by 10 SSR markers in this study is relatively higher than the overall mean values of 0.49 and 0.50 which were reported by Ganapathy et al. (2012) and Olweny et al. (2014). According to DeWoody et al. (1994), markers with PIC more than 0.5 are efficient in discriminating genotypes and extremely useful in detecting the genetic variation among genotypes. This is evidenced by the findings of Taramino et al. (1997) and Kong et al. (2000) who previously reported higher overall mean PIC values of 0.80 and 0.89 respectively. A narrow range of amplified fragment sizes (difference between the shortest and longest fragments) from 1 to 36 bp obtained in this study is similar to the range of 2 to 44 b p reported by Olweny et al. (2014).

On the other hand, the range of amplified fragment sizes was lower than 51 to 421 bp and 4 to 134 bp reported by Geleta et al. (2006) and Amelework et al. (2015) respectively.

The overall mean value (0.61) of the average gene diversity within genotypes (H_e) across all loci obtained in this study was similar with 0.61 and 0.66 reported by Deu et al. (2010) and Ngugi and Onyango (2012) respectively. On the other hand, the low overall mean value (0.33) of the observed gene diversity within genotypes (H_o) was slightly higher than 0.17 and 0.10 reported by Olweny et al. (2014) and Amelework et al. (2015). Similarly, most of the genetic studies in sorghum using SSRs supported these findings (Deu et al., 2010; Ngugi and Onyango, 2012). The observed low heterozygosity and moderately high fixation index observed in this study signified that the genotypes were different and mainly homozygous, which is maintained by self-fertilization. The assessment of SSR classes based on the number of repeats and motifs showed that there is an association between the repeat length and type, and the degree of polymorphism using 10 SSR markers. This is in conformity with the results obtained by Amelework et al. (2015) who highlighted that this association can have an implication in marker-assisted selection and selection of SSR markers for genetic diversity analysis.

Cluster and PCA analysis are multivariate techniques that were used to visually observe genetic relationship among ICRISAT and ACCI materials in this study. The dendrogram generated from cluster analysis indicated that all the accessions apart from one (AS 244) could be distinguished and clustered into four main clusters, I, II, III, and IV. The four groups created suggests that genotypes in each group could be closely related and might have the same genetic background (Geleta et al., 2006). This observation is supported by other researchers who published similar results (Kamuntu, 2010_ENREF_77; Olweny et al., 2014). For example, Cluster I was observed to contain one genotype from ICRISAT, suggesting that there might be a close relationship between this genotype with the other two genotypes obtained from ACCI (URJA and SS 27). On the other hand, Cluster II had three generated sub-clusters and a singleton (AS 244) which was obtained from ACCI. This genotype had the highest genetic distance, suggesting that this particular genotype was distantly related with other genotypes in different groups. The distinction is clearly visualized in the generated dendrogram.

Although the PCA is based on different algorithms using the Eigen vector, the groupings of the genotypes was the same with that of the dendrogram. This suggested that both

multivariate techniques were accurate in assessing the genetic relationship among the genetic materials in this study. Therefore, the inter- and intra-cluster variations of genotypes generated by both methods can be useful in the selection of unrelated genotypes in relation to biomass yield and sugar-related traits for hybrid development. The obtained results can further enable breeders to maximize the genetic variability in the breeding program, and to produce good segregating progenies.

3.6 Conclusion

The use of SSR markers in this study showed that there is genetic variation among sweet sorghum lines obtained from ICRISAT and ACCI. Four potential heterotic groups (C I, C II, C III and C IV) with one singleton were identified using the SSR markers. Therefore, appropriate parental lines can be selected from the clusters for hybrid development. Genetic diversity can further be maintained by subsequently making crosses between clusters and between sub-clusters within the broad clusters.

3.7 References

Ali, M., J. Rajewski, P. Baenziger, K. Gill, K. Eskridge, 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.

Amelework, B., H. Shimelis, P. Tongoona, M. Laing, and F. Mengistu. 2015. Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. Plant Genetic Resources 13:131-141.

Antonopoulou, G., H.N. Gavala, I.V. Skiadas, K. Angelopoulos, and G. Lyberatos. 2008. Biofuels generation from sweet sorghum: fermentative hydrogen production and anaerobic digestion of the remaining biomass. Bioresource Technology 99:110-119.

Aremu, C. 2012. Exploring statistical tools in measuring genetic diversity for crop improvement. INTECH Open Access Publisher.

Assar, A., R. Uptmoor, A. Abdelmula, M. Salih, F. Ordon, and W. Friedt. 2005. Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). Crop Science 45:1636-1644.

Avise, J.C. 2012. Molecular markers, natural history and evolution. Springer Science & Business Media. Chapman and Hall, New York.

Cholastova, T., M. Soldanova, and R. Pokorny. 2013. Random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) marker efficacy for maize hybrid identification. African Journal of Biotechnology 10:4794-4801.

Deu, M., F. Sagnard, J. Chantereau, C. Calatayud, Y. Vigouroux, J.-L. Pham, et al. 2010. Spatio-temporal dynamics of genetic diversity in *Sorghum bicolor* in Niger. Theoretical and Applied Genetics 120:1301-1313.

DeWoody, J., R. Honeycutt, and L. Skow. 1994. Microsatellite markers in white-tailed deer. The Journal of Heredity 86:317-319.

Elangovan, M., N. Seetharama, and J. Patil. 2014. Genetic diversity and heritability characters associated in sweet sorghum [*Sorghum bicolor* (L.) Moench]. Sugar Technology 16:200-210.

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

Ganapathy, K., S. Gomashe, S. Rakshit, B. Prabhakar, S. Ambekar, R. Ghorade, et al. 2012. Genetic diversity revealed utility of SSR markers in classifying parental lines and elite genotypes of sorghum (*Sorghum bicolo*r L. Moench). Australian Journal of Crop Science 6:1486-1493.

Geleta, L., M. Labuschagne, and C. Viljoen. 2005. Genetic variability in pepper (*Capsicum annuum* L.) estimated by morphological data and amplified fragment length polymorphism markers. Biodiversity & Conservation 14:2361-2375.

Geleta, N., M.T. Labuschagne, and C.D. Viljoen. 2006. Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers. Biodiversity & Conservation 15:3251-3265.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3).

Govindaraj, M., M. Vetriventhan, and M. Srinivasan. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics Research International 2015.

Kamuntu, S.P. 2010. Phenotypic and genetic diversity of sweet sorghum (*sorghum bicolor* I.) varieties and their suitability for ethanol production in Tanzania. Morogoro. Sokoine MSc thesis, University of Agriculture, Tanzania.

Karp, A., D.S. Ingram, and P.G. Isaac. 2012. Molecular tools for screening biodiversity: plants and animals. Springer Science & Business Media, Devon.

Khalil, S.R., A. Abdelhafez, and E. Amer. 2015. Evaluation of bioethanol production from juice and bagasse of some sweet sorghum varieties. Annals of Agricultural Sciences 60:317-324.

Kong, L., J. Dong, and G. Hart. 2000. Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). Theoretical and Applied Genetics 101:438-448.

Kumar, C.G., S.K. Mamidyala, B. Sreedhar, and B.V. Reddy. 2011. Synthesis and characterization of gold glyconanoparticles functionalized with sugars of sweet sorghum syrup. Biotechnology Progress 27:1455-1463.

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.

Mace, E.S., K.K. Buhariwalla, H.K. Buhariwalla, and J.H. Crouch. 2003. A high-throughput DNA extraction protocol for tropical molecular breeding programs. Plant Molecular Biology Reporter 21:459-460.

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

Mofokeng, A., H. Shimelis, P. Tongoona, and M. Laing. 2014. A genetic diversity analysis of South African sorghum genotypes using SSR markers. South African Journal of Plant and Soil 31:145-152.

Motlhaodi, T.M. 2016. Genetic diversity and nutritional content of sorghum [*Sorghum bicolor* (L.) Moench] accessions from Southern Africa. PhD dissertation. Swedish University of Agricultural Sciences, Sweden.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Ng'uni, D., M. Geleta, and T. Bryngelsson. 2011. Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) accessions of Zambia as revealed by simple sequence repeats (SSR). Hereditas 148:52-62.

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

Olweny, C., J. Jamoza, M. Dida, W. Kimani, J. Njuguna, D. Githae, et al. 2014. High genetic diversity for improvement of sweet sorghum (*Sorghum bicolor* (L.) Moench) genotypes for sugar and allied products. Molecular Plant Breeding 5:29-35.

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

Pei, Z., J. Gao, Q. Chen, J. Wei, Z. Li, F. Luo, et al. 2010. Genetic diversity of elite sweet sorghum genotypes assessed by SSR markers. Biologia Plantarum 54:653-658. doi:10.1007/s10535-010-0116-x.

Rao, S., S. Rao, N. Seetharama, A. Umakath, P.S. Reddy, B. Reddy, et al. 2009. Sweet sorghum for biofuel and strategies for its improvement. Information Bulletin No. 77. International Crops Research Institute for the Semi-Arid Tropics, Patancheru - 502 324, A.P, India.

Reedy, S., and V. Khanna. 2010. Molecular fingerprinting of sorghum (*Sorghum bicolor* L. Moench) germplasm using simple sequence repeat markers. Pantnagar Journal of Research 8:180-186.

SAS, S. 2011. Windows Version 9.3. SAS Institute Inc., Cary, NC, USA.

Shehzad, T., H. Okuizumi, M. Kawase, and K. Okuno. 2009. Development of SSR-based sorghum (*Sorghum bicolor* (L.) Moench) diversity research set of germplasm and its evaluation by morphological traits. Genetic resources and crop evolution 56:809-827.

Singh, B., and A.K. Singh. 2015. Marker-assisted plant breeding: principles and practices. Springer, India.

Soriano, J.M., C. Romero, S. Vilanova, G. Llácer, and M.L. Badenes. 2005. Genetic diversity of loquat germplasm (*Eriobotrya japonica* (Thunb) Lindl) assessed by SSR markers. Genome 48:108-114.

Taramino, G., R. Tarchini, S. Ferrario, and M. Lee. 1997. Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theoretical and Applied Genetics 95:66-72.

Tomar, S.S., S. Sivakumar, and K. Ganesamurthy. 2012. Genetic variability and heritability studies for different quantitative traits in sweet sorghum [*Sorghum bicolor* (L.) Moench] genotypes. Electronic Journal of Plant Breeding 3:806-810.

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.

Turki, D.A., M. Al Jamali, and A. Kanbar. 2011. Genetic variation of sorghum (*Sorghum bicolor* L. Moench) varieties assessed by ISSR markers. Advances in Environmental Biology 5:3504-3511.

van Berloo, R. 2008. GGT 2.0: versatile software for visualization and analysis of genetic data. Journal of Heredity 99:232-236.

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

4 A general overview of the research findings

4.1 Introduction and objectives of the study

South Africa is known to be the largest emitter of greenhouse gases in Africa due to its energy intensive economy. The increase in energy demand, volatile oil prices and climate change has led the country to reduce its dependency on fossil fuels and promote biofuels. In the interest of contributing to a sustainable renewable bioethanol production, there is need to establish genetic variability in sweet sorghum. This is in order to select parents for hybrid development to maximise biomass yield and sugar-related traits. The objectives of the present study are highlighted and achieved from which suggestions can be drawn based on the obtained results. The chapter gives a summary of the core findings and their implications in developing superior sweet sorghum genotypes that will be used to extract ethanol needed in the South African biofuel industry.

To summarize, the objectives of this study were:

1. To assess sweet sorghum lines for agronomic performance and genetic diversity using quantitative morphological traits.
2. To assess sweet sorghum lines for genetic diversity and interrelationships using simple sequence repeat (SSR) markers.

4.2 Summary of the major findings

4.2.1 Assessment of sweet sorghum lines for agronomic performance and genetic diversity using quantitative morphological traits

- The sweet sorghum lines revealed highly signficant variations for 13 quantitative characters assessed in this study. The extent of variation was highly influenced by environment and genotype by environment interaction.
- Genotypes designated as IS 2331, IESV 92008 DL, ICSV 700 AS 244, URJA and SS 27 were identified as suitable genotypes with high plant height, dry matter, fiber, ^{o}brix, ^{o}total brix, total fermentable sugars and ethanol. The specified genotypes also exhibited medium to late maturity with relatively high fresh biomass and fresh stalk yield.
- Genotype 91018 LT showed the highest fresh biomass yield, fresh stalk yield, stalk diameter and relatively high grain yield.

- High levels of trait heritability were observed for fresh stalk yield (98%), stalk diameter (93%), fresh biomass yield (81%), panicle length (76%), fiber (73%) and plant height (66%). Heritability estimates were influenced by the environment and genotype by environment interaction.
- The first three principal components showed 83% of the total variability among the genotypes. Ethanol, total fermentable sugars, ^{o}total brix, fresh stalk yield and ^{o}brix contributed mainly to PC 1, whereas fresh biomass yield and stalk diameter contributed mainly to PC 2.
- The dendrogram generated from cluster analysis divided the genotypes into two main clusters and three singletons (ICSB 324, ICSB 654 and ICSV 700). Cluster I comprised 54% of the total germplasm and included only one ACCI genotype (SS 27), while cluster II comprised of 33% of the total variation.

4.2.2 Assessment of sweet sorghum lines for genetic diversity and interrelationships using simple sequence repeat markers

- The morphological variability examined above was complemented by high molecular markers diversity. Variation was observed for all the markers with allelic size ranging from 1 to 36 bp. A total of 61 alleles were generated with an average of 6.1 alleles per locus.
- The PIC values ranged from 0.32 to 0.86 with an overall mean value of 0.62, showing a high discriminating ability of the markers used.
- The genetic distance ranged from 0.5 to 1.9 with AS 244 (GD = 1.9) showing the largest genetic distance, while IESV 92001 DL and IESV 92008 had the smallest genetic distance (GD = 0.5).
- The dendrogram generated from cluster analysis using SSR markers classified the 24 sweet sorghum lines into two major clusters. Cluster I comprised of 13% of the total genotypes which included URJA, SS 27 and ICSB 654. It was observed that all ACCI genotypes apart from AS 244 were grouped in Cluster I, with URJA and SS 27 being very closely related.
- Cluster II was observed to be the largest (88%) with 21 genotypes, which further formed 3 sub-clusters (A, B and C) and a singleton (AS 244).
- The same results were obtained using PCA analysis which grouped genotypes into four clusters with same type of genotypes in each group.

4.3 Implications of the research findings in breeding for sweet sorghum cultivars with high biomass yield and sugar-related traits

- The genetic variability examined for biomass yield and sugar-related traits in sweet sorghum lines obtained from ICRISAT-Kenya and ACCI-South Africa implied that there is huge potential for selection of source germplasm for improved sweet sorghum cultivar breeding needed in the South African biofuel industry.
- The two dendrograms generated from cluster analysis using phenotypic traits and SSR markers grouped the sweet sorghum lines into different clusters. Genotypes like ICSB 654 and ICSR 93034, which are morphologically different were not found to be distinct using SSR markers. This is expected because morphological traits are highly influenced by environmental conditions and have a major impact on the relative performance of genotypes. In addition, the genotypes may not be genetically different from each other and the morphological differences between these genotypes may have little genetic basis. Therefore, the use of both phenotypic traits and molecular markers is prudent in providing more representative sampling of the genome
- The three distinct sub-groups observed from cluster analysis using SSR markers would enable plant breeders to select distinctive alleles and exploit the potential of transgressive segregation between the two sub-groups.
- Crossing distinct genotypes such as SS 27, IESV 92008 DL, ICSV 700 and AS 244 possessing highest quality traits and plant height with IESV 91018 LT having highest yield components and thickest stalk may result in a good combination for selection of progenies with desirable characteristics to be used as superior varieties needed for bioethanol production.
- Bioethanol production from sweet sorghum in South Africa will offer income generating opportunities for small holder farmers who will be key players in its production and in turn boost the socio-economic sectors. At national level, the bioethanol production from sweet sorghum will open up new markets, and generate new technologies and industries which will boost the country's economy.

4.4 Challenges encountered and recommendations

Although assessment of sweet sorghum lines for genetic diversity using quantitative traits and SSR markers was successful in this study, challenges were encountered. Proper sowing dates need to be explored to prevent depressed plant height which has a major impact on fresh biomass yield, fresh stalk yield and the sugar quality traits. The study was also carried out in a single season and in two locations to test for the potential of biomass yield and sugar-

related traits and for identification of promising genotypes for possible breeding programme. It is recommended that more multi-locational trials in different seasons should be conducted to evaluate the performance of these improved sweet sorghum varieties before recommending the best to farmers. Apart from that, genotypes possessing high biomass and stalk yields with low sugar content should be crossed with sweet sorghum genotypes having low sugar content but with high biomass and stalk yields. This is in order to maximize biomass yield and sugar-related traits for hybrid development, and the development of segregating populations to map genes controlling sugar content in sweet sorghum.