GENETIC STRUCTURE, PHENOTYPIC DIVERSITY AND SALINITY TOLERANCE POTENTIAL OF SELECTED SWEET AND GRAIN SORGHUM POPULATIONS

JOHN MBALUKA KISUA

MASTER OF SCIENCE

(Genetics)

JOMO KENYATTA UNIVERSITY OF

AGRICULTURE AND TECHNOLOGY

2017

Genetic Structure, Phenotypic Diversity and Salinity Tolerance Potential of Selected Sweet and Grain Sorghum Populations

John Mbaluka Kisua

A Thesis Submitted in Partial Fulfillment for the Degree of Master of Science in Genetics in the Jomo Kenyatta University of Agriculture and Technology

2017

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University

Signature…… Date …….

John Mbaluka Kisua

This thesis has been submitted for examination with our approval as University supervisors.

Signature…… Date …….

Prof. Martha Makobe, PhD

JKUAT, Kenya

Signature…… Date …….

Dr. Kaibui Mwikamba, PhD

JKUAT, Kenya

Signature…… Date …….

Prof. Anne Muigai, PhD

JKUAT, Kenya

DEDICATION

To my mum Katûkû for inspiring me and my late father Kisua for instilling discipline in me.

ACKNOWLEDGEMENT

Special thanks to Professor Martha Makobe, my first supervisor and a senior member of sorghum research community in Kenya for the original concept of exploring the genetic diversity of sorghum from diverse background. Her suggestion for looking at a new way of application of sound statistical tools in analysis of population genetics research in the genetics department at the School of Biological Sciences in Jomo Kenyatta University of Agriculture and Technology was the problem that provided the original idea for my research work. She suggested to me that this study was a good project and demonstration of stringency and accuracy in analysis of data would elevate research level at Department of Genetics. It is my greatest pleasure to discuss research work with her. Her knowledge, boundless enthusiasm and encouragement saw me through the tough times of my research.

Professor Anne Muigai, my supervisor and the leader Department of Genetics deserves the credit for her expert guidance and the ideas behind presentation of molecular work and genetics in general for my thesis. It is the great deal of help from her that gave me a most enlightening experience. She kept inaccuracies and omissions of my thesis at minimum. It is because of her patience that i give my utmost gratitude. I am eternally grateful to this professor.

Thanks also need to be given to Dr. Kaibui my supervisor, department of Research Production and Extension for financial support through KAPPI funding. He greatly facilitated my life with travel in the arid and semi arid lands of Kenya, to study this orphaned but food security crop growing in real situations. I was aided immeasurably by his advice. My thanks to the lab technicians in JKUAT, in particular to Rose Mingate, Mr. Ndakala, Richard Rotich and Mr. Mbogo and also Mr. Ngesu of Ministry of Agriculture in Kathonzweni and Agricultural officers of Rongai Sub Counties for generously sharing their knowledge with me. Big thanks to my mother. She kept inspiring me to persevere during tough times in my research work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ABBREVIATIONS

ANOVA: Analysis of Variance

CTAB: CetylTrimethyl-Ammonium Bromide x

DNA: Deoxyribonucleic acid.

EDTA: Ethylene Diamine Tetra acetic Acid

FAOSTAT: Food and Agriculture Organization of the United Nations

HWE: Hardy-Weinberg Equilibrium

IBPR: International Board for Plant Genetic Resource

ICRISAT: International Crop Research Institute for Semi-Arid Tropics

JKUAT: Jomo Kenyatta University of Agriculture and Technology

KALRO: Kenya Agricultural and Livestock Research Organization

PCR: Polymerase Chain Reaction.

RCBD: Randomized Complete Block Design

ROS: Reactive Oxygen Species

SSR: Simple sequence repeats

STR: Short Tandem Repeats

UNJ: Un-weighted neighbour joining

UPGMA: Un-weighted Pair Group Method using Arithmetic Averages.

ABSTRACT

Sorghum (*Sorghum bicolor* L. Moench), a diploid with 2n = 20, has a genome estimated to contain 811Mbp of DNA. *Sorghum* Moench belong to the botanical family Poaceae and sub-generic Eusorghum. It is a subject of plant population genetics research based on its importance as one of the world's leading grain crops and of desirable syrup. Sorghum cultivars (*S. bicolor*) classified within primary gene pool, are genetically impoverished, therefore determination of genetic divergence based on phenotypic diversity, salinity tolerance potential, heritability and molecular diversity, is timely and crucial. The findings, as revealed in this study could be utilized in selection of heterotic parents for crossbreeding. Seven qualitative and 16 quantitative traits were measured and the phenotypic data analyzed. A number of fitness traits including seedling dry weight, germination percentage, root length, shoot length, shoot length and wet weight were measured under different levels of selection pressure. The hydroponic growth medium was contaminated with NaCl at different (EC5, EC 10, and EC 15) concentration levels. The phenotypic variance, broad sense heritability (h^2) and genotypic variance for all characters based on salinity levels was calculated. Molecular genetic diversity study involved use of microsatellite markers where cetyltrimethyl ammonium bromide (CTAB) protocol was used for extraction of DNA. Gel electrophoresis, DNA quantification, polymerase chain reaction and SSR markers optimization was done. The data obtained was analyzed using statistical softwares; GenStat v.12, DARwin v.6 and PowerMarker v.3.25. The data were standardized prior to cluster analysis. Statistical methods employed clearly discriminated these sorghum genotypes into distinct groups. Phenotypic diversity analysis distinctively demarcated grain sorghums from sweet sorghums. The first four principal components accounted for 97.19 % of the total variation with PC4 scoring 1.14Eigen value. Both the biplot and dendrogram formed through (HCA) classification with UPGMA technique sorted genotypes into 3 main clusters. Three sweet sorghums 3018, 3022 and 3008 gathered in one sub-cluster. Genetic distance matrix based on phenotypic traits ranged from 2.36 to 8.80. The phenotypic diversity index ($\bar{H}$) for traits studied recorded a mean value of 0.70

and a maximum 0.95.Based on tolerance index Seredo (83.31%) was identified as a tolerant population and Serena (74.09%) as a considerably sensitive tolerant population. Tolerance index (TI) showed a decrease in TI (5 dS/m>10 dS/m >15 dS/m) with increase in salinity stress levels. The biplot display showed that Kivila, a landrace population had a close relationship with 3008 and 3022 sweet sorghums. The traits were projected by vectors thereby indicating wide salinity tolerance diversity. The genetic distance between Seredo (2.88) and Serena showed a narrow genetic base indicating that they could have descended from a common ancestral population. Minimum distance coefficient (1.14) of Manhattan dissimilarity was recorded between Kivila and 3022. Coefficient of variance values as well as scores for heritability was higher at control than that of respective selection pressures (5dS/m, 10dS/m and 15dS/m). Heritability in broad sense values ranged from 38.57 % for wet weight to 62.58 % germination percentage at 15dS/m. The SSR markers used were polymorphic (50%) and informative with PIC value ranging from 0.35 to 0.37 with a mean of 0.36. The frequency of the major alleles at different SSR loci ranged from 0.50 to 0.64. Recorded expected heterozygosity (H_e) ranged from 0.46 to 0.50 with an average of 0.47 whereas heterozygosity (H_O) values observed ranged from 0.43 to 0.89 scoring a mean of average of 0.66. Hardy-Weinberg equilibrium computation produced a highly significant p value (0.047) for locus Xtxp 021. Coordinate (Jaccard similarity coefficient) calculated for the two first axes had positive Eigen values, 0.08517 and 0.06173 for the first and second coordinates in that order. PCoA provided a better diversity structure than the dendrogram. Ultimately the patterns of genetic diversity and relationships of sorghum revealed in this study is insightful in providing useful information for identification and utilization of the genetic resource in breeding programmes to break performance barriers in sorghum.

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Sorghum [*Sorghum bicolor* (L.)Moench], a diploid with 2n = 20, a progenitor of one of the world's most noxious weeds (*S. halpense*), is the first sequenced plant genome of African origin (Mutegi *et al*., 2011). The *Sorghum bicolor* genome is estimated to contain 811Mbp of DNA (Price *et al*., 2005). The molecular size of each chromosome and size-based nomenclature for sorghum chromosomes (SBI-01–SBI-10) and linkage groups (LG-01 to LG-10) has been estimated based on chromosome size, arm length, and arm orientation (Howell *et al.,* 2002). Its small genome makes sorghum an attractive model for plant genetics studies. The existence of high levels of DNA polymorphism and divergence within *Sorghum bicolor* allows inferences about the possible role of selection in the evolution of particular loci (Paterson, 2008).

Understanding how genotypic variation translates into phenotypic variation and how it is shaped in populations and species is fundamental to our understanding of sorghum population genetics (Ayana & Bekele, 1998). Furthermore, genetic techniques examine individual variation to discern the emergent properties of populations. The central theme of population genetics is to make sense of these variations, which includes the inference of past evolutionary histories as well as the dissection of molecular mechanisms leading to the variation we observe today (Folkertsma *et al.,* 2005). In order to infer the diversity within sorghum genotypes under study, the choice of statistical models is prerequisite, whether the research is done in the laboratory or in the field. Therefore, the need to study the behavior of alleles using formal mathematical analysis of the properties of ideal populations, and experimental population genetics, which examines the behavior of real alleles in natural or laboratory populations, is key.

Phenotypic diversity index of qualitative traits and multivariate analysis of quantitative traits have been used to measure genetic relationships within sorghum populations. Previous researchers such as Abdi *et al.* (2002) conducted studies on phenotypic diversity in sorghum using phenotypic markers. Phenotypic traits such as stem height, leaf area, and panicle size, number of internode and stalk diameter are subject to effects of genes and the environment (Geleta *et al.*, 2006). Studies on diversity in sorghum genotypes based on these traits have provided evidence that sorghum has appreciable genetic variation that has been poorly used in terms of crop improvement programme (Dillon *et al*., 2007). Furthermore, the relationships with historic sweet cultivars and grain sorghums are poorly documented. The pattern of genetic relationships between and within such cultivars, based on phenotypic polymorphism could be shown by multivariate analyses.

Salinity tolerance represent combined effects of genetic and environment which include the integration of physiological mechanisms that confer sorghum genotypes ability to survive salinity stress. Salinity stress to sorghum in arid and semi arid lands generates some potential effects on plant tissues, including osmotic stress, ionic stress and oxidative stress. Conventional selection and breeding techniques have commonly been used by previous researchers on improvement for salinity tolerance character in sorghum (Weimberg *et al.,* 1984). Typical selection traits included germination percentage, germination rate, and root and shoot length and fresh root and shoots weights (Jamil *et al*., 2006). Therefore, the analyses of data on salt stress effects on seedling growth are based on fitness depending on data treatment.

Molecular analysis remains superior in analysis of genetic diversity in sorghum crop (Ayana & Bekele, 1998). Available molecular techniques could prove useful in the detection variation of heritable traits present in sorghum crop (Agarwal *et al.* 2008). It may not be feasible to measure the exact genetic state of an individual which constitute the DNA sequence of all its chromosomes in genetic diversity studies (Kong *et al*., 2000). It is sufficient getting information on the states of a handful of specific sites along the genome which exhibit polymorphism (Hartl & Clark, 1998). These indicate the presence of variant states (marker) or alleles in the population

under study that could be utilized in detection of underlying diversity and attribute this to forces that contribute to variability (Iqbal *et al.,* 2010). With the sorghum genome sequence available, one can anticipate renewed interest and accelerated progress in relating sorghum genes to their functions (Carrera *et al*., 1996).

A number of DNA markers have been applied to determine population divergence and their utilization could generate novel sources of genes providing superior combinations of sorghum genotypes (Mehmood *et al.,* 2008; Kumar *et al.,* 2008). Different types of DNA markers have been used in many studies of sorghum. For example AFLPs (Ritter *et al.,* 2007), RAPDs (Iqbal *et al.,* 2010), RFLPs (Ahnert *et al.,* 1996), and SSRs (Folkertsma *et al.,* 2005) have been successively used to determine genetic divergence among sorghum germplasm.

The genetic diversity existing in cultivated grain and sweet sorghum in semi-arid tropical environments represents one of the world's most important genetic resources for future sorghum breeding efforts and global food security. Knowledge of the levels and patterns of genetic diversity in this curious group under this study will help to reveal the path by which to follow so as to improve and commercialize the production of both grain and sweet sorghum in the arid and semi-arid lands in Kenya. Furthermore, the genetic diversity of these genetic resources is specifically important to farmers, breeders and other users in the strategic planning of exchange of materials and in the identification of particular divergent sorghum populations. Therefore, the need to provide baseline data on the genetic diversity of both grain and sweet sorghum populations is crucial. Finally, a relatively advanced state of knowledge of the genetic diversity of sorghum, add to its promise as a salt tolerant crop for arid and semi arid lands of lower eastern Kenya and offers novel learning opportunities. It is this baseline knowledge that could provide a good beginning for designing a sound breeding programme for sorghum crop improvement.

1.2 Taxonomy

Sorghum Moench belong to the Andropogoneae tribe in the botanical family Poaceae. The genus is divided into five sub-generic sections: Eusorghum,

Parasorghum, Heterosorghum, Chaetosorghum and Spitosorghum. Based on the gene pool concept, the genus *Sorghum* has been classified into three gene pools. The primary gene pool includes *S. bicolor* (2n = 20) according to Stenhouse *et al*. (1997). The secondary gene pool consists of *S. halepense*. . (L.) Pers. (2n = 40). According to Mutegi *et al.* (2011), the primary and secondary gene pools of sorghum, which include the cultivars and their wild and weedy relatives, are classified within the section Eusorghum. The tertiary gene pool of the genus *Sorghum* includes all wild sorghum belonging to subgenera *Chaetosorghum*, *Heterosorghum*, *Parasorghum* and *Stiposorghum* (Ng'uni *et al*., 2012).

According to records of Food and Agriculture Organization (2011), sorghum ranks fifth in the world after wheat, rice, maize and barley in terms of production, but in Africa the crop is the second highly produced cereal crop after maize. It is a food security crop providing dietary staple for many people, especially in the semi-arid tropics (Iqbal *et al.,* 2010).The crop's wide adaptability to conditions such as drought, water logging and salinity makes sorghum a crop of choice in marginal soils where growth of other cereals such as maize cannot be supported (IBPGR & ICRISAT, 1993). Africa is considered as both the continent of domestication and a centre of genetic diversity for cultivated sorghum (Li *et al*., 2010). Numerous varieties of sorghum were created through the practice of disruptive selection, whereby selection for more than one level of a particular trait within a population occurs.

1.2.1 Uses of sorghum

Sorghum in these areas is mainly consumed in the forms of couscous, porridge and fermented drinks. The sorghum grain is used for baking of unleavened bread, biscuits and tortilla in the confectionary industry (Schober *et al*., 2005). Sorghum is an important alternative grain to use in gluten-free diets (Grenier *et al*., 2004).The crop has a wide range of uses, such as beer brewing and livestock feed and fodder (Chakauya *et al.*, 2006), in the confectionary, and in the production of commercial alcohol, adhesives, waxes, construction materials and bio-ethanol from sweet

sorghum varieties (Antonopoulou *et al.*, 2008).Sorghum is mainly a rain-fed crop of lowland, semi-arid areas of the tropics and sub-tropics (Craufurd *et al.*, 1999). Sorghum is adapted to drought conditions due to a number of morphological and physiological characteristics, including an extensive root system, waxy bloom on leaves that reduces water loss, and the ability to stop growth in periods of drought and resume it when the stress is relieved (Stenhouse *et al.*, 1997).

1.3 Statement of the problem

World Hunger Education Service, 2012 data shows that the number of hungry people in the world is estimated to be 925 million. Moreover, Sub-Saharan Africa is reported to have the highest prevalence of nutritional related ailments in the world and more than 2 billion people in the world have been affected according to FAO (2011) report. Increased use of fertile agricultural lands for human activities other than crop production pushes sorghum crop cultivation, a staple food for poor people in arid and semiarid lands of Kenya, to less productive lands in the marginal areas. Furthermore, demand for biofuel, in light of Kenyan vision 2030 has caused a re-evaluation of sweet sorghums as a source of energy. Despite its importance, sorghum productivity is far below the genetic potential of the crop. There are few genetic studies that have investigated grain sorghum germplasm including some sweet sorghums. Currently, there are no discrete objective criteria, such as a molecular marker or sugar concentration level, to differentiate sweet sorghums from grain sorghums. This is a major obstacle to their potential use in breeding programs. Additionally, little has been documented about the genetic relationships and diversity within sweet sorghums and how sweet sorghums relate to grain sorghums. It is also known that, sorghums are extremely diverse phenotypically, genetically and geographically. However, there is an apparent lack of knowledge on the allelic frequency and information on Hardy Weinberg Equilibrium of various grain and sweet sorghum populations within and between various eco-geographical regions. Numerous varieties of sorghum in Kenya have been continuously selected by both farmers and research institutions. As such selection generally removes genetic

variation from a population suggesting that there may be a narrow genetic base for Kenyan sweet and grain sorghum cultivars.

1.4 Justification

Sorghum is a pillar of food security in the arid and semi-arid regions of Kenya (Makobe *et al.,* 2006). It provides steady income for poor farmers and this makes it an ideal crop for arid and semi-arid lands, contributing to household food security. Sorghum grain yield and quality are challenged by salt stress and the use of phenotypic, seedling traits, heritability as well as molecular techniques (SSR) in assessing genetic diversity is precise and fast in selection of high yielding and salt tolerant cultivars. These markers provide an efficient way to effectively utilize favorable genes. Additionally, the need to combine seedling traits under salt conditions and genetic aspects of sorghum demands studying this crop growing in real situation. Furthermore, migration is known to bring variation into a population. A population geneticist carrying out this research in Makueni County plays this vital role of expanding variation by use of the effects of migration in the research programs by introduction. This justifies the choice of study site, coupled with the fact that other cereal crops like maize do not do well here. Little has been documented about the genetic relationships and diversity within sweet sorghums and how they relate to grain sorghum racial types. Sorghums are extremely diverse phenotypically, genetically and geographically. However, there is an apparent lack of knowledge on the genetic diversity of grain and sweet sorghum populations. This is a major obstacle to both their effective conservation and potential use in breeding programs. Moreover, assessment of genetic variation among sorghum accessions is a gateway to the study of evolutionary forces that has strong impact on conservation and breeding. Therefore, the need to provide baseline data on the genetic diversity of both grain and sweet sorghum populations under study is both urgent and justified. Knowledge of the levels and patterns of genetic diversity and population structure in this curious group under this study will help to reveal the path by which to follow so as to improve and commercialize the production of both grain and sweet sorghum in the arid and semi-arid lands in Kenya. Furthermore, the genetic diversity of these

genetic resources is specifically important to farmers and breeders in the strategic planning of collections, exchange of materials and in the identification of particular divergent accessions with pertinent traits of interest.

1.5 Hypothesis

There is no molecular diversity, phenotypic diversity, broad sense variation of traits, and salinity tolerance variability among the studied genotypes

1.6 General objective

To determine genetic structure, phenotypic diversity, broad sense variation of traits, and salinity tolerance variability among the studied genotypes

1.6.1 Specific objectives

1. To determine phenotypic diversity of sweet and grain sorghum genotypes using phenotypic markers
2. To determine salinity tolerance variability of sorghum genotypes under hydroponic salt stress
3. To estimate broad sense heritability of sorghum seedlings in salt stress conditions
4. To determine genetic structure of sorghum genotypes using allelic frequency in sorghum populations

CHAPTER TWO

LITERATURE REVIEW

2.1 Evolutionary and population genetics in sorghum

Research showed that, *Saccharum officinarum* L. (sugarcane), *Zea mays* L. (maize), and *Sorghum bicolor* which belong to grass family diverged from rice approximately 50million years ago. Furthermore, analysis of nuclear rRNA, molecular data shows *S. bicolor* and *S. officinarum* are closely evolutionary related. Their divergence from a common ancestor is recent, roughly estimated to be 5 million years ago according to Sobral *et al*. (1994).

As the search for potentially useful alleles of sorghum moves from landraces to cultivars or from natural populations to cultivars, machinery from evolutionary and population genetics may prove useful in explaining the changes in terms of selection. The issues of concern to population geneticists involve estimating the nature and amount of selection on a defined suite of characters and the population genetics of its evolution. Furthermore methods for estimating the nature of natural selection on any traits may be of interest to plant geneticists. Plant geneticists therefore can estimate the nature of natural selection on any measurable suite of characters, separating selection into direct and indirect effects (Willis, 1996). A detailed understanding of the nature of natural selection in either improved cultivated or landrace sorghum populations can provide the geneticists with valuable insight into characters that can further improve performance.

According to Ulanch *et al.* (1996), there is a rich literature from population genetics dealing with detection of selection from a sorghum population sample of DNA sequences. An interesting application of these methods would be the finding of the levels of polymorphism in *Sorghum bicolor* (Mehmood *et al*., 2008). Other issues of concern to quantitative geneticists include estimating the degree of relationship based on molecular data and efforts to be made on collection of candidate alleles associated with certain characters. Therefore the study of genetic polymorphism (DNA

variation) and its organization at various levels is of importance as it can reveal how genetic variability is organized among sorghum populations.

Fundamental principles of population genetics play a role of revealing the levels of relatedness among individuals through analysis of inbreeding (Nei *et al.,* 1975). Moreover, newer techniques exist, based on the very high levels of variation found with simple sequence repeats. Characterization of basic population genetic parameters is an essential prerequisite to any approach that analyzes variation in population samples. Population genetic analysis can also provide an identification of loci in sorghum genome that have been targets of selection as reported by Mutegi *et al.* (2011)

When a DNA sequence change arises, depending on its genetic context and environment, its effect can be neutral, deleterious (negative) or advantageous to the reproductive success (fitness) of a species. Based on this idea, when a variant confers a higher fitness its frequency will tend to increase and it will eventually become the only variant present in the whole population (it becomes fixed). On the other hand, a variant that is deleterious will tend to disappear from the population. Chakauya *et al.* (2006) suggested that genetic variability in a sorghum population is the result of natural selection acting on mutations and drift. Mating system is an important variable in population genetics. Sorghum is self-pollinating; therefore, it has high levels of linkage disequilibrium (LD) and homozygosity. This mating system reduces the level of heterozygosity (Casa *et al.,* 2005).The amount of heterozygosity across loci, which is synonymous with allelic variation, indicates the amount of genetic variability which has a bearing on the survival of a sorghum population and allows these populations to adapt to changing environments provided that some loci have adaptive values (Ghebru *et al.,* 2002).

Merits of sorghum in population genetic analyses of phenotypes, assessed sequence variation throughout the genome including coding and noncoding sequences and the relationship between polymorphism and divergence (Guo *et al.,* 2014) allows inferences about the possible role of selection in the evolution of particular loci.

Identification of targets of selection may prove valuable in the search for candidate alleles underlying important phenotypes (Al-Sadi *et al.,* 2013). Hardy-Weinberg equilibrium (HWE) can serve as an approximation in many cases or be used to test the violation of any of the conditions. Deviations from HWE can provide important insights into forces that determine the distribution of variation in natural populations (Liu, 2003).

The allelic data obtained in a population genetics study is used to estimate the average number of alleles per locus, percent polymorphic loci, observed and expected heterozygosity and genetic variation measures estimated by Shannon's diversity index (I) and Nei's gene diversity (h). High heterozygosity means lots of genetic variability. Low heterozygosity means little genetic variability (Nei *et al*., 1975). Often, the observed level of heterozygosity is compared to what is expected under Hardy-Weinberg equilibrium (HWE).

Nei *et al*. (1975) indicated that, very low heterozygosities indicate severe effects of small population sizes, that is, population bottlenecks. When a bottleneck occurs in a population, allelic diversity is reduced faster than is heterozygosity, which is a result of loss of rare alleles from the population contributing little to the overall heterozygosity. Rare recessive alleles are usually found in heterozygous individuals in populations, with a selective advantage for the heterozygote, and alleles that have low selection pressure may remain in the heterozygous condition for many generations. If the observed heterozygosity is lower than expected, we seek to attribute the discrepancy to forces such as inbreeding. If heterozygosity is higher than expected, we might suspect an isolate-breaking effect, that is, the mixing of two previously isolated populations.

A microsatellite or short tandem repeat locus is a stretch of sequence, whose location in the genome is defined by a pair of unique sequences at its ends, with a region in the middle containing a repeat sequence. Microsatellite loci are found throughout the genomes of eukaryotes (Tóth *et al.,* 2000). With codominant markers, the heterozygotes show an intermediate or distinguishable from homozygote phenotype.

Codominance will be very important even for non-coding microsatellites because the variants differ in size.

2.2 Phenotypic traits in sorghum diversity study

The variability exhibited by many traits in sorghum fails to fit into discrete phenotypic classes (discontinuous variability), but instead forms a spectrum of phenotypes that blend imperceptively (continuous variability). Economically important traits such as sorghum plant heights, internode length, stem diameter and yield of grain per hectare are quantitative or complex traits with continuous variability (Mukuru, 1993). The basic differences between qualitative and quantitative traits involve the number of genes contributing to the phenotypic variability and the degree to which the phenotype can be modified by environmental factors. Quantitative traits are governed by many genes; each contributing a small amount to the phenotype and their individual effects cannot be detected by Mendelian methods. For this reason, quantitative traits are also referred to as polygenic traits.

Furthermore, each gene usually has effects on more than one trait. Some genes encode products such as enzymes that participate in multistep biochemical pathways or proteins that regulate the activity of one or more other genes in metabolic, regulatory, or developmental pathways. Because of the complex interactions within these pathways, a gene product acting at any one step might have phenotypic effects on other characters. For a given gene, some of its pleiotropic effects may be relatively strong for certain traits, whereas its effects on other traits may be so weak that they are difficult or impossible to identify by Mendelian techniques (Ayana *et al.,* 1998). It is the totality of these pleiotropic effects of numerous loci that constitutes the genetic base of a quantitative trait. In addition to this genetic component, the phenotypic variability of a quantitative trait in a population usually has an environmental component. Heritability analysis can help determine the magnitude of the genetic and environmental components of the total phenotypic variability of each quantitative trait in a population.

Phenotypic characterization was the first method used by researchers to select superior genotypes. These constitute diverse taxonomic characteristics which have been used to assess the patterns of phenotypic diversity in sorghum genotypes. Selection has been carried out based on the phenotype, which is the result of genotypic and environmental effects. Relatively high level of variability exists in both quantitative and qualitative traits within and among sorghum cultivars. Although experimental results may sometimes be applicable over a range of populations or environments, quantitative genetics still serves primarily as a method of local analysis and results are applicable to particular populations in particular environments (Falconer & Mackay, 1996).

Evaluation of genetic diversity can be determined using quantitative and qualitative traits such as maturity, height, midrib colour, yield, panicle length and width, panicle compactness and shape, glume colour, grain colour, size and weight (Geleta *et al*., 2006). Phenotypic traits are simple, rapid and inexpensive to score and measure. Phenotypic estimates show the level of genetic diversity in genotypes. Similarity in the phenotype traits equally presents genetic similarity of genotypes under study (Cox *et al*., 1985). Genetic diversity studies involve the cultivation of sub-samples in real situation so as to comprehend the degree of interaction between genetic and phenotypic traits to establish a suitable selection criterion.

Phenotypic markers are valid in genetic diversity studies due to the fact that, statistical tools for analysis of phenotypic traits are available. Moreover, existing data of genetic resources can be used; this enhances plausible description of heterosis since phenotypic measures of distances included are independent variables. It is a practical means which can be employed in population genetics, selection and genetic diversity studies. Additionally, discrete phenotypic traits are the basis for description of identity, distinctness and uniformity of cultivars in plant variety protection and registration. However, several studies show that phenotypic markers are not suitable for traits with low heritability and are highly affected by environments (Sankarapandian *et al*., 1998). It is difficult to separate genetic and environmental effects for many plant traits. Furthermore, it is an uphill task for the testing

procedures have been reported difficult, tedious, or expensive because of the nature of certain traits.

Some traits in sorghum are controlled by one or two genes but appear to be quantitatively expressed due to low heritability. It is impractical to determine many plant traits at the phenotypic level whether or not a specific gene, or genes, is present in a plant or cultivar. Overall, the efficiency of selection based on phenotypic variability may be reduced by the complex inheritance of multigenic traits and the environment. In other words, the genetic control of phenotypic traits is complex, involving epistatic interactions. Among other limitations, phenotypic markers are known to be recessive with a large number having deleterious effects on traits and most elite cultivated genotypes do not grow vigorously with observable phenotypic markers (Smith, 1987). They cannot adequately describe genotypes without many trials.

2.3 Variability in Salt tolerance of Sorghum Genotypes

Sorghum is among the most versatile of crop species due to its potential for wide adaptation and diversity of end uses. More than half of the world's sorghum is grown in semi-arid tropics of Africa and India, where it is a staple food for millions of poor people as reported by Mehmood *et al.* (2008). Sorghum is an annual C4 plant of African origin and is grown as a source of sugar for syrup. The grain sorghum is mainly consumed in the forms of couscous, porridge and fermented drinks. Despite its importance, sorghum productivity is far below the genetic potential of the crop. Under vision 2030, the discovery of underground water in the arid, Turkana (Kenya) and in other marginal lands of this country, together with the existing water sources, there is a tendency to introduce and implement large-scale irrigation agriculture by many county governments, so as to meet the demands of the ever-increasing human population by elevating productivity. Many Counties in Kenya have salt affected soils; this suggests that, this problem may be severe in some years to come. Moreover, increased use of fertile agricultural lands for human activities other than crop production pushes crop cultivation to less productive lands, including saline

area. Krishnamurthy *et al*. (2007) suggested salinity as one of the major factors that reduce the productivity of sorghum. Nevertheless, sorghum is moderately tolerant to salinity, and a threshold of 6.8 dS m^{-1} and 50% yield reduction at salinity level of 9.9 dS m^{-1}(Shannon, 1997).

The major ions contributing to salinity include the cations Na+, Ca++ and anions Cl- and SO4. Sodium (Na+) ions are the most abundant ion in soil (Ramesh, 2005). According to Noreen (2008), persistent salt stress may generate some potential effects on plant tissues, including osmotic stress, ionic toxicity, mineral deficiencies, physiological and biochemical perturbations. This effect decreases the growth of leaves and roots, even when the concentrations of ions (Na+ and Cl-) in the cells are below toxic concentrations.

Plant species differ in their salt tolerance depending on their genetic makeup. Netondo (2004) studied the effect response of plant growth, ion concentrations in roots, shoots, leaf blades, and sheaths of two sorghum varieties in relation to NaCl salinity. He reported that increasing NaCl concentration significantly reduced the relative shoot growth rate and shoot dry weight.

Salinity like other abiotic stresses alters general metabolic processes and enzymatic activities in sorghum plant. Other effects of salinity on plants include shrinkage of leaf size, reduced ATP, and induced reactive oxygen species (ROS) damage to lipids, proteins and nucleic acids involved in metabolism and cellular homeostasis (Jaleel *et al*., 2008). Makobe *et al.* (2006) successfully demonstrated salinity as a prominent abiotic factor affecting sorghum crop yields in arid and semiarid lands in Kenya.

Classic genetic studies have demonstrated that the ability of plants to tolerate salt stress is a quantitative trait involving the action of many genes. Azhar and McNeilly (1989) found that, for salinity tolerance of young sorghum seedlings, both additive and dominant effects were involved, the latter being of greater importance. A number of genes that encode Na^+ and Cl^- transporters, such as the *HKT*, *NSCCs, PPase* and *NHX* gene families, and specific genes such as *PpENA1* have been studied.

Members of the P450 gene family as well as HKT, *PPase* and *NHX* gene families have been reported to protect sorghum leaves from salinity stress (Ren *et al.,* 2005).

Furthermore, the understanding of the role plaid by the *HKT* gene family in Na+ exclusion from leaves is increasing, as is the molecular basis for many other transport processes at the cellular level. Accordingly, the *HKT* gene-family is expressed in xylem parenchyma cells and protects leaves from salinity stress by removing sodium from the xylem-sap (Uozumi *et al.,* 2000; Zhang *et al.,* 2010). Depending on their genetic makeup, sorghum genotypes have multiple Na+ transport systems to bypass Na+ toxicity. Therefore, such genotypes vary in both the functions of major Na+ transporters and their mechanisms that mediate salinity resistance; the *SOS1* (Salt Overly Sensitive) is a good example. Moreover, proline is one osmolyte that has received much attention and accumulates in sorghum genotypes under salt stress conditions. Recent findings confirmed by genetic studies as well as transgenic studies demonstrated that proline can increase the tolerance of plants to abiotic stress (Tester *et al.,* 2003). It is the genetic basis of these variations in encoding substances by genes that confer genotypes with the ability to tolerate varied levels of salt stress. In other words, sorghum genotypes depend on the broad genetic base of variation for resilience and adaptability for different levels of salinity. Such genotypes respond to the salt stress in part by modulating gene expression, which eventually leads to the restoration of cellular homeostasis, detoxification of toxins and recovery of growth (Volkov & Amtmann, 2006) Seedling characteristics are the most viable criteria used for selecting salt tolerant plants due to the fact that the final plant stand of a crop primarily depends on seedling characteristics (Yilmaz *et al.,* 2004). Exploiting genetic variability to identify salt tolerant genotype is one of the strategies that could be used by geneticists and breeders to overcome salinity. Ali *et al*. (2008) proposed that genotypes possessing salt tolerance characteristics could help in boosting up plants production in salt-affected soils. The growth and yield are measured as determinants of salt stress.

Growing evidence suggests that selection for salt tolerance characteristics in field grown plants is slow and may not be economically feasible. This is primarily because

of the labor-intensive nature of root recovery and quantization, but also because of additional constraints posed by soils which may greatly modify expression of genetic potential. Nevertheless, to proceed with this approach, affirming the presence of genetically based variation for salt tolerance in a particular crop is a prerequisite (Kaya *et al.,* 2007). Few reports deal with genotypic variations in either roots length or shoot distribution patterns among sweet and grain sorghum cultivars.

Geressu, and Gezahagne, (2008) showed that representatives of several sorghum genotypes have a single primary or temporary root in contrast to other cereals. Using a solution culture technique in detailed studies of sorghum root systems, Blum (1979) demonstrated a high level variation between sorghum genotypes. Results of many studies support the suitability of hydroponics as method to investigate details of sorghum root and shoot growth. Faster root extension rate for sorghum seedlings was associated with faster shoot growth as reported by many studies in sorghum. This research attempted to investigate the salinity tolerance variability of selected sweet sorghum and grain sorghum genotypes in hydroponics.

2.4 Heritability Values and Population Divergence

Falconer and Mackay (1996) defined broad sense heritability as the proportion of phenotypic variance that is attributable to an effect for the whole genotype, comprising the sum of additive, dominance, and epistatic effects. Therefore it is a function of the response to selection since it is a useful parameter in quantitative genetics.

None the less, heritability of a population may give information of selection over a few generations because it varies with the changes in allele frequency. Based on this property rare favorable alleles could become more frequent thereby increasing the level of heritability.

According to Tazeen (2009), heritability reveals large variation between traits among populations and through time which could be determined by gene frequencies as well as selection and drift through the evolutionary process.

Substantial differences in broad sense heritability have been reported from previous research findings, for salinity tolerance in diverse crop species. In maize broad sense heritability was 82 % under salt stress conditions (Rao & McNeilly, 1999). Selection of salinity tolerance traits based on genetic variability need be known considering that quantitative traits are under influence of environmental effects. The information on variation components of a trait and the heritability is therefore significant for any sound breeding programme since it determines the scope of selection.

2.5 Sorghum genetics and molecular basis for adaptation

Sorghum *bicolor* is diploid with 20 chromosomes. The complete genetic information of sorghum is distributed in a series of 10 pairs of chromosomes (Figure1.1).These chromosomes have different lengths of DNA sequences. Linkage group nomenclature and chromosomal designations for metaphase chromosomes of *Sorghum bicolor* ($2n = 20$), are usually based on chromosome size, arm length, and arm orientation (Anderson *et al.,* 2003).

The molecular size of each chromosome and size-based nomenclature for sorghum chromosomes (SBI-01–SBI-10) and linkage groups (LG-01 to LG-10) has been estimated. Chromosomes are ordered and designated according to total length at metaphase, namely SBI-01 (longest) to SBI 10(shortest). The three-letter acronym SBI designates the genus and species, and the two-digit numeric code denotes the chromosome number.

Menz *et al.* (2002) represented Linkage groups in sorghum chromosomes as LG-A, LG-B LG-C, LG-D, LG-J, LG-I, LG-E, LG-H, LG-F and LG-G .The consistent use of two digits facilitates data sorting by computers that is LG-01 to LG-10. Therefore, they are numbered from SBI-01 to SBI -10 as follows SBI-01, SBI-02, SBI-04, SBI-05, SBI-06, SBI-07, SBI -08, SBI- 09 and SBI -10. Sorghum *bicolor* has an exceptionally long pair of chromosomes (SBI-01); morphologically SBI-01 is the most distinct chromosome (Figure1.2).

Within a sorghum plant and throughout its life, its genome stays mostly unchanged. However cells make different use of stretches of the DNA, called genes, as templates to build functional cellular products in a process called gene expression. It's for this reason, that, different sorghum genotypes can have very distinct appearances, functions and respond differently to environment.

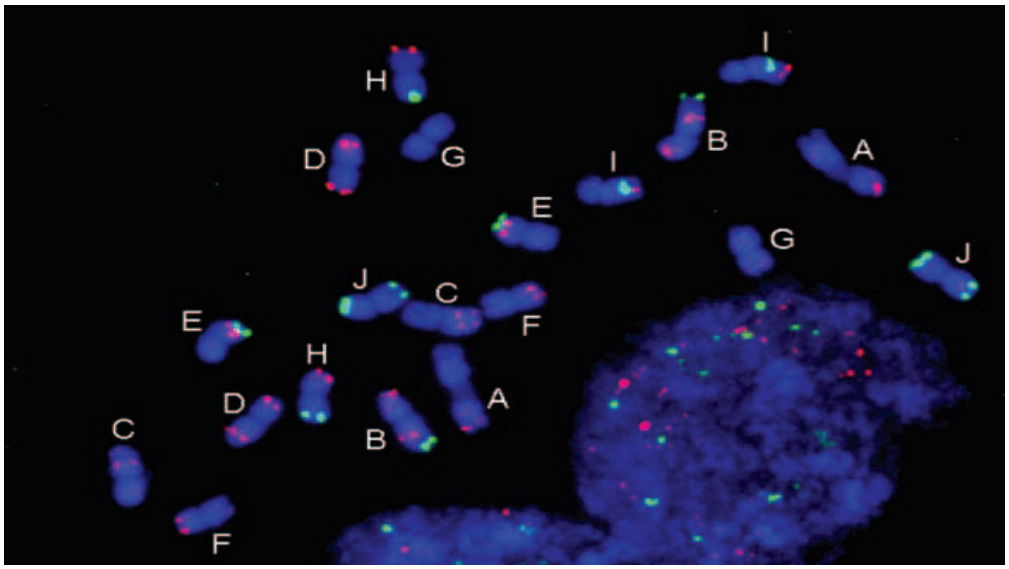

Figure 1.1: FISH-based karyotype of sorghum and the linkage groups showing the various linkage groups of the sorghum genome modified from Menz *et al.*, 2002.

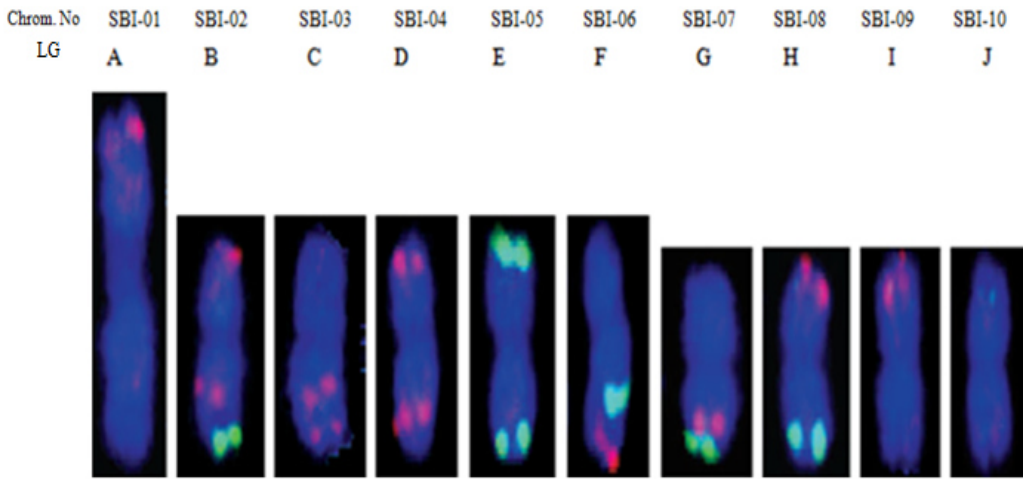

Figure 1.2: Sorghum chromosome number and linkage group modified from Menz *et al.,* 2002.

Blum, (1979) reported that drought resistance in sorghum is a complex trait influenced by many genes coding for various traits contributing towards drought tolerance and this characteristic may be partly related to expansion of one miRNA and several gene families. Studies have shown that, cytochrome P450 domain-containing genes, often involved in scavenging toxins such as those accumulated in response to salt stress, are abundant in sorghum. Additionally, expansins, enzymes that break hydrogen bonds and are responsible for a variety of growth responses that could be linked to the durability of sorghum, occur in sorghum. Furthermore, gene families involved in biotic and abiotic stress tolerance, which do not contain transposons, also harbor enriched large-effect SNPs. It is known that, over-expression of lecithin:cholesterolacyltransferase can increase lipid metabolism and the fluidity of membranes and hence the resistance to heat shock (Ito *et al.* 1997). Shiringani*et* (2010) suggested that five genes in the starch and sucrose metabolism pathway are located on chromosomes 2, 6 and 9. Schnable *et al.* (2009) revealed that transposon elements are abundant in sorghum; with organellar DNA insertion taking 0.085% of the sorghum nuclear genome. It is argued that transposons in sorghum may relocate genes and gene fragments resulting in change in gene structure and

gene activity which in essence is a source of new genetic variation for stressed sorghum populations.

Genetic variation consists of sequence variation and structure alteration which may explain the relevance of various loci that would help understand the various alleles that could define a variety of sorghum populations. This could explain the variation underlying these phenotypes on the basis of their intrinsic nature of their molecular and morphological differences. Ultimately this has contributed to the development of DNA markers by scientists for sorghum crop improvement.

Overall, large phenotypic differences based on observable trait, between individuals of the same genotypes are driven by differences in the environment they are exposed to, and, by differences in genetic sequence. This suggests that phenotypic diversity between organisms is partially explained by DNA sequence variation yet the environment and genes do interact. Therefore, an individual plant derives its phenotype from the genotype, which is the genetic makeup of the plant hence the environment modifies the expression of the genotype. Based on the phenotypic traits of the inflorescence, grain and glumes, *S. bicolor* subsp. *bicolor* as classified today include five basic races: *bicolor*, *guinea*, *caudatum*, *durra* and *kafir*a described by Harlan and de Wet. (1972).

2.6 Genetic diversity

Genetic diversity refers to the variation of heritable characteristics present among alleles of genes in different individuals of sorghum populations (Geleta *et al.,* 2006) that serves as an important role in evolution by allowing a species to adapt to a new environment (Agrama, 2003).Variation in allele frequencies at many unlinked loci is the preferred way to describe genetic structure and diversity (IPGRI, 1993), and hence to estimate the various forces shaping these loci (Agarwal *et al*. 2008). Several molecular techniques are available for the detection of genetic variability in natural populations.

The ultimate source of genetic diversity is gene mutation; it is the way in which new genes are formed, determined by selection, gene flow, genetic drift, and migration (Hartl & Clark, 1997). Abu Assar, (2005) studying sorghum reported that natural selection chooses the best fit among and within a population and there can be no adaptive evolution without genetic variation. Genetic diversity enables populations of sorghum crop to survive, adapt new circumstances, and evolve to produce new genetic variants; it is regarded as a raw material for evolution of both wild and domesticated sorghums (Hedrick, 2000). Research in population genetics focuses at the determination of the amount of genetic variation in both wild and domesticated sorghum populations so as to face biotic and abiotic stress challenges (Yeh, 1997).

Genetic diversity is vital in selection by enabling the integration of genes or traits for tolerance to both biotic and abiotic stresses (Allard, 1999), in order to address food security problem in arid and semi-arid lands. Genetic diversity can be measured using different approaches within and between sorghum populations as the number of organisms differing from others and the relationships among individuals of their relative frequency at genus, species, population, individual, genome locus and DNA base sequence levels (Gaston, 1998).Therefore, quantitative assessment of genetic diversity is of pivotal importance for determination of the extent of genetic differences between and within sorghum cultivars. Studies on genetic diversity in sorghum genotypes based on biomass, yield index, grain and nutritional quality, phenotypic traits and DNA markers, has provided evidence that sorghum has appreciable genetic variation that has been poorly used in terms of crop improvement programme.

2.7 Genetic distance

Genetic distances are measures of the average genetic divergence between two sequences, species or between populations within a species or taxa. Smaller genetic distances are an indication of a close genetic relationship while large genetic relationships indicate a more distant genetic relationship. According to Carrera *et al.* (1996) genetic distances among progeny confirm their origin and the genetic

relationships between them and their parents. Research on the genetic relatedness and variation present within *sorghum bicolor* and its closely related wild species is important for sorghum population genetics studies. Mutegi *et al.* (2011) studied sorghum genotypes and their wild relatives using SSR markers and applied genetic distances to infer their divergence. Such population genetics studies give information that indicates the rate of adaptive evolution and the extent of response in crop improvement. It is possible to select individuals of sorghum crop but with desired traits for grain or syrup extraction based on their existing genetic distances (Mukuru, 1993). According to Nei, (1975) genetic distance is the extent of gene differences between cultivars, as measured by allele frequencies at a sample of loci. Genetic similarity is the converse of genetic distances, which is the extent of gene similarities among cultivars. Thus, the measure of distance or similarity among cultivars is the covariance of allele frequencies summed for all characters as stated by Smith, 1984. A number of genetic distance measures are used to quantify genetic relationships among cultivars or germplasm accessions. For example, Euclidian distance (Kisua *et al*., 2015) and Rogers coefficient distance (Ng'uni *et al.,* 2012) have been applied in sorghum.

Genetic distance estimates could be determined by molecular markers, which have been found suitable to assess genetic diversity and to identify diverse sources of sorghum genetic resources. In essence each variable of molecular bands such as DNA-based marker bands are treated as a locus so that every locus has two alleles. The banding profiles of each accession can be scored as present (1) or absent (0). Dissimilarity indices concerned with these modalities where presence express presence of information and absence modality indicate absence of information could be described by Jaccard similarity index. Multivariate analyses could be used to reveal the pattern of genetic relationships between and within populations. Closely related populations could be studied through application of a suitable statistical tool such as cluster analysis.

2.8 Polymorphism in sorghum population genetics

Polymorphism is related to genetic variation and adaptation. Two or more clearly different phenotypes existing in the same population of a species may present polymorphism. Differences in forms of the same basic structure, that is, DNA sequences in either their coding regions (exons) or noncoding (introns) regions of genes are called polymorphisms.

A locus is said to be polymorphic when modification of the sequence of a gene at a specific locus in a population occurs. Polymorphism may occur as a result of change in a single nucleotide base or in the number of tandemly repeated sequence in a DNA molecule. Ultimately, DNA markers provide polymorphism information about a genetic resource pool, useful for selection and genetic diversity studies in plant populations. Polymorphic information content (PIC) range from 0 (monomorphic) to 1, that is very discriminative, with many alleles in equal frequencies and hence their usefulness in sorghum genetic analysis according to Agrama *et al.* (2004). Utilization of polymorphism information could assist in identification and selection of parents for breeding highly performing progenies.

2.9 Molecular markers

Molecular genetic markers show differences at the genotype level that can be used to answer and explain questions of genetics. Molecular markers have provided a powerful approach to analyze genetic diversity and evolutionary relationships among and within sorghum populations (Geleta *et al.,* 2006). Molecular markers are useful DNA techniques that complement morphological and physiological characterization of cultivars since they are found in the whole genome according to Altintas *et al.* (2008). Lokko *et al.* (2005) suggested that markers are identifiable DNA sequences found at specific locations of the genome and could be transmitted by the laws of inheritance from one generation to the next. Variation in nucleotide sequence has been exploited to assess the genetic diversity and relationships in sorghum populations. A number of DNA-based molecular markers have been utilized for analysis of genetic divergence of sorghum germplasm including RAPDs (Agrama &

Tuinstra, 2003) and SSRs (Mutegi *et al*., 2008). A DNA marker such as simple sequence repeats (SSRs) have been commonly used to estimate the genetic diversity in sorghum populations.

2.9.1Simple sequence repeats (SSRs)

Simple sequence repeats are highly polymorphic. Polymorphic information content (PIC) of an SSR marker provides an estimate of discriminatory power of that SSR marker by taking into account not only the number of alleles that are detected but also the frequencies of those alleles. Moreover, each SSR is controlled by one locus and each of the loci is controlled by many alleles (2-16 alleles). The DNA sequences flanking SSRs are critical as they allow geneticists to develop locus-specific primers to amplify the microsatellites with polymerase chain reaction (Agrama *et al.,* 2004) Therefore, different alleles can be detected at a locus by the polymerase chain reaction, using conserved DNA sequences flanking the SSR as primers. The primers for PCR are the sequences from the unique flanking regions (Figure1.3).SSR markers are useful at a wide range of scales of analysis as reported by Mutegi *et al*. (2008).

In an evolutionary context they are useful as markers for parentage analysis. SSRs can also be used to address questions concerning the level of relatedness of individuals or groups, and genetic distances. Potential drawback of using SSR markers is that, there are relatively few loci to work with. The most common way to genotype a SSR locus is by a combination of polymerase chain reaction (PCR) and electrophoresis. This method does not give the full sequence information about the alleles. Only the information about (relative) length of the alleles can be obtained. Therefore, the main objective in SSR genotyping is to determine the relative length of the alleles.

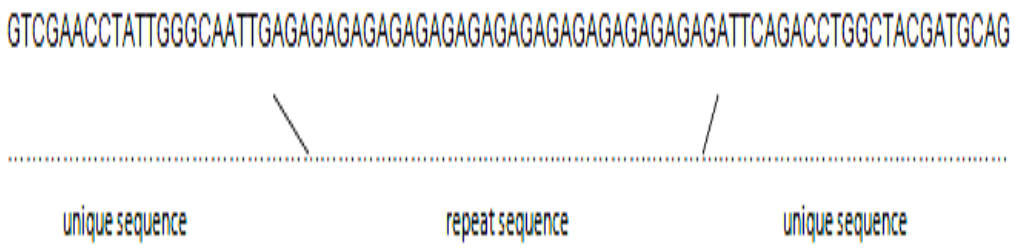

Figure 1.3: A di-nucleotide (GA) n repeat locus of a typical DNA sequence of microsatellite showing a repeat sequence flanked by unique sequences that identify the marker.

CHAPTER THREE

MATERIAL AND METHODS

3.1 Plant material

Seven sorghum populations that consisted sweet and grain sorghum types were assembled for study. The plant material comprised of 3 improved grain (Serena, Seredo and Gadam) sorghum populations from KARLO, 3 sweet (3008, 3018 and 3022) sorghum populations from Chinese origin and a landrace (Kivila) obtained from farmers of Akamba ethnic community in Makueni county. *Sorghum halpense* was included in the study for comparison.

3.2 Study site

The experiment on quantitative genetics was conducted at Kyemole sub-location (1° 53' 18.95" S and 37° 41' 5.07" E) experimental site (Figure 3.2) in Makueni County, during 2013 cropping season. The site located 194 km from Nairobi city, at an elevation of 1,027 meters above sea level is hot semi-arid lowland receiving an average rainfall between 450 and 588 mm annually. The site has sandy soils with a pH 6.5.The average maximum and minimum temperature is 28.67°C and 18°C respectively. Salinity tolerance study and molecular work were conducted at JKUAT in plant physiology and molecular laboratories.

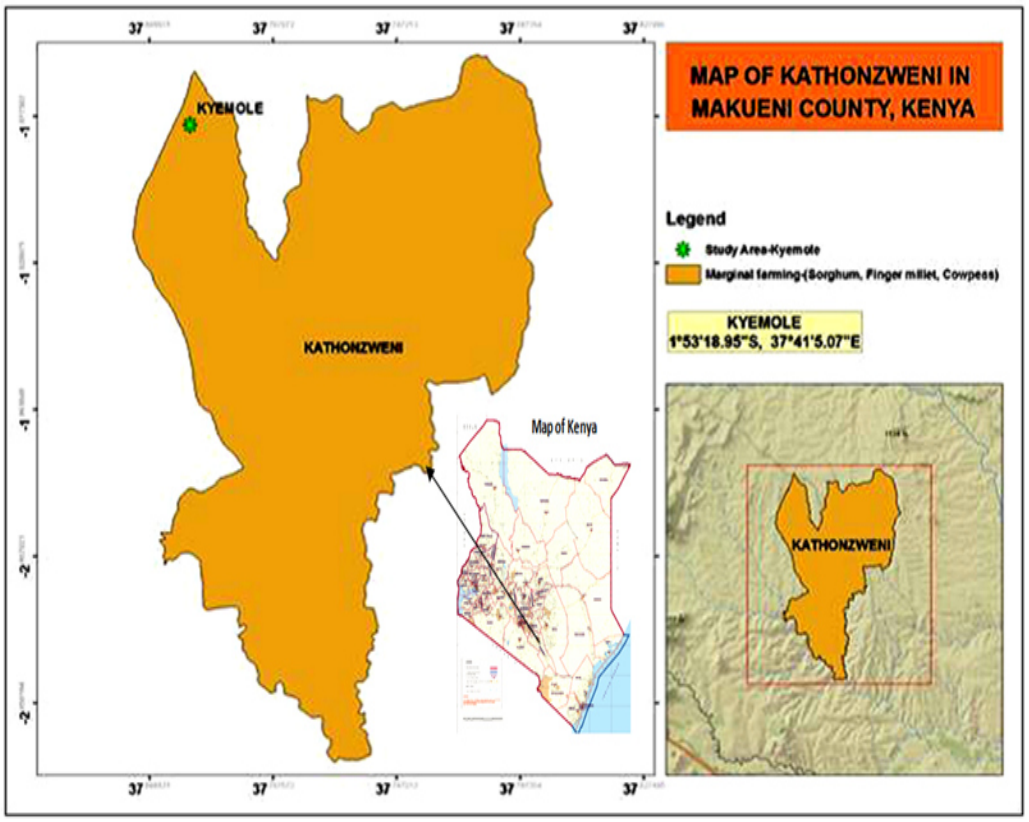

Figure 3.1: Kyemole sorghum study site in Kathonzweni sub-county, Makueni County

3.3 Phenotypic Diversity Using Phenotypic Markers

3.3.1 Experimental design and layout

The study was laid out in a Randomized Complete Design (RCD). Each genotype was grown in three replications for one season in3 x 2 m plots at spacing of 75 x 20cm leaving 1 m path between blocks. Uniform crop management practices including weeding, pest control (stalk borer using Karate) and fertilizer (DAP) was applied.

3.3.2 Scoring for quantitative and qualitative traits

Ten randomly selected and tagged plants (Plate 3) were sampled from each plot to measure 16 quantitative (Table 3.2) and seven qualitative traits (Table 3.3).The data was recorded according to sorghum descriptors (IBPGR, 1993).Leaf area was measured (Plate 1b) and calculated (leaf length x leaf width x 0.747) as recommended by Stickler *et al.* (1961).Stalk diameter was measured with high precision electronic calipers (Plate 1a). Other linear measurements such as in Plate1 (c), (d), (e), and (f) were taken using a standard measuring tape recommended by World Agro Forestry Organization (Table 3.2).Leaf number, number of tillers as well as primary branches per panicle was counted. Grain yield per panicle was obtained by measuring the yield from each panicle using a high precision electronic weighing balance. Munsellcolour film (1990) was used to score for leaf mid rib colour, glume colour and seed colour.

Table 3.1: Quantitative traits studied

Observed trait	Code	Observed trait	Code
Leaf number (count)	LN	Leaf sheath length (cm)	L SL
Leaf length (cm)	LL	Plant height (cm)	PH
Leaf width (cm)	LW	Peduncle exertion length (cm)	LPE
Leaf area (cm^2)	LA	Panicle length (cm)	PL
Stalk diameter (cm)	SD	Panicle width (cm)	PW
length of internode (cm)	IL	Number of internodes (count)	NI
Number of tillers (count)	NT	Grain yield per panicle (g)	GYP
Primary branches per panicle (count)	NPB	Hundred seed weight (g)	HSW

Plate 3.1: Phenotyping for quantitative traits in the field at Kathonzweni study site. The plates shows the researcher taking; a) stalk diameter using electronic calipers b) leaf width c) panicle length d) Peduncle exertion length e) panicle width f) internode length.

3.3.3 Phenotypic Description

Analysis of variance (ANOVA) was computed using GenStat Discovery, v12.1. Program. Means for each trait were separated by the least significant difference (LSD) at ($p<0.05$). Phenotypic correlation coefficients were calculated to estimate the level of relationship among the quantitative traits. The data were standardized prior to cluster analysis. Application of an agglomerative, hierarchical cluster analysis (HCA) classification with UPGMA technique effected to obtain a dendrogram. Euclidean distance measure is a commonly used statistic for estimating genetic distance between populations by phenotypic traits. Let d_{ij} be dissimilarity between units i and j where X_{ik}, and X_{jk} be values of variable k for units i and j while

K represents the number of variables. Euclidean distance which is commonly used for continuous data can be described as:

$$d_{ij} = \sqrt{\sum_1^K (x_{ik} - x_{jk})^2} \qquad (1)$$

The diversity index ($\acute{H}$) of Shannon and Weaver (1949) was used as a measure of phenotypic diversity for each trait. The diversity index was computed for each trait over all populations. This index is expressed as:

$$\acute{H} = -\sum_{i\ 1}^{n} Pi \log_2 Pi \qquad \text{……………………………………………………….(2)}$$

Where n is the number of classes of traits and Pi is the proportion of populations in the i^{th} class of a trait. Each H' value was normalized by dividing it by its maximum value (log_2), which ensured that all H' values were in the range of 0 to 1. Σ is the sum of the calculations. The various qualitative traits across the populations were gathered to estimate the diversity of traits between the studied populations. The average diversity index (H´) for n traits was computed using $H' = \sum H'/n$ statistic. The statistical analysis was done using the GenStat Discovery, v12.1, statistical software.

3.4 Determination of Salt Tolerance potential in selected sorghum genotypes

3.4.1 Germination test

Seven genotypes of *Sorghum bicolor* (Table. 3.4) were evaluated for their salinity tolerance measured at three salinity levels [5ds/m, 10ds/m, and 15ds/m against a control (0.22dS/m)]. Sorghum seeds were disinfected with 3 % sodium hypochlorite solution for 2 minutes. Petri dishes with a diameter of 90mm lined with12.5 mm diameter What man filter paper were supplied with 10ml of each treatment solution and the control. Forty (40) uniform seeds of each sorghum population were placed on each petri dish and the petri dishes were arranged in a RCD with three replications.

The petri dishes were covered with a polyethylene sheet to avoid the loss of moisture through evaporation. The germination test was done in germinator with temperature 25^{o}C, 70% relative humidity; and a light condition was 12 hour regime. Treatment application continued every other day and germination count was started after 48 hours of sowing and continued until the 14th day. The seeds were considered to have germinated when their radicle length was more than 2 mm. In order to calculate germination indices, germinated seeds were counted daily. Germination Percentage (GP) and Germination Rate (GR) were calculated; Germination percentage is an estimate of the viability of a population of seeds expressed as:

$$GP = Ni / N \times 100; \quad (3)$$

Where GP is germination percentage; Ni is the number of germinated seed till i^{th} day) and N is the total number of seeds

While germination rate (GR) which is the average number of days needed for plumule or radical emergence was calculated using MAGUIRE's equation (Maguire 1962) as:

$$GR\ (M) = n_2/t_2 + n_4/t_4 + n_6/t_6 \ldots + n_{14}/t_{14}; \quad (4)$$

Where n_2, n_4, n_6 .., n_{14} represent the number of germinated seeds at times t_2, t_4, t_6 …, t_{14} (in days).

3.4.2 Seedling growth test under hydroponics

Sorghum seeds were disinfected with 3 percent sodium hypochlorite solution for 2 minutes and germination conducted in JKUAT Plant Physiology laboratory under controlled conditions at 27^{o}C. The seeds were placed in 90mm petri dishes lined with moistened 12.5 mm diameter Whatman filter paper with each petri dish containing 40 uniform seeds of a particular cultivar and placed in a germination chamber. The petri dishes were covered with polythene sheet to avoid the loss of moisture through evaporation. After three days, the seedlings chosen for uniformity

of both the plumule and radicle length of 0.5 cm were transferred to 4 liters containers containing 3 liters of Shive and Robbins nutrient solution culture. A hydroponics system in the laboratory was set for the growth of seedlings with provision for aeration to provide enough gaseous exchange for plant growth.

The different levels of salinity were imposed in Shive and Robbins nutrient hydroponic solution using NaCl corresponding to electrical conductivities 5, 10, and 15 dS/m against a control, i.e. 0.22 dS/m. Treatment application was done immediately after transfer of the seedlings and the nutrient solution was replaced after every two days. The pH was adjusted to 5.5 (by use of HCl for lowering pH and NaOH for pH increment) throughout the growing season in order to maintain a suitable root growth environment. Supplemental lighting of 4750 lux for 12 hours a day was provided by use of fluorescent tubes throughout the growing season. A temperature of 28^0C was maintained automatically for the day and at 25^0C for the night. The seedlings were allowed to grow in this condition for seven days. . The results of non- destructive daily shoot and root length measurements during the rapid phase of leaf elongation (seven consecutive days) starting on the second day after transferring seedlings into the nutrient solution was obtained. Fresh weight and dry weight of all the selected seedlings (10) were measured using an electronic precision weighting balance. Dry weight was obtained after drying the seedlings in the oven 65^0C for 24hrs and cooled before weighing. Root-shoot ratio was calculated using measurements of root and shoot dry weight.

3.4.3 Data analysis

3.4.3.1 Multivariate Analyses for Seedling Traits under NaCl Treatment

The data of seedling traits were transformed to arc sine $\sqrt{\frac{x}{100}}$ transformation prior to subjecting it to data analysis. Data transformation is regarded as a remedial measure for variance heterogeneity where the variance and the mean, are functionally related. This technique converts the original data into a new scale that result in a new data set that is aimed at satisfying homogeneity of variance condition. To determine the index

of tolerance (TI), the growth character is considered to be fitness because the parameters measured are correlated with fitness. Therefore, the following formula was used to calculate tolerance index (TI):

$$\mathrm{TI} = \frac{mean\ of\ trait\ in\ contaminated\ environment}{mean\ value\ under\ control} \qquad (5)$$

Where, ω represents the mean value of a given trait in a contaminated environment, and α the mean value of the trait under control.

Tolerance index was represented graphically showing the variations based on salinity levels. Better still, the salt tolerance values for all the traits studied were subjected to analyses of variance (ANOVA) for interaction between population and salinity. The trait correlation coefficients were calculated to determine the level of relationship among the studied growth and seedling traits. Multivariate analysis of variance was conducted to show the genetic variability based on salinity tolerance for salinity at E.C 15 dS/m level. The means of each of the character studied were standardized prior to subjecting it to the principal component analysis (PCA) and hierarchical cluster analysis.

Manhattan and Euclidean belong to the same distance family and are linked. Manhattan distances are commonly used for continuous data. The continuous data obtained was used for calculation of Manhattan distance using the dissimilarity matrix.

This formula was used:

$$d(i,j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \ldots + |x_{ip} - x_{jp}| \qquad (6)$$

Where $i = (x_{i1}, x_{i2}, \ldots, x_{ip})$ and $j = (x_{j1}, x_{j2}, \ldots, x_{jp})$ are two p-dimensional data objects. Dendrogram was produced using DARwin v.6.1 software. All other computations were made using GenStat v.12.

3.5 Heritability and Variance Components Analysis of Traits under NaCl Treatment

The seven sorghum populations namely; Kivila, Serena, 3018, Seredo, 3008, Gadam and 3022 were investigated under four NaCl E.C levels control, 5dS/m, 10dS/m, and 15dS/m. Seven traits associated with fitness were considered. The mean values obtained in the determination of genetic diversity under salt treatment were used to determine heritability levels.

3.5.1 Estimation of magnitude of variation

The mean values obtained in the determination of genetic diversity under different salt treatment were subjected to analysis of variance (ANOVA) for each NaCl concentration level. The phenotypic and genotypic variances were computed using the formula shown below:

$$\sigma^2{}_g = \frac{(MS_g - MS_e)}{r} \quad (7)$$

$$\sigma^2{}_P = \sigma^2{}_g + \sigma^2{}_e \quad (8)$$

$$\sigma^2{}_e = MS_e \quad (9)$$

Where, σ^2_g = genotypic variance, σ^2_p = phenotypic variance, σ^2_e = environmental variance, MS_g= Mean square due to genotypes/populations, MS_e= Error mean square, and r = number of replications.

The coefficient of variations at phenotypic and genotypic level was estimated according to Falconer (1981). This is expressed as:

$$PCV = \left[\sigma_p / \bar{x}\right] x100 \quad (10)$$

$$GCV = \left[\sigma_g / \bar{x}\right] x100 \quad (11)$$

$$ECV = \left[\sigma_e / \bar{x}\right] x100 \quad (12)$$

Where, σ_p = phenotypic standard deviation ($\sigma_g + \sigma_e$), σ_g = genotypic standard deviation, σ_e = Environmental standard deviation, $\bar{X}$ = Grand mean for the characteristic x; PCV, GCV, and ECV = phenotypic, genotypic and environmental coefficient of variation, respectively.

3.5.2 Estimate of heritability

Broad sense heritability (h^2) for all characters based on salinity levels (EC5, EC 10, EC 15) was calculated as suggested by Allard (1960) using the formula:

$$h^2 = \left[\sigma^2_g / \sigma^2_p\right] x100 \quad (13)$$

$$\sigma^2_p = \sigma^2_g + \sigma^2_e \quad (14)$$

Where, σ^2_g = genotypic variance, σ^2_p = Phenotypic variance, σ^2_e = error variance.

3.6 Molecular Diversity among Sorghum populations

3.6.1 DNA extraction

DNA was isolated from shoots and roots of etiolated, 4 day old sorghum seedlings grown in an incubator (Heraus CO_2-auto-zero incubator at 30°C). Five sorghum seedlings from each population were randomly sampled and used for DNA extraction. Individually sampled fresh shoots and roots tissue was placed in eppendorf tubes, frozen in liquid nitrogen, freeze dried and ground into powder. The isolation of DNA was done using cetyltrimethyl ammonium bromide (CTAB) according to the modified method from Mace *et al*., (2004). CTAB buffer [3% (w/v) CTAB, 1.4m NaCl, 20mM hydrodymethylaminomethane hydrochloride (Tris-

HCl),20mM ethylenediaminetetraacetic acid (EDTA) pH 8.0, 0.17% (v/v) β-mercaptoethanol] corresponding to 450µL per sample was dispensed into a glass bottle and incubated in a water bath maintained at 65 ^{o}C. Extraction Buffer was added to the samples and ground to fine powder. The macerated product was transferred to fresh microfuge tubes and incubated for 15 minutes at 65^{o}C with occasional mixing. The DNA was then purified by adding chloroform: isoamyalcohol (24:1) to each sample and mixed thoroughly by inversion. The tubes were centrifuged at 12000 xg for 10 minutes at 24^{o}C and the upper portion transferred into fresh tubes (about 400µl). About an equal volume of Iso-propanol was added to the aliquot to precipitate the DNA and inverted once to mix and the tubes centrifuged after 20-30 minutes at 12000 xg for 15 minutes. The supernatant was decanted into a fresh tube and the pellet air dried for 30 minutes. The DNA pellet was washed with 70% ethanol air dried at room temperature over night then dissolved in 100µL 1X Tris-EDTA. A volume of 3 *µL*RNAse (10 mg/ml) was added to DNA sample before incubating in an oven at 37 ^{O}C for one hour. 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. Later the aqueous layer was transferred into fresh tubes. DNA purification was done by adding 315µL absolute ethanol and 1/10 volume of 3M sodium acetate solution (pH 5.2) to each sample and then placed in -20^{O}C for 5 minutes. The tubes were then centrifuged at 12000 xg for 5 minutes and the supernatant decanted. A volume of 200µL of 70% ethanol was added and centrifuged at 3500xg for 5 minutes. This is for washing the DNA pellet. DNA pellet was air-dried for one hour. The pellet was then re-suspended in 100µL low salt TE (10mM Tris, 1mM EDTA [pH 8]) buffer and stored at 4^{o}C.

3.6.2 Gel Electrophoresis

The quality and quantity of the isolated DNA was confirmed by loading 2 µL of DNA and 0.5 µLof loading dye (bromophenol blue) in 0.8% agarose gel (Sigma, UK) in 1x TBE buffer (89.2 mMTris, 89.0 mM Boric acid, 1.25 mM EDTA pH 8.0) and electrophoresed at 100 volts/hour. Water was placed in the first and last wells in the gel to act as a negative control. The DNA fragments emitted a luminous glow

under UV light and were photographed using a video capture system (Flowgen IS 1000). The concentration and quality was further determined spectrophotometrically by using Nanodrop ND8000 (Thermo Scientific NanoDropCHEM-CF-8 (1) at 260 nm and 280 nm.

3.6.3 DNA Quantification

The concentrations were used to determine the dilution factors for each sample in order to achieve a concentration of 50 ng/µL. DNA samples were diluted to 10 ng μL^{-1} and used to perform polymerase chain reaction. Ten Simple Sequence Repeats (SSRs) markers (Table 3.6) were used for the amplification of the DNA samples.

Table 3.2: Simple Sequence Repeat locus, repeat motif type, primer sequences and Linkage group (LG) location on sorghum linkage map.

Locus	Motif	Forward primer (5'-3')	Reverse primer (5'-3')	LG
Xtxp 34	(CT)29	*TGGTTCGTATCCTTCTCTACAG*	*CAT ATA CCT CCT CGT CGC TC*	SB1-03
Xtxp141	(GA)23	*TGTATGGCCTAGCTTATCT*	*CAACAAGCCAACCTAAA*	SB1-10
Xtxp21	(AG)18	*GAGCTGCCATAGATTTGGTCG*	*ACC TCG TCC CAC CTT TGT TG*	SB1-04
Xcup60	(CGGT)4	*GTATGCATGGATGCCTGATG*	*GCGAGGGTATGTAGCTCGAC*	SB1-01
Xgap84	(AG)14	*CGC TCT CGG GAT GAA TGA*	*TAA CGG ACC ACT AAC AAA TGA TT*	SB1-02
Xcup07	(CAA)8	*CTAGAGGATTGCTGGAAGCG*	*CTGCTCTGCTTGTCGTTGAG*	SB1-07
Xcup37	(AG)9	*CCCAGCCTTCCTCCTGATAC*	*GTACCGACTCCAATCCAACG*	SB1-09
Xcup02	(GCA)6	*GACGCAGTTTGCTCCTATC*	*GTCCAACCAACCCACGTATC*	SB1-06
Xtxp14	(GA)15	*GTAATAGTCATGACCGAGG*	*TAA TAG ACG AGT GAA AGC CC*	SB1-05
Xtxp273	(TTG)20	*GTACCCATTTAAATTGTTTGCAGTAG*	*CAGAGGAGGAGGAAGAGAAGG*	SB1-08

3.6.4 PCR and SSR primer optimization

The SSR primer screening for amplification, optimization of PCR conditions, SSR series, repeat motif and detection of polymorphism lead to the selection of ten SSR primer pairs for genetic diversity analyses of both grain and sweet sorghum germplasm. To carry out amplification, a 5 µL PCR mix consisting of 10 ng of DNA, 1X reaction buffer, 1.5 mM $MgCl_2$, 0.1 mM dNTPs (dATP, dCTP, dGTP and dTTP),

0.2 U *Taq* polymerase, 0.2 pmols of forward and reverse primers and 2.23 µL of sterile water was amplified in and amplified in GeneAmp PCR system 9700 (Applied Biosystems, Foster City, Ca, USA).

The reaction mixture was denatured at 95°C for 5 minutes initially, then subjected to four more cycles of 95°C (45seconds), 45 seconds annealing at 55 ° C reduced by 1 ° C every cycle, and 60 seconds extension at 72 ° C. This was followed by 30 cycles of 45 seconds at 95 ° C, 45 seconds at 55 ° C, 60 s at 72 ° C and extension phase of 5 minutes at 72 ° C prior to cooling at 4°C.

A sample lacking genomic DNA was also loaded in each PCR run to serve as a negative control. Confirmed amplified PCR products were separated on 3 % agarose gel. Allelic data for each locus was recorded as fragment size in comparison with a standard 50 bp DNA ladder and also as binary data coded as 1 or 0 for the presence or absence for each allele.

3.6.5 Analysis of allelic variation for microsatellite loci and allele frequency in studied populations

Bands from polymorphic SSR markers were scored as present (1) or absent (0) for each population by manual inspection. Determination of total number of alleles and polymorphism information was estimated from the genotyping data, considering that each SSR primer pair represents a unique locus. The degree of polymorphism of each SSR locus was calculated based on the polymorphic information content (PIC), as described in Botstein *et al*. (1980):

$$PIC = 1 - \sum_{i=1}^{k} p_i^2 - \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} 2p_i^2 p_j^2 \qquad (16)$$

Where k is number of alleles, pi and pj are the frequency of the i^{th} and the j^{th} alleles respectively. Allelic size data for each SSR locus was used to determine the percentage of polymorphic loci.

Nei's gene diversity (h) was used to summarize the data of SSR alleles. The value provides an estimate of the discriminatory power of a locus by taking into account not only the number of alleles, but also their relative frequencies. The genetic diversity is equivalent to the expected heterozygosity (H_e).

It is computed as:

$$h_j = 1 - \Sigma p_i^2 \qquad (17)$$

Where, p^{ji} denote the frequency of the i^{th} allele at the j^{th} locus. For a single locus, H ranges from 0 (monomorphic) to 1 which is described as very highly discriminative with many alleles in equal frequencies. So the average diversity over all *r* loci for each population is computed as:

$$H=1-\sum_{j=1}^{i=r}\sum_{i=1}^{i=n} p_{ji}^2/r \qquad (18)$$

Observed heterozygosity (H_o) and expected heterozygosity (H_e) was calculated according to Nei (1987)using Power Marker version 3.25 (Liu, 2003).Test of Hardy Weinberg equilibrium was done by computing expected genotypic frequencies under random mating using the algorithm used by Falconer and Mackay (1996),and performing chi-square (X^2) tests. Probability values were used to determine the populations under HW equilibrium. Genetic similarity (GS) indices were calculated using the Jaccard's coefficient for all possible pairwise comparisons. The Jaccard's coefficient disregards the conjoint absence of alleles in the pairwise comparison, reducing the risk of over estimating similarity.

This was calculated as:

$$d_{ij} = \frac{b + c}{a + (b + c)} \quad (19)$$

Where *dij* is the dissimilarity between units *i* and *j*;X_i, X_j represent variable values for units *i* and *j*; *a* represent the number of variables where X_i= presence and X_j= presence, represent the number of variables where X_i= presence and X_j= absence and *c* represents the number of variables where X_i= absence and X_j= presence. Jaccard's coefficients were calculated using DARwin 6.0.12 dissimilarity analysis software and used to cluster genotypes according to similarity. Cluster analysis based on Jaccard similarity coefficient with Unweighted Neighbor Joining method as proposed by Gascuel (1997). Let i and j be the elements combined in a new element s, ci and cj the unit numbers of these elements, and k be another element. Taking d(s,k) as the weighted average of dissimilarities between k and elements i and j; then unweighted solution is given as:

$$d(s,k) = \frac{c_i}{c_i + c_j} d(i,k) + \frac{c_j}{c_i + c_j} d(j,k) \quad (20)$$

To ascertain the statistical strength of genetic relationships identified through this analysis, bootstrapping (≥ 40) of the data was performed.

Principal co-ordinate analysis was performed based on Jaccard's distance matrix using DARwin 6.0.12 dissimilarity analysis software. Principal coordinates analysis (PCoA) is a member of the factorial analysis family working on distance matrices. Let *dij* be the observed distance between *i* and *j*, and δ*ij* be the distance between the projections of *i* and *j* on the first axis.

PCoA extracts the first axis using the formula:

$$\sum_{i,j}(d_{ij} - \delta_{ij})^2 \qquad (21)$$

Then it extracts a second axis, orthogonal to the previous one minimizing the squared differences. Eigenvectors and eigen values produce the solutions of the matrix W of scalar products between elements that is defined from the d_{ij} according to the Torgers on formula which is given as:

$$W_{ij} = -(d_{ij}^2 - d_{i.}^2 - d_{.j}^2 + d^2)/2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(22)$$

3.7 Population equilibrium within seven sorghum populations based on five loci

Test of Hardy Weinberg equilibrium (HWE) was computed using expected genotypic frequencies under random mating according to Falconer and Mackay (1996). Chi-square (X^2) tests was performed. Probability values were used to determine the populations under HW equilibrium.

CHAPTER FOUR

RESULTS

4.1 Phenotypic Diversity using Phenotypic Markers

4.1.1 Analysis of Variance for Eight Sorghum Populations

The eight populations of sorghum differed significantly differences ($p \leq 0.01$) in all the phenotypic traits studied (Table 4.1a). The analysis of variance (ANOVA) recorded significantly high values in population 3022 in comparison with all other populations for grain yield per panicle, internode length, leaf number, leaf sheath length, plant height and stalk diameter.

Considerably high mean values for panicle width were observed in two populations (3008 and 3018) than in 3022. However, low mean for number of primary branches per panicle and considerably high mean value for panicle length were observed in Kivila (landrace). Although Gadam is a grain sorghum genotype, it recorded significantly low mean value for grain yield per panicle. Additionally, *Sorghum halpense* had the lowest values of grain yield per panicle. Sweet sorghums were phenotypically diverse from the grain sorghums as was observed.

Table 4.1a: Coefficient variance percentage, least significant difference, means and mean squares for 16 quantitative traits

Population	Trait															
	GYP	HSW	IL	LA	LL	LN	LPE	LSL	LW	NI	NPB	NT	PH	PL	PW	SD
3018	36.2	1.6	10.1	403.4	60.4	10.6	48.4	20.8	8.9	8.9	39.6	0.1	181.7	20.4	10.2	14.9
Seredo	30.3	2.8	10.7	411.5	68.9	9.3	34.8	20.9	8.1	7.4	58.1	0.6	120.9	24.4	9.3	15.3
Serena	28.8	1.9	12.4	360.8	65.6	9.1	33.8	20.2	7.4	6.6	50.9	0.3	134.3	24.2	7.1	13.1
Kivila	38.4	1.8	12.4	380.5	71.2	8.9	44.7	19.7	6.9	8.9	25.2	1.3	166.8	23.9	10.5	16.6
3022	40.2	1.9	11.8	435.9	69.4	11.1	47.4	21.9	8.4	9.3	45.2	0.2	190.2	21.9	9.7	17.1
Gadam	28.5	1.4	9.1	328.2	66.5	8.2	33.4	16.4	6.7	6.7	34.3	0.7	104.8	20.6	7.9	15.2
3008	26.5	1.9	12.6	447.9	70.0	10.8	46.6	21.3	8.3	8.9	44.2	0.3	196.2	23.1	10.2	16.6
S.halpense	4.3	1.4	18.4	108.5	54.1	8.7	70.4	15.2	2.7	6.0	41.3	0.0	194.8	34.7	25.0	3.7
Mean squares	3.8**	5.8**	23.6**	3.6**	101.1**	3.7**	4.4**	17.8**	11.5**	5.1**	300.5**	0.5**	3.9**	61.5**	96.9**	57.3**
LSD	0.12	0.04	0.19	0.12	0.18	0.17	0.15	0.16	0.14	0.14	0.3	0.15	0.21	0.17	0.16	0.15
CV %	0.20	0.10	0.60	0.60	0.20	0.10	0.70	0.10	0.40	0.80	0.70	6	14	0.10	0.30	0.5

Trait abbreviations: PH-plant height, GYP-grain yield per panicle,, LL-leaf length , IL-internode length, LA-leaf area, LN-leaf number, LPE-peduncle exertion length, LSL-leaf sheath length, LW-leaf width, NI-number of internodes, HSW-hundred seed weight , NPB-primary branches per panicle, NT-number of tillers, PL-panicle length, PW-panicle width, SD-stem diameter;**= significant at p≤0.01

4.1.2 Bivariate statistics

Significantly positive association was observed between grain yield per panicle ($r=0.87$, $p<0.01$) and leaf area (Table 4.1b). However, significant negative relationship between grain yield per panicle and panicle length ($r=-87$, $p<0.01$) as well as panicle width ($r=-0.82$, $p<0.01$) was observed. Leaf area was positively and significantly correlated with leaf length ($r=0.81$, $p<0.01$) and width ($r=0.97$, $p<0.01$). Leaf number revealed positive correlation with plant height ($r=0.64$, $p<0.01$) and negative correlation with number of tillers ($r=-0.41$, $p<0.01$) as well. High correlation coefficient value was recorded between stalk diameter($r=0.75$, $p<0.01$) and leaf sheath length. Significant and negative correlation coefficient scores were observed between stalk diameter and internode length ($r=-0.81$, $p<0.01$). Such a negative relationship was noted between panicle exertion ($r=-0.72$, $p<0.01$) and a number of other traits. Nevertheless a positive relationship ($r=0.83$, $p<0.01$) between panicle exertion and internode length was noted. Strong correlation coefficient values for grain yield per panicle with leaf area ($r=0.87$, $p<0.01$) and stalk diameter ($r=0.91$, $p<0.01$) was observed.

Table 4.1b: Correlation coefficient matrix for 16 phenotypic traits.

No.	Trait	GYP	HSW	IL	LA	LL	LN	LPE	LSL	LW	NI	NPB	NT	PH	PL	PW	SD
1	GYP	-	NS	*													
2	HSW	0.334	-														
3	IL	-0.782*	-0.269	-													
4	LA	0.871**	0.534	-0.751*	-												
5	LL	0.730*	0.555	-0.583	0.811*	-											
6	LN	0.443	0.207	-0.153	0.618	0.230	-										
7	LPE	-0.625	-0.465	0.841**	-0.666	-0.696	0.128	-									
8	LSL	0.757*	0.622	-0.455	0.890**	0.630	0.789*	-0.405	-								
9	LW	0.878**	0.401	-0.808*	0.977**	0.670	0.648	-0.651	0.887**	-							
10	NI	0.762*	0.197	-0.381	0.760*	0.555	0.807*	-0.070	0.789*	0.738*	-						
11	NPB	-0.134	0.675	0.005	0.162	-0.022	0.245	-0.221	0.350	0.217	-0.234	-					
12	NT	0.461	0.234	-0.383	0.272	0.649	-0.427	-0.471	0.037	0.152	0.184	-0.502	-				
13	PH	-0.145	-0.272	0.579	-0.093	-0.266	0.649	0.798*	0.192	-0.110	-0.481	-0.183	-.445	-			
14	PL	-0.870**	-0.169	0.941**	-0.867**	-0.656	-0.384	0.743*	-0.607	-0.890**	-0.605	0.093	-.308	0.317	-		
15	PW	-0.829*	-0.386	0.892**	-0.862**	-0.764**	- 0.218	0.923**	-0.639	-0.867**	-0.415	-0.119	-.390	0.511	0.913**	-	
16	SD	0.919**	0.413	-0.839**	0.956**	0.884**	0.441	-0.762*	0.754*	0.907**	0.738*	-0.070	0.498	-.217	-0.918**	-0.890**	-

Trait abbreviations : PH-plant height, GYP-grain yield per panicle,, LL-leaf length , IL-internode length, LA-leaf area, LN-leaf number, LPE-peduncle exertion length, LSL-leaf sheath length, LW-leaf width, NI-number of internodes, HSW-hundred seed weight , NPB-primary branches per panicle, NT-number of tillers, PL-panicle length, PW-panicle width, SD-stem diameter.*=significant at $p \leq 0.05$, **= significant at $p \leq 0.01$

4.1.3 Principal component analysis

Principal component analysis of the quantitative traits revealed that the first four principal components cumulatively accounted for 97.19 % of the total variation (Table 4.1c). The first PC accounted for 57.61 % of the total variation and was positively associated with internode length, leaf area, leaf number, leaf sheath length, panicle width and plant height. The first and the second PCs cumulatively scored 77.29 % of the total variation. The second, third and fourth principal components explained 19.68 %, 12.76 % and 7.14 % of total variation, in that order.

Table 4.1c: Proportion, cumulative variances and Eigen-vectors on four principal components (PC) based on16 quantitative traits in eight studied sorghum populations

	Eigenvectors			
Trait	**PC 1**	**PC 2**	**PC 3**	**PC 4**
Grain yield per panicle	0.301	0.032	0.164	-0.013
Hundred seed weight	0.171	-0.014	-0.468	0.502
Internode length	-0.288	0.183	-0.031	0.324
Leaf area	0.320	0.102	-0.044	0.020
Leaf length	0.280	0.080	0.051	0.050
Leaf number	0.157	0.497	-0.053	-0.112
Leaf peduncle exertion	-0.257	0.322	0.152	0.101
Leaf sheath length	0.274	0.274	-0.164	0.131
Leaf width	0.314	0.127	-0.072	-0.153
Number of internodes	0.239	0.345	0.264	0.135
Number of primary panicle	0.033	0.075	-0.690	-0.032
Number of tillers	0.136	-0.329	0.317	0.513
Plant height	0.091	0.523	0.193	0.121
Panicle length	-0.305	0.056	-0.136	0.334
Panicle width	-0.318	0.158	0.040	0.174
Stalk diameter	0.324	-0.012	0.114	0.045
Eigen Value	9.220	3.154	2.042	1.141
% Variation	57.61	19.68	12.76	7.14
% Cumulative variance	57.61	77.29	90.05	97.19

4.1.4 A two-dimension plot

The biplot display pattern (Figure 4.1) is explained by the similarities in number of internode, leaf sheath length, leaf width and leaf area of the studied populations. Grain sorghums clustered together as well as sweet sorghums. However, *S. halpense* and Gadam showed divergence from these groups. The distance between the panicle width and internode length traits was narrow as well as leaf width with leaf area. Plant height trait and number of tillers were projected by the vectors in opposite directions on the biplot display. It was observed that the number of primary branches and hundred seed weight were projected close to the core. Many of other traits were projected distant away from the core depending on whether they were closely related or had a wide phenotypic divergence.

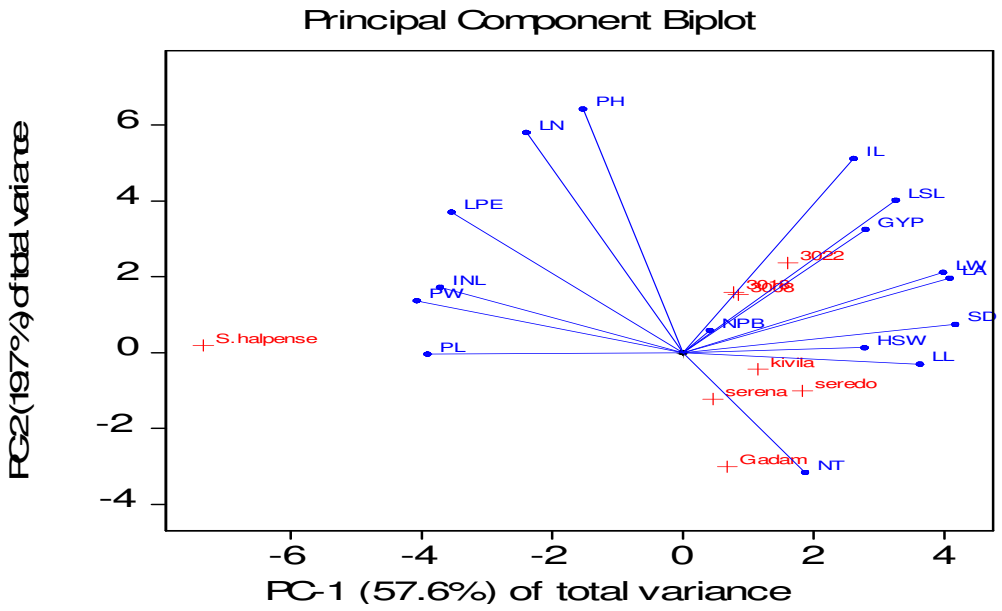

Figure 4.1: Principal component biplot accounting for genetic variation in terms of sorghum phenotypic traits (Blue) and populations (Red). Trait abbreviations : PH-plant height, GYP-grain yield per panicle,, LL-leaf length , IL-internode length, LA-leaf area, LN-leaf number, LPE-peduncle exertion length, LSL-leaf sheath length, LW-leaf width, NI-number of internodes, HSW-hundred seed weight , NPB-primary branches per panicle, NT-number of tillers, PL-panicle length, PW-panicle width, SD-stem diameter.*=significant at p≤0.05, **= significant at p≤0.01.

4.1.5 Genetic distance and cluster analysis

The dedrogram (Figure 4.2) obtained from the cluster analysis grouped the eight sorghum populations into three main clusters. Cluster 1 comprised two sub-clusters. Su-cluster (1a) consisted of three sweet sorghum populations namely; 3018, 3022 and 3008.two populations 3018 and 3022 separated at a minimum distance coefficient 2.36. Sub-cluster (1b) contains kivila which is a landrace separated at a genetic distance 4.53. Grain sorghums were grouped into cluster 2 forming two sub-clusters. Sub-cluster (2a) gathered Serena and Seredo at a Euclidian coefficient 2.56. Gadam a dwarf variety was grouped into sub-cluster (2b) at a genetic distance 4.00 (Table 4.1d). Cluster 3 consisted *S.halpense* which is a wild sorghum population at the maximum genetic distance 8.80.Cluster means (Table 4.1e) recorded substantially high values in cluster 1 for the traits studied. However, cluster 3 produced the highest mean in peduncle exertion trait.

Table 4.1d: Genetic distance based on pair-wise combination of phenotypic traits studied

No.	Populations	1	2	3	4	5	6	7	8
	Gadam	1							
2	Kivila	4.53	1						
3	3008	4.53	4.28	1					
4	Serena	4.00	4.53	4.53	1				
5	3018	4.53	4.28	2.80	4.53	1			
6	Seredo	4.00	4.53	4.53	2.56	4.53	1		
7	3022	4.53	4.28	2.80	4.53	2.36	4.53	1	
8	*S.halpense*	8.80	8.80	8.80	8.80	8.80	8.80	8.80	1

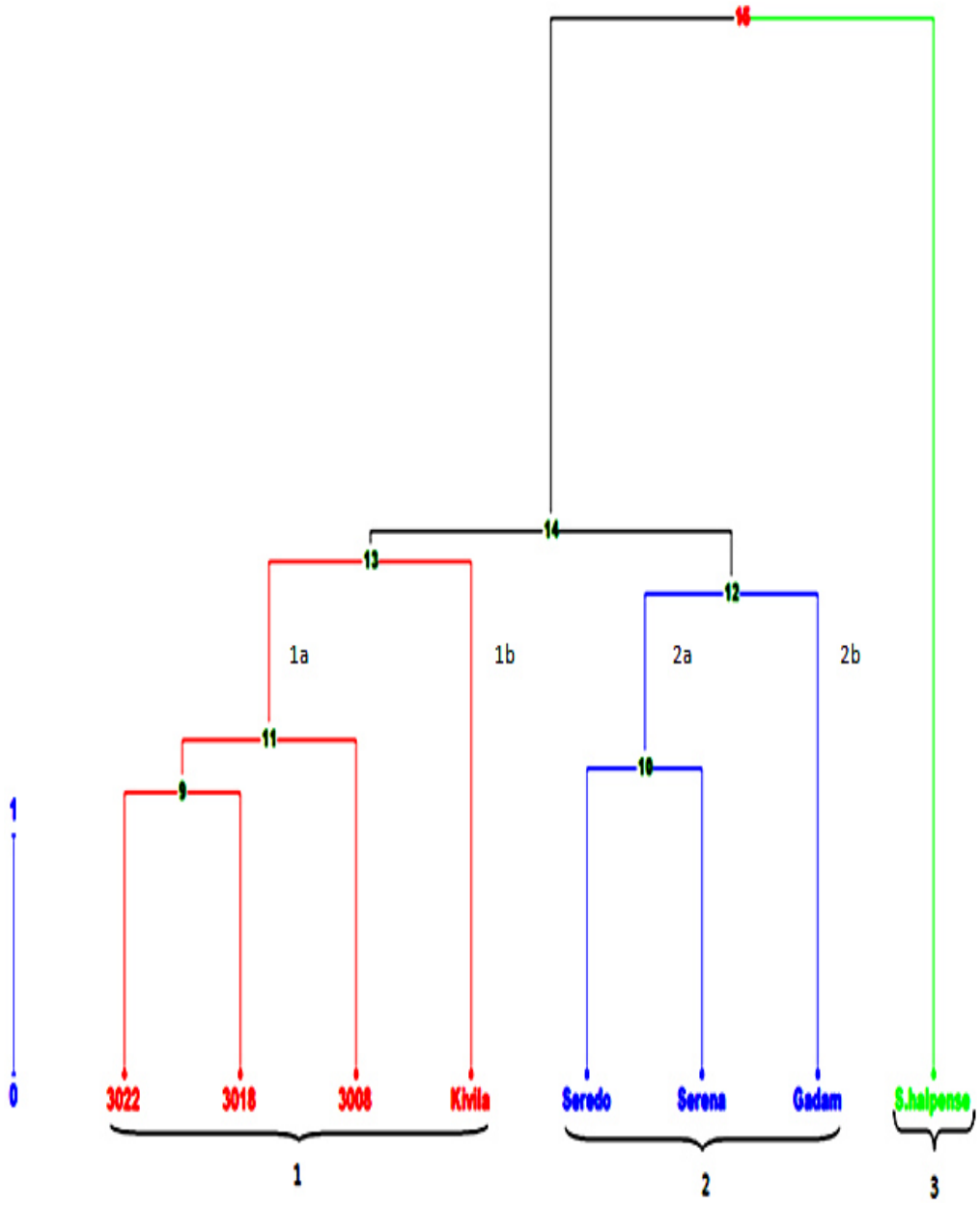

Figure 4.2: Genetic distance among eight sorghum populations revealed by UPGMA cluster analysis. The scale to the left represents Euclidian distance coefficient and the values between branches are the bootstrap values generated.

Table 4.1e: Summarized cluster means for the 16 phenotypic traits.

Traits	Cluster Means			
	1	2	3	Mean
Grain yield per panicle (g)	35.27	29.17	22.4	28.95
Hundred seed weight (g)	1.81	2.02	1.14	1.66
length of internode (cm)	11.79	10.65	18.34	13.59
Leaf area (cm^2)	416.93	366.83	107.5	297.09
Leaf length (cm)	67.71	66.94	54.09	62.91
Leaf number (count)	10.40	8.87	13	10.76
Length of peduncle exertion (cm)	46.80	33.98	70.38	50.39
Leaf sheath length (cm)	20.93	19.08	15.15	18.39
Leaf width (cm)	8.13	7.33	2.69	6.05
Number of internodes (count)	9.08	6.93	6	7.34
Primary branches per panicle (count)	38.65	47.87	41.3	42.61
Number of tillers (count)	0.48	0.50	0	0.33
Plant height (cm)	183.75	119.90	194.7	166.12
Panicle length (cm)	22.35	23.06	34.68	26.70
Panicle width (cm)	10.23	8.11	25.01	14.45
Stalk diameter (cm)	16.29	14.56	3.73	11.53

4.1.6 Qualitative traits

Shannon - Weaver diversity index computation for qualitative traits (Plate 4) studied revealed a considerably high level of variation. The phenotypic diversity index ($\bar{H}$) for traits studied recorded a mean value of 0.70 and a maximum 0.95. Panicle compactness and shape trait scored 0.73, while glume colour recorded 0.76 (Table 4.1f). However values >0.80 were observed in grain shape, grain colour and plant color.

Table 4.1f: Diversity index ($\bar{H}$) values for qualitative explaining the genetic diversity of populations studied

Qualitative traits	Diversity index (H)
Grain covering	0.70
Panicle compactness and shape	0.73
Grain shape	0.87
Glume colour	0.76
Grain colour	0.80
Leaf mid rib color	0.95
Plant color	0.86
Total diversity index	0.70

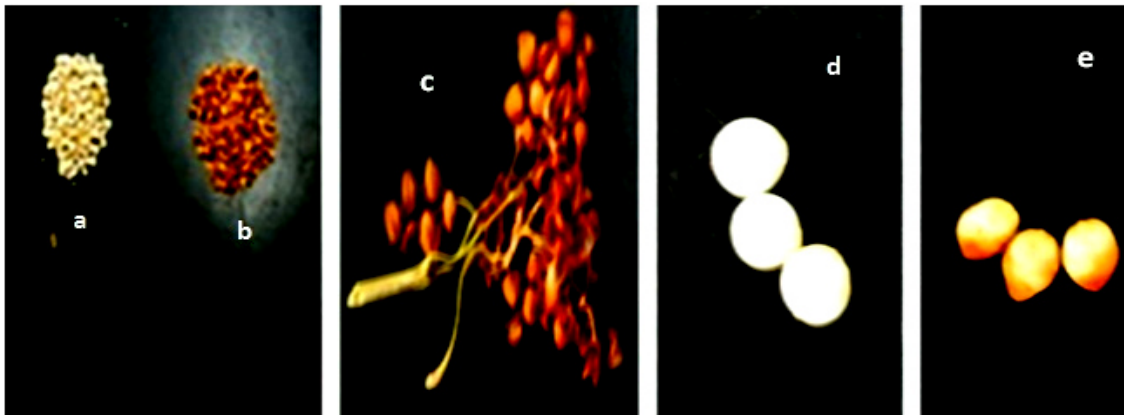

Plate 4.1: Genetic variability in sorghum populations based on qualitative traits. a) white grains b) brown grains c) brown glumes d) white-round shaped grains f) yellow-brown oval shaped grains

4.2 Germination and Seedling Trait under hydroponics study

4.2.1 Effect of population and salinity tolerance

Analyses of variance (ANOVA) showed that mean squares for all the traits studied differed highly significantly ($P<0.001$) for the seven populations based on three salinity levels ($P<0.001$) except root shoot ratio ($P<0.01$) and population x salinity except for dry weight ($P< 0.05$) for interaction (Table 4.2a). Coefficient of variation ranged from 3.0 % to 19.2%. Leaf necrosis was observed in Serena (Plate 4.1a) under 15 EC treatments whereas chlorosis which is a symptom of nutrient imbalance was observed in 3008 (Plate 4.1b). Poor root formation (Plate 4.2) was observed in these two populations (Serena and 3008) as well. Calculated tolerance index showed a decline trend for mean values for growth and germination traits with respect to increase in salinity levels.

Table 4.2a: Mean squares and significances from the analysis of variance for salt tolerance traits in seven sorghum populations under study.

		Traits						
Sources of variation	DF	Dry Weight	Germination percentage	Germination rate	Root length	Root shoot ratio	Shoot length	Wet weight
Replications	2	0.0000019^{ns}	5.356^{ns}	1.0364^{ns}	9.3989^{ns}	0.0022677^{ns}	2.565^{ns}	0.0010039^{ns}
Populations(Pop)	6	0.0003478***	399.454***	19.5305***	45.1639***	0.0218190***	98.357***	0.0018018***
Salinity (Sal)	3	0.0017276***	3997.491***	29.2653***	63.8169***	0.0010949**	672.073***	0.0085492***
Pop x Sal	18	0.0000213*	117.254***	1.4597***	2.3960***	0.0020347***	29.179***	0.0007764***
Error	54	0.00001	6.082	0.2901	0.5828	0.00203	3.214	0.00024
CV %		4.9	4.2	9.8	5.7	6.2	3.0	19.2

**indicates differences significant at* P≤0.05, ***indicates differences significant at* P≤0.01, ****indicates differences significant at* P≤0.001, ^{ns} *denotes differences which are non-significant*, Pop = population, Sal = salinity

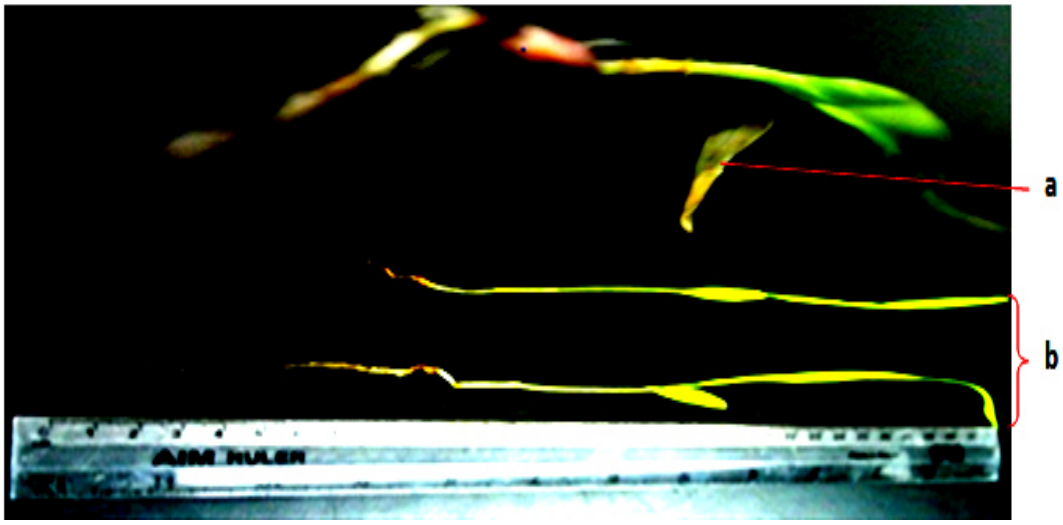

Plate 4.2: Leaf toxicity symptoms in seedlings under15 dS/m NaCl treatment. a) Necrotic leaf of Serena population b) chlorotic leaves of 3008 population

.

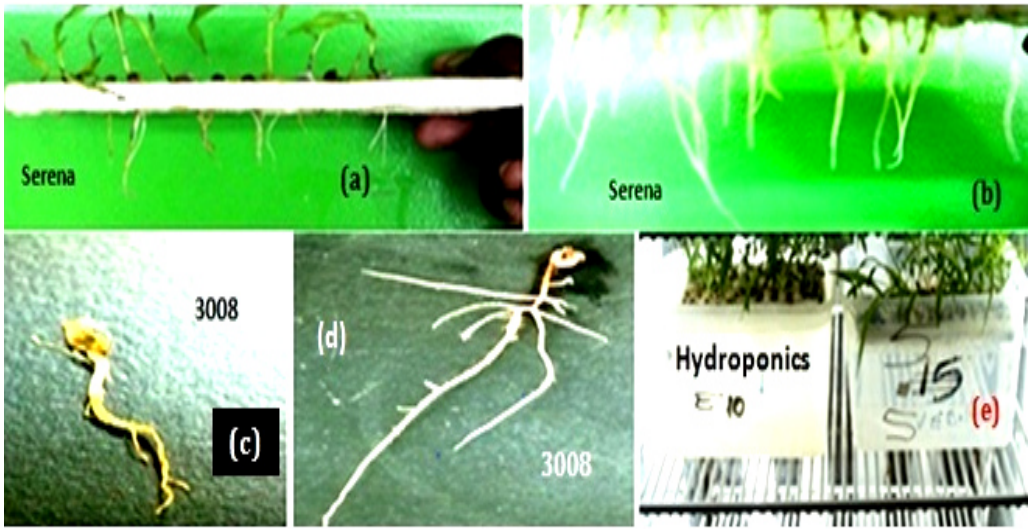

Plate 4.3: Effect of salt treatment on root structure of Serena and 3008 populations. a) Altered root structure of Serena under NaCl 15 dS/m treatment b) nature of root structure of Serena at control c) altered root architecture of 3008 under 15 dS/m NaCl.

4.2.2 Salinity tolerance index

Salinity tolerance indices revealed significant differences in reduction of dry weight. Lowest values were observed at 15 dS/m salinity level where Serena recorded a TI = 19.14% and Gadam considerably higher values for dry weight (84.89%) and wet weight (84.42%) traits. Germination percentage ranged from 36.33% % to 84.21% with Kivila and 3022 scored low values of 40.46% and 36.33% respectively. A higher value 84.21% was observed in Seredo which is grain sorghum.

Germination rate varied from 40.91% to 79.09 %. Shoot length scored a higher (85.74 %) value than root length (79.33 %). Low scores of wet weight were observed in Serena (26.28%) but 3018 scored a high 86.48 % tolerance index. Summarized data (Table 4.2b) for tolerance index (TI) shows a decrease in TI (5 dS/m>10 dS/m >15 dS/m) with increase in salinity stress levels. Based on calculated total percentage variation for TI on individual populations high genetic variability was noted where Seredo, Gadam,3022, kivila,3018, 3008 and Serena recorded 83.31%, 81.99%, 82.48%,79.77%, 78.39%, 78.10% and 74.09% respectively (Table 4.2b).

Table 4.2b: Mean tolerance index (TI) values at different selection pressures

Population	S_2= 5 dS/m	S_3= 10 dS/m	S_4= 15 dS/m	Mean TI (%)
Kivila	0.9847	0.7862	0.62227	79.77
3018	0.9111	0.76255	0.6779	78.39
Seredo	0.9594	0.813	0.7268	83.31
3008	0.9087	0.7456	0.6888	78.10
Gadam	0.9461	0.7977	0.7158	81.99
Serena	0.9462	0.7466	0.5300	74.09
3022	0.8965	0.8604	0.7175	82.48
Average	0.9361	0.7874	0.6684	-

4.2.3 Bivariate Analysis

The significance ($p<0.01$) pattern of the genetic correlations (Table 4.2c), related to germination percentage, germination rate, root length, root- shoot ratio and shoot length with the dry weight at r=0.7153**, r=0.5628**, r=0.7219**, r=0.3318** and r=0.6067** values respectively, was largely the same. Similar pattern of genetic correlations were recorded for germination rate (r= 0.3122**) with root length. Significant positive correlation (r= 0.2424*) was scored between shoot length and root length (<0.05). This has a survival value of utilization of available moisture in the soil by a large number of seeds during early development stage. Notably, root-shoot ratio(r=0.8189**) and wet weight ($r=0.4841^{**}$) showed strong correlation with root length. Nor was there observed positive significant correlation of root- shoot ratio trait with germination rate (-0.0177^{ns}) at probability level of 5 %. Strong negative correlation was observed between shoot length ($r= -0.3478^{**}$) and root-shoot ratio. Nonetheless positive correlation coefficient observed between this trait with germination percentage ($r =0.0296^{ns}$) was non significant.

Table.4.2c: Correlation coefficient matrix for seven studied salt tolerance traits.

Traits	DWT	GP	GR	RL	RSR	SL	WWT
DWT	1	0.7153**	0.5628**	0.7219**	0.3318**	0.6067**	0.6174^{**}
GP		1	0.6934**	0.4318*	0.0296^{ns}	0.6597**	0.5093**
GR			1	0.3122**	-0.0177^{ns}	0.5235^{**}	0.3536^{**}
RL				1	0.8189**	0.2424*	0.4841^{**}
RSR					1	-0.3478^{**}	0.2605^{*}
SL						1	0.3373^{**}
WWT							1

Trait abbreviation: dry weight (DWT), germination percentage (GP), germination rate (GR), shoot length (SL) root length (RL), root-shoot ratio (RSR) and wet weight (WWT),*=significant ≤0.05, **=significant≤ 0.01, ^{ns}=Non-significant.

4.2.4 Principal component analysis

The first three principal components, with eigenvalues greater than unity (eigenvalue >1), explained 79.98 % of the total variation among the seven populations. The first and second principal component accounted for 45.10% and 20.07% of the total variation, in that order with a cumulative variation of 65.17 %. PC1 axis was mainly contributed by wet weight, root length and root-shoot ratio (Table 4.2d). Germination percentage and tolerance index had high coefficients and were major contributors of PC2 variation. However dry weight showed the highest variation in PC3 axis.

Table 4.2d: Principal Components for eight quantitative traits in sorghum genotypes.

	PC1	**PC2**	**PC3**
Eigenvalue	3.157	1.405	1.037
Proportion δ^2	45.10	20.07	14.81
Cumulative δ^2	45.10	65.17	79.98
		Eigenvector	
Dry Weight	0.32102	0.29917	0.61738
Germination %	0.08489	0.65187	-0.00570
Germination rate	-0.06527	0.43004	0.05884
Root Length	0.50983	0.08766	0.10756
Root Shoot Ratio	0.52688	0.02192	-0.12883
Shoot Length	-0.41678	0.08238	0.51933
Tolerance Index	-0.09174	0.47230	-0.56210
Wet Weight	0.40722	-0.25021	-0.03916

4.2.5 A two-dimension plot

The populations were scattered within the four quadrants of the two dimension plot as well as the traits which were projected by vectors. Kivila (landrace) and 3018 (sweet sorghum) were projected far away from the core in comparison with the rest. Germination rate showed a close relationship with tolerance index. Similar pattern was shown by root length and root shoot ratio. Germination percentage, wet weight and shoot length were projected away from the center core and in different quadrants. The biplot (Figure 4.5) statistic separated 3018 away from other sweet sorghums (3008 and 3022) including Kivila which is landrace. It was noted that Kivila clustered with the sweet sorghums. This statistic showed a similar pattern among the grain sorghums Serena diverged away from Gadam and Seredo.

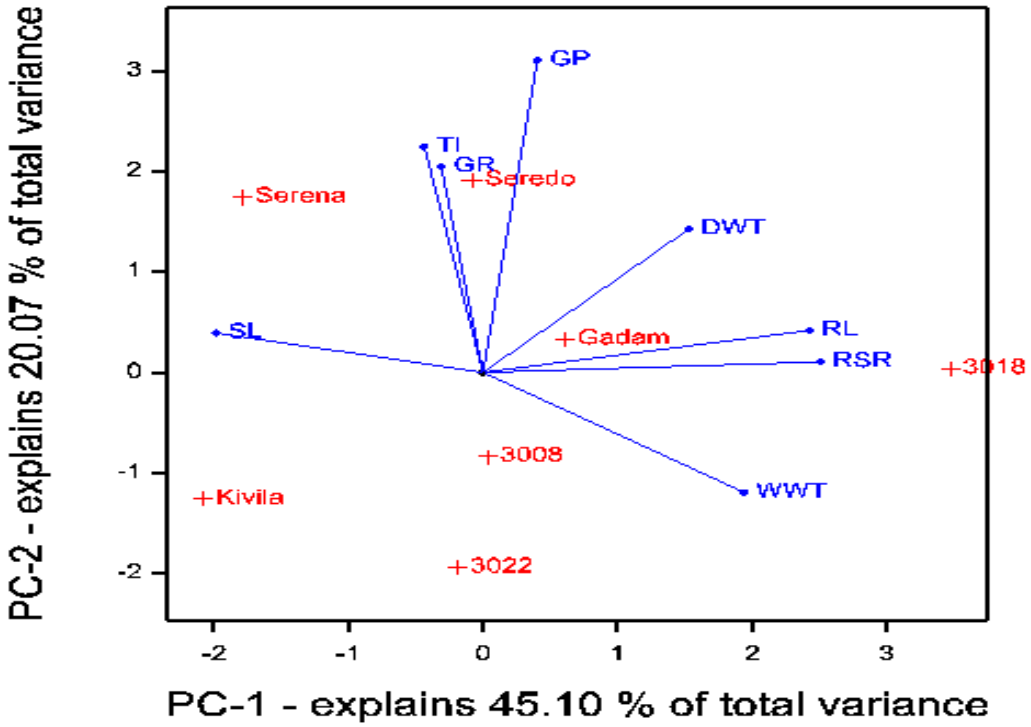

Figure 4.3: A two-dimension plot of sorghum traits (blue) and populations (red). The traits used are dry weight (DWT), germination percentage (GP), germination rate (GR), shoot length (SL) root length (RL), root-shoot ratio (RSR), tolerance index (TI)and wet weight (WWT).

4.2.6 Scree plot

The scree plot pattern (Figure 4.6) displayed the seven dimension PCA analysis in a linear fashion, thereby presenting 100% variation. The graphical statistic delineated the PC axes from 45.10% to 0.29%. PC 1 recorded the highest variation at 45.10% followed by PC2 with 20.07%. An elbow was detected at Eigenvalue 1.037 scoring 14.81% of the total variation. The remaining four, PC4, PC5, PC6 and PC7 received 13.24%, 4.44%, 2.04%, and 0.29% of the total variation respectively.

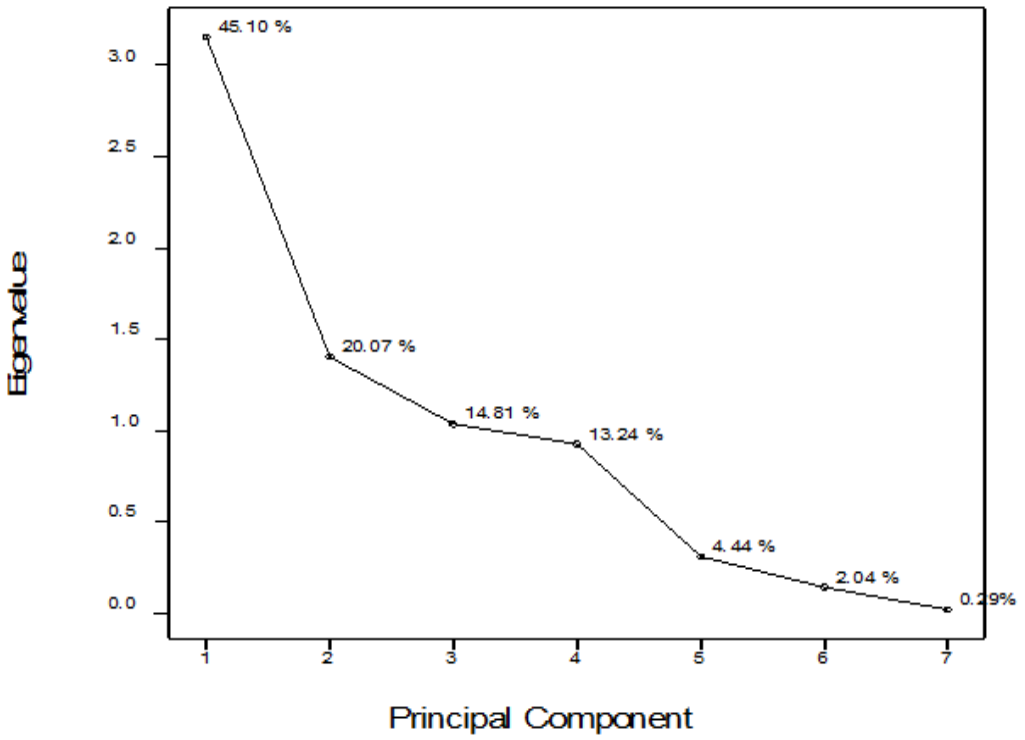

Figure 4.4: Scree plot of Principal Component analysis illustrating proportion of variation for each component sorted in decreasing in a decreasing linear fashion.

4.2.7 Cluster analysis

Cluster analysis (Figure 4.7) gathered sorghum populations in three main clusters. Serodo and Serena were classified in cluster1 at a genetic distance of 2.88. Cluster 2 comprised Gadam a short variety with a tall sweet sorghum (3018) genotype with 1.83 coefficient distance. Cluster 3 consisted of three genotypes 3022, Kivila and 3008. Kivila and 3022 separated at a Manhattan distance 1.14 (Table 4.2e) whereas3022 was placed at a genetic distance 2.25 from the other two members of this cluster.

Table 4.2e: Manhattan distance matrix based on Darwin's grouping of the seven sorghum populations studied.

Population	1	2	3	4	5	6	7
Kivila	1						
3018	3.74	1					
Seredo	3.74	3.17	1				
3008	2.25	3.74	3.74	1			
Gadam	3.74	1.83	3.17	3.74	1		
Serena	3.74	3.17	2.88	3.74	3.17	1	
3022	1.14	3.74	3.74	2.25	3.74	3.74	1

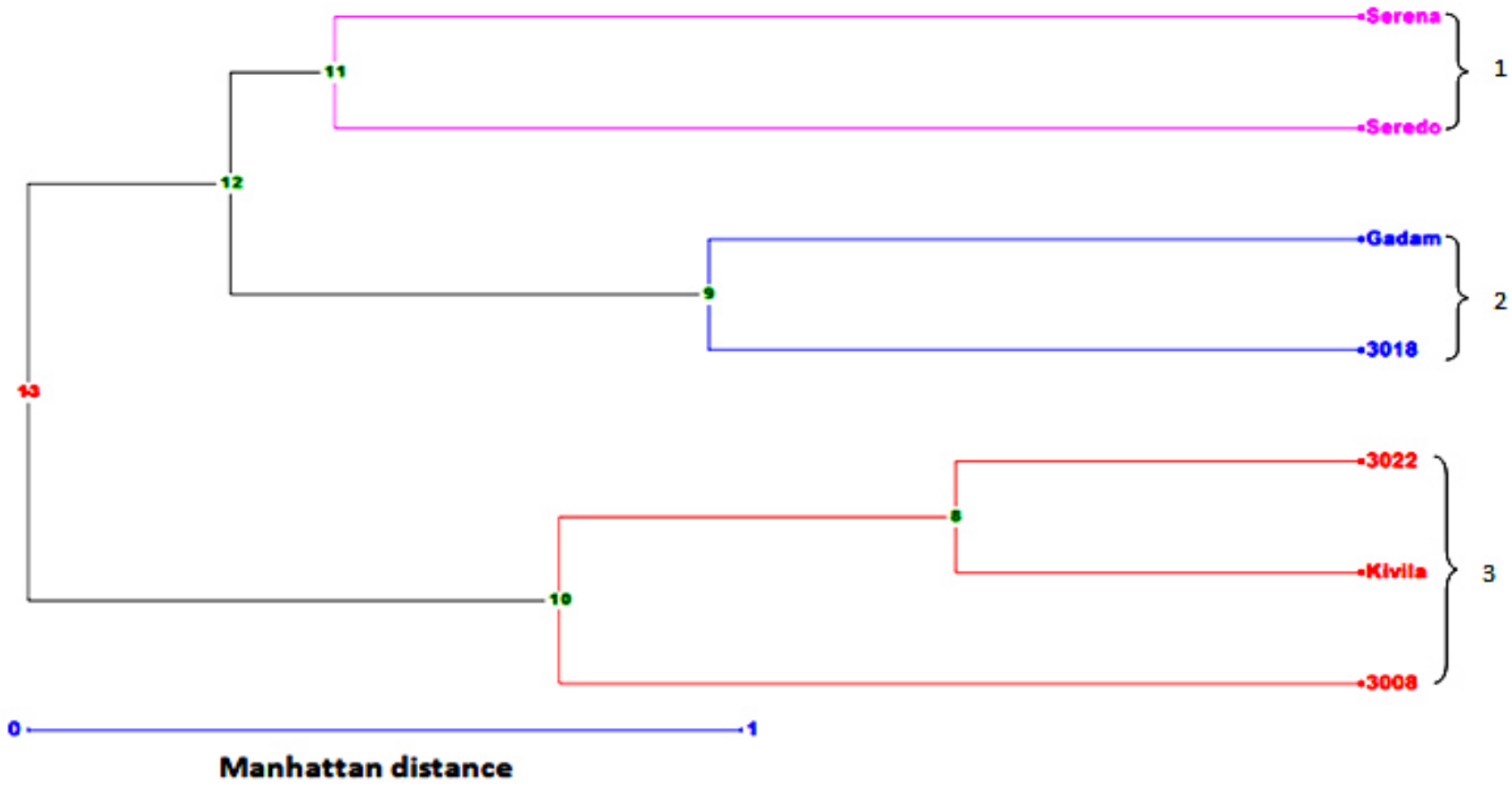

Figure 4.5: Manhattan distance and UPGMA clustering of sorghum populations under NaCl concentration treatments

4.3 Heritability Values and Population Divergence

4.3.1 Estimates of variance components

Substantially high genotypic and phenotypic variances were observed in shoot length, root length, germination percentage and germination rate at 5dS/m treatment (Table 4.3). Coefficient of variance values as well as scores for heritability was higher at control than that of respective selection pressures (5dS/m, 10dS/m and 15dS/m). Genotypic coefficient of variability (GCV) ranged from 2.32 % for shoot length to 21.65 % for germination rate whilst phenotypic coefficient of variability (PCV) ranged from 3.58 % for shoot length to 28.63 % for germination rate at maximum salt concentration (15dS/m) . Germination rate (28.67%), exhibited high phenotypic (PCV) coefficient of variance, whereas that of root- shoot ratio (13.7%) and wet weight (13.7%) was moderate. Shoots length (3.58 %), germination percentage (6.6%), dry weight (9.7%) and root length (9.6%) scored considerably low PCV values. Genotypic (GCV) coefficient of variance scores were high for germination rate (28.64%) and low for traits including root- shoot ratio (9.3%), wet weight (8.5%), shoots length (2.3 %), germination percentage (5.2%), dry weight (7.0%) and root length (6.7%). Generally, the values of PCV were higher than their corresponding GCV values.

4.3.2 Heritability in broad sense (h^2)

Heritability in broad sense values ranged from 38.57 % for wet weight to 62.58 % germination percentage at 15dS/m. Moderate heritability (40 % to 80%) for dry weight (51.76%), germination percentage (62.6%), germination rate (57.2%), root length (48.6%), root-shoot ratio (45.8%) and shoot length (42.2%) was observed at this selection pressure. Heritability of wet weight (38.57%) was low. However, high heritability (>80 %) in dry weight (85.8%) and germination percentage (81.1%) was recorded at control.

Table 4.3: Estimates of phenotypic (δ2p) and genetic (δ2g) variance, phenotypic (PCV) and genotypic (GCV) coefficients of variation and heritability (h2) of seven traits of seven sorghum populations.

Components	Traits	Control	5dS/m	10dS/m	15dS/m
$\delta^2 p$	Dry weight	2.8884×10^{-4}	3.5334×10^{-5}	3.4697×10^{-5}	4.6696×10^{-5}
	Germination percentage	23.34	16.0936	13.461	15.5725
	Germination rate	4.789	3.667	3.2529	3.2303
	Root length	3.509	2.4532	1.902	2.0080
	Root- shoot ratio	4.2183×10^{-3}	9.4031×10^{-4}	1.0039×10^{-3}	9.989×10^{-4}
	shoot length	3.2045	5.219	4.1385	5.2798
	Wet weight	1.0916×10^{-4}	1.0908×10^{-4}	9.62177×10^{-5}	9.4607×10^{-5}
$\delta^2 g$	Dry weight	2.4788×10^{-4}	2.5115×10^{-5}	2.417×10^{-5}	2.424×10^{-5}
	Germination percentage	18.93	12.6802	10.3812	9.7453
	Germination rate	3.814	2.7147	2.3271	1.8461
	Root length	2.679	1.7263	1.2455	0.9757
	Root- shoot ratio	3.1127×10^{-3}	6.223×10^{-4}	5.7243×10^{-4}	4.577×10^{-4}
	shoot length	2.2297	0.2947	2.229	2.2296
	Wet weight	6.9067×10^{-5}	5.6737×10^{-5}	4.7175×10^{-5}	3.649×10^{-5}
PCV (%)	Dry weight	36.24	11.09	8.3410	9.6764
	Germination percentage	11.39	4.253	6.1559	6.6211
	Germination rate	34.86	28.98	11.3211	28.633
	Root length	12.70	10.06	9.3500	9.607
	Root- shoot ratio	0.2809	13.07	2.148	13.6701
	shoot length	2.7435	1.098	2.536	3.5763
	Wet weight	14.2985	12.47	12.083	13.6859
GCV (%)	Dry weight	33.56	6.84	6.9616	6.9616
	Germination percentage	10.26	13.44	5.4060	5.2378
	Germination rate	26.84	24.93	24.3027	21.646
	Root length	11.09	8.4362	7.566	6.6968
	Root- shoot ratio	0.2413	10.63	1.00391	9.253
	shoot length	2.2885	0.8253	2.1105	2.324
	Wet weight	11.3735	8.99	8.4607	8.4999
h^2	Dry weight	85.82	71.07	69.66	51.76
	Germination percentage	81.11	78.79	77.12	62.58
	Germination rate	79.64	74.03	71.54	57.15
	Root length	76.35	70.37	65.47	48.59
	Root- shoot ratio	73.79	66.18	57.02	45.82
	shoot length	69.58	56.46	53.86	42.23
	Wet weight	63.27	52.01	49.03	38.57

$\delta^2 p$= Phenotypic variance; $\delta^2 g$ = genotypic variance; PCV = phenotypic coefficient of variation; GCV = genotypic coefficient of variation and h^2 = broad sense heritability.

4.4 Allelic variation for microsatellite loci and allele frequency in studied populations

Five (50%) SSR markers were monomorphic and the rest were polymorphic (Plate 4.4). This study detected a total of 10 alleles for the SSRs genotyped over seven genotypes (Table 4.4a). The frequency of the major alleles at different SSR loci ranged from 0.50 (Xtxp 021) to 0.64 (Xtxp 014, Xcup02, Xgap08). Loci Xtxp34 scored 0.57. The mean PIC was 0.36 and this ranged from 0.35 to 0.37. Recorded expected heterozygosity (H_e) ranged from 0.46 to 0.50 with an average of 0.47 whereas heterozygosity (H_O) values observed ranged from 0.43 to 0.89 scoring a mean of average of 0.66. However expected heterozygosity (H_e) was lower than heterozygosity (H_O) values in Xcup02.

Table 4.4a: Molecular diversity indicators revealed by five SSR primer combinations indicating, locus, repeat motif, repeat length, major allele frequency, genotype number, allele number, gene diversity (He), heterozygosity (Ho), and PIC for seven studied populations

Locus	Repeat motif	Repeat length	Major allele Frequency	Genotype Number	Allele number	Gene diversity	Heterozygosity	PIC
Xtxp34	(CT)29	2	0.5714	3	2	0.4898	0.5714	0.3698
Xtxp 014	(GA)15	2	0.6429	2	2	0.4592	0.7143	0.3538
Xtxp 021	(AG)18	2	0.5000	1	2	0.5000	0.8942	0.3750
Xcup02	(GCA)6	3	0.6429	3	2	0.4592	0.4286	0.3538
Xgap084	(AG)14	2	0.6429	2	2	0.4592	0.7143	0.3538
Mean		_	0.6000	2	2	0.4735	0.6646	0.3612

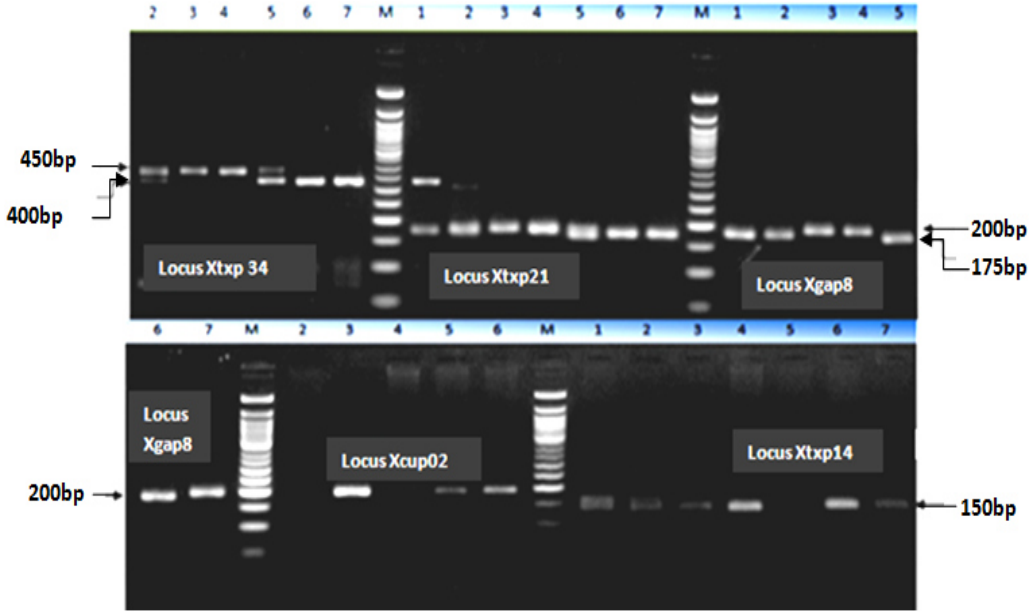

Plate 4.4: PCR products for the 5 optimized markers. M denote 50bp molecular size ladder. Lane (1, 2, 3,4,5,6 and 7) represents DNA for serena, seredo, kivila, gadam, 3022, 3008 and 3018 respectively.

4.4.1 Population equilibrium within seven sorghum populations

Hardy-Weinberg equilibrium computation recognized four loci Xtxp34 (p-value=0.6592), Xtxp 014 (p-value=0.1416), Xcup02 (p-value=0.2703), Xgap084 (p-value=0.1416) conforming to this test. However, locus Xtxp 021 recorded a highly significant p value (0.047) showing non - adherence to Hardy-Weinberg equilibrium (Table 4.4b).

Table 4.4b: X2test of HW equilibrium within studied sorghum populations

Marker	χ^2	P
Xtxp34	0.1944	0.6592
Xtxp 014	2.1605	0.1416
Xtxp 021	3.9375	0.0472
Xcup02	1.2153	0.2703
Xgap084	2.1605	0.1416

4.4.2 Cluster analysis and principal coordinate analysis (PCoA)

Cluster analysis based on Jaccard similarity coefficient with Un-weighted Neighbor Joining method was performed on the 5 SSR markers for seven populations. The cluster analysis discriminated the populations and recognized three main clusters (Figure 4.9a). Cluster I had two sub-clusters (Ia and Ib) and contained four populations. The sub-cluster Ia contained 3018 and 3008 and sub-cluster Ib contained Gadam (grain sorghum) and Kivila a landrace from Akamba community at a genetic distance 0.333 (Table 4.4c). Cluster II contained Serena and Seredo and cluster III contained population 3022. The first two principal coordinates of PCoA explained 78.10 % of the total variation and the populations scattered in the four quadrants (Figure 4.9b). Moreover, the first and second coordinates explained 45.28% and 32.82% of the total variation in that order.

Table 4.4c: Jaccard similarity distance for seven sorghum populations using five SSR primers.

Population	1	2	3	4	5	6	7
Serena	1						
Seredo	0.286	1					
Kivila	0.812	0.699	1				
Gadam	0.755	0.642	0.333	1			
3022	0.556	0.444	0.827	0.770	1		
3008	0.685	0.572	0.682	0.625	0.700	1	
3018	0.711	0.598	0.708	0.651	0.726	0.286	1

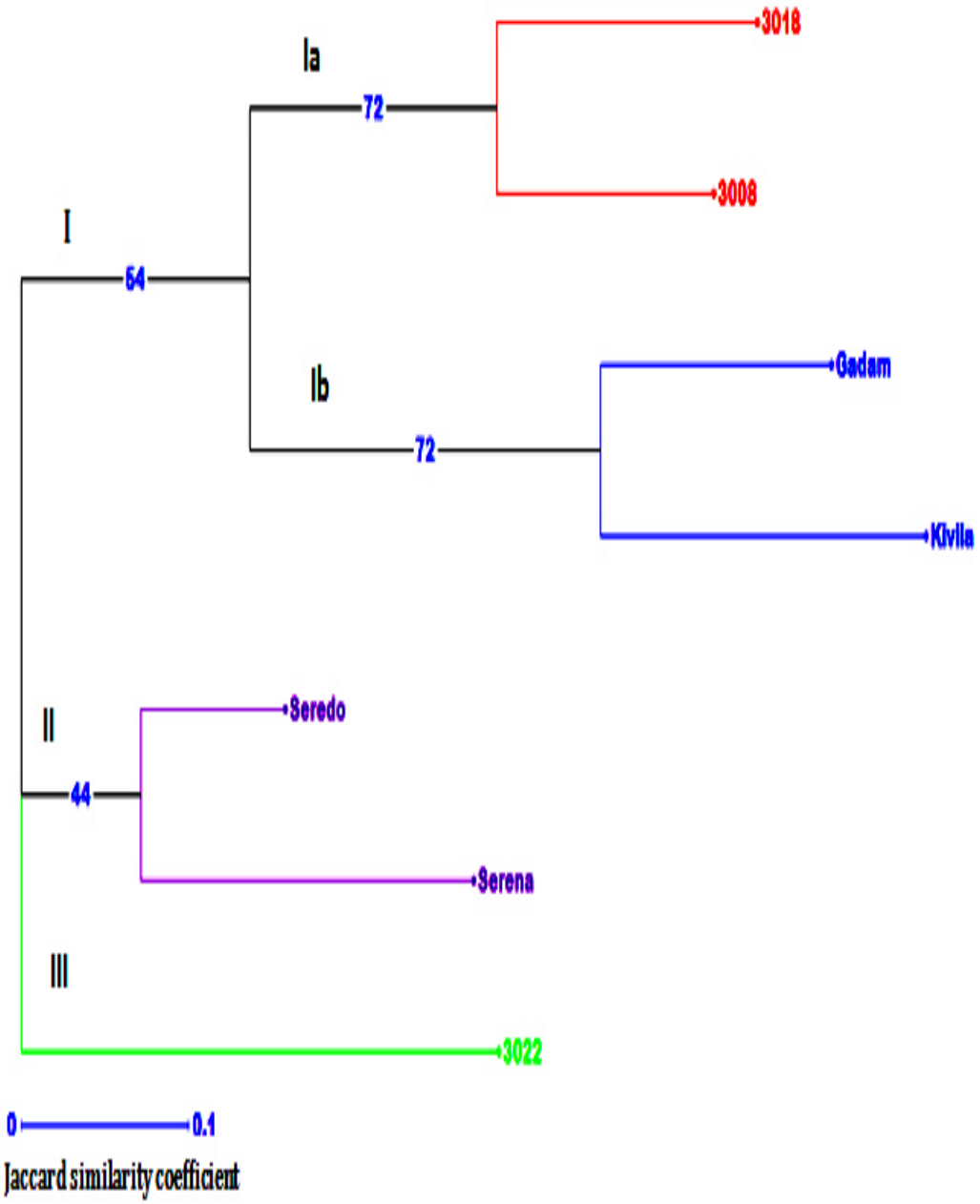

Figure 4.6a: UNJ phenogram of seven sorghum populations based on Jaccard similarity coefficient with ≥ 40 bootstrap support .The scale shown at the bottom is the measure of genetic similarity based on Jaccard coefficients.

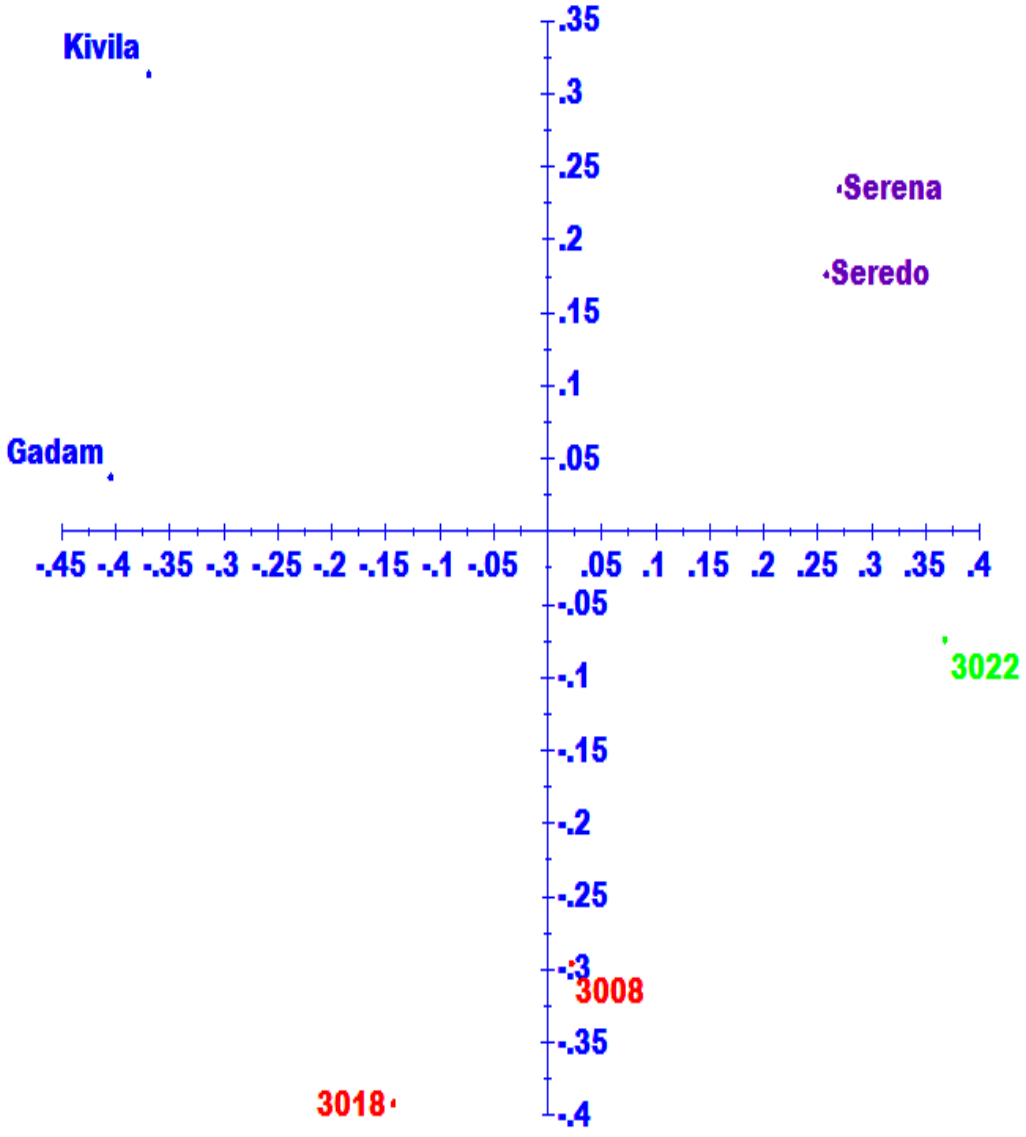

Figure 4.6b: The first two principal coordinates of the principal co-ordinate analysis (PCoA) using SSR genetic similarity matrix of seven sorghum populations. The plot was generated from a Jaccard similarity matrix using DARwin 6.0.12 dissimilarity analysis software.

CHAPTER FIVE

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

5.1 Discussion

5.2 Phenotypic Diversity using Phenotypic Markers

5.2.1 Analysis of Variance

The high significant differences ($p<0.01$) among the sorghum populations investigated traits suggests a substantial degree of genetic divergence among population studied. Population 3022 scored high mean values for a number of traits. This is an example of a tall variety producing considerably high yields as well as harboring desirable attributes of high quantity syrup. Despite Kivila and other tall populations being selected for syrup extraction, majority are susceptible to lodging. This is in congruence with what Murray *et al.* (2009) found. High mean values for number of leaves per plant and number of internodes per plant observed in 3008 and 3018 are indicators of accumulation of sucrose in stems of tall sorghum populations. Gadam is a short variety which qualifies to be selected for early maturing because it can escape drought. Although this population recorded high mean values for primary branches per panicle (yield component), it could be a source of genes for shortness with ability to cope with water stress in arid and semi arid lands. A high mean value for panicle length which is a yield component was observed in Kivila (landrace) although a low mean for number of primary branches per panicle was recorded. Many primary branches could have aborted probably due to water stress thereby recording low mean grain yield per panicle in Gadam. A wild sorghum population (*Sorghum halpense*) recorded significantly low mean for grain yield per panicle which could be due to considerably small size and low weighted grains. Ali *et al.,* (2008) observed moderate heritability for tillering in sorghum. Kivila recorded high mean values for number of tillers than other populations. This trait has an effect on total grain yield and so could be targeted for selection. Overall the mean separation of phenotypic traits clearly demarcated sweet and grain sorghum populations. Similar

argument was staged by Beyene, *et al.* (2013) working with highland maize populations. Additionally, the genetic divergence based on mean grain yield per panicle of sorghum populations significantly distinguished grain from sweet sorghums. The highly significant mean square values (Table 4.1a) observed in this study reveal high level genetic divergence among the studied populations indicating a wide genetic diversity among the populations under present study.

5.2.2 Bivariate statistics

Strong positive correlation observed in many traits under present study is indicative of high level of heritability and could be genetically controlled. Such a positive relationship was observed between grain yield per panicle (r=0.87, <0.01) and leaf area. The results of the present study are in consistence with the findings of Slewinski (2012). Deformed grains as well as primary panicle branch abortions due to water stress in the soil during grain filling stage could explain the existence of significantly negative correlations between grain yield per panicle and panicle length (r=-87, <0.01) . Similar relationship between this trait and panicle width (r=-0.82, <0.01) was observed. Positive correlation between leaf number and plant height depicted substantial level of competitive ability of sorghum populations for resources when intercropped (r=0.64, <0.01). Stem diameter and leaf sheath length (r=0.75, <0.01) are associated with high level of heritability as argued by Zheng *et al.* (2011). This feature has an implication on tolerance to lodging suggesting the significance of selecting both traits in any breeding programme. Significant positive association between leaf area and stalk diameter (r=0.91, <0.01) observed contributes substantially to food synthesized through photosynthesis and storage in stems. Bivariate analysis showed significant positive correlations between quantitative traits. Such positive correlations suggest gene linkage and similar biological background between correlated traits. Clearly these results support the suitability of use of phenotypic markers in genetic diversity studies.

5.2.3 Principal component analysis

The first four principal components cumulatively explained 97.19 % of the total variation (Table 4.1c). The first PC axes with a variance of 9.22 explained 57.61% of the total variation while the second, third and fourth principal component accounted for 19.68%, 12.76% and 7.14% of the total variance ,in that order. The variation in the first principal component was positively related with leaf area, leaf width and grain yield per panicle. The second PC was positively associated with plant height and leaf number. A third axis was moderately related to hundred seed weight and number of primary panicle branches while the fourth axis was positively associated with number of tillers. The relatedness of a trait to a given PC is based on the magnitude of the coefficient of a given trait in that PC. The higher the coefficient the more it is related to the component axis. Considering eigenvalues > 1 as significant, four principal components out of seven PCs were meaningful.

5.2.4 A two-dimension plot

The biplot display pattern (Figure 4.1) is explained by the similarities in number of internode, leaf sheath length, leaf width and leaf area of the studied populations. Grain sorghums clustered together as well as sweet sorghums. However, *S. halpense* and Gadam showed divergence from these groups. The distance between the panicle width and internode length traits was narrow as well as leaf width with leaf area. Plant height trait and number of tillers were projected by the vectors in opposite directions on the biplot display.

It was observed that the number of primary branches and hundred seed weight were projected close to the core. Many of other traits were projected distant away from the core depending on whether they were closely related or had a wide phenotypic divergence. The biplot analysis displayed the sweet sorghums (3018, 3008 and 3022) in the upper right quadrant indicating a close genetic relationship. Grain sorghums (Serana, Seredo) and a landrace (Kivila) clustered together as well, but in the lower right quadrant due to phenotypic relatedness. Gadam and *S. halpense* showed high divergence from the rest of the other populations. The loading biplot revealed that

some quantitative traits are closely related. Leaf area and leaf width showed a narrow genetic distance and probably could influence the grain yield per panicle. The two-dimension display is not only useful as a selection criterion but also an assessment tool for measuring the implementation of selection objectives in a breeding programme. Generally a wide phenotypic diversity among studied population was revealed.

5.2.5 Genetic distance and cluster analysis

Genetic distance matrix based on phenotypic traits ranged from 2.36 to 8.80. The pair-wise combinations for phenotypic traits measured indicated that the populations in the present study were diverse. A minimum genetic distance of 2.36 was observed between 3022 and 3018 while the highest genetic distance (8.80) was recorded for S. halpense and the rest of the populations. The wide genetic distance observed between the populations indicated variation and diversity. Cluster segmentation demarcated grain from sweet sorghum populations (Figure 4.2).Specific use of the sorghum genotypes could impose a positive selection pressure especially for the syrup extraction types. Ali *et al.* (2008) working with US sweet sorghums reported similar results.

Moreover, cluster means (Table 4.0) revealed differences in respect to the 16 quantitative traits studied. Thus, the populations clustered in different groups based on these traits. The dendrogram displayed the populations into 3 main clusters. Cluster 1 comprised three sweet sorghums (3022, 3018 and 3008) with a landrace (Kivila). The landrace was projected distantly from the rest within this cluster indicating that it was genetically diverse. A similar behavior was observed in cluster 2 comprising grain sorghums where Gadam separated from the rest. Members of cluster 1 recorded the highest values for leaf lengths, leaf area, internodes, stem diameter and leaf sheath. Cluster 2 was identified with high values of primary branches per panicle and lowest leaf count values. Cluster 3 was characterized by narrowest leaf, lowest values of number of tillers, thinnest stalk but highest leaf count, plant height and peduncle exertion. Clearly sweet sorghums were

phenotypically dissimilar with grain types. Therefore, the populations included in the present study which are genetically diverse could be a valuable genetic resource in sorghum improvement. This justifies the utility of cluster analysis statistical tool in classification of sorghum populations based on phenotypic traits.

5.2.6 Qualitative traits

Shannon - Weaver diversity index computation for qualitative traits (Plate 4) studied revealed a considerably high level of variation. However values >0.80 were observed in grain shape, grain color and plant color. Estimated phenotypic diversity (Table 4.1f) for the individual qualitative traits of the eight populations studied showed a significantly different pattern of diversity among the populations. Phenotypic diversity index ranged from 0.70 for grain covering to 0.95 for leaf mid rib color with an average phenotypic diversity index of 0.70. Plant color (0.86), grain shape (0.87), grain colour (0.80), and leaf mid rib color (0.95) were highly polymorphic whereas grain covering (0.70), panicle compactness and shape (0.73) and glume colour were relatively polymorphic. Since variation depended on the trait as shown by the wide diversity, the index proves useful in selection for the desired traits. The high H' values observed indicate a high level genetic diversity for quantitative traits. Therefore, Shannon Weaver diversity index is a useful statistic in determining phenotypic diversity of individual qualitative traits.

5.3 Multivariate Analysis of Germination and Seedling Trait under NaCl Treatment

5.3.1 Analysis of variance for effect of population and salinity tolerance traits

Calculated tolerance index showed a decline trend for mean values for growth and germination traits with respect to increase in salinity levels. The genetic variations observed were attributed to the intrinsic (genetic) differences between the populations and salinity levels. Additionally the variability in response based on difference in ionic concentration could represent genotype x environment interaction. The seed germination percentage decreased with increase in salinity level. High salt

concentration significantly inhibited germination. Similar results were reported by Duan *et al*. (2007). High germination rate does not necessarily reflect subsequent growth in biomass in a particular genotype as was observed with Serena. This is in line with what Azhar and Mcneilly (1987) expressed.

Toxicity effects of salts varied with populations, whereas tolerant populations showed fitness, the less tolerant were affected by these noxious salts at varying levels. Thus the observable effects in one population are compared with another trait which is a subject of environment and genetics. It is understood that fitness cannot be measured since it is difficult to interpret; this compels the use of germination and growth traits of seedlings under hydroponics.

The results revealed high mean significant differences (Table 4.2a) indicating that variation in terms of salinity tolerance was due to the effects of genetic nature of a population and levels of salt concentration. The coefficient of variation values observed in is indicative of high experimental reliability as well as diversity. Coefficient of variation is a measure with which treatments in an experiment are compared and it also expresses the experimental error as a percentage of mean according to Gomez and Gomez (1984). Therefore, the lower the CV value, the higher is the reliability of the experiment. A genotype x environment interaction is significant to a plant breeder. Therefore population x salinity interaction was relevant for characterizing the populations under present study based on performance of individual populations at varying salinity levels. Increase in growth resulted from increased cell number and expansion. Under continuous exposure to saline conditions the cell number decreases and a number of cells may succumb resulting to low growth. The effect of salt on tissue and organ development is reflected in altered patterns of plant growth. Munns *et al.* (2006) reported that continuous exposure to elevated root-zone salinity progressively decreases leaf size over time, necrosis, chlorosis and poor lateral root formation. Leaf necrosis (Plate 4.1a) in sensitive populations is a manifestation of toxicity due to damage of membrane associated with Reactive Oxygen Species (ROS) as was observed in Serena under 15 EC levels eventually leading to cell death. Additionally, chlorosis observed in 3008 (Plate 4.1b)

is a symptom of nutrient imbalance due to uptake of Na and Cl ions which could replace K, Ca and NO_3- ions. The net effect of chlorophyll content reduction could result in low photosynthetic rate with poor conductance and CO_2 assimilation. This also explains the poor lateral root formation observed in Serena and 3008. Research shows that the tolerant populations are more endowed with the ability of their genes to encode for antioxidant enzymes which includes superoxide dismutase (SOD), ascorbate peroxidase (APX),and catalase (CAT) and Glutathione Reductase (GR) which scavage the ROS. McKersie and Leshem (1994) reported the existence of substantial quantities of CAT within the cytosol, mitochondria and peroxisomes of cells of plants that are salt tolerant. This particular enzyme is known to dismutate H_2O_2 into its constituent harmless hydrogen and oxygen atoms.

The findings under the present study indicate that tolerant varieties suffered less from salinity effects suggesting that they have genes that encode antioxidants that dismutate ROS. Therefore germination potential and seedling traits as well as other measures of fitness against salinity stress could serve as good markers for salinity tolerance. The results from the present study indicated that all salinity treatments affected all the traits adversely. Evidently the mean values for growth and germination traits decreased drastically with increase in salinity levels as shown by the calculated tolerance index and the effects varied with individual sorghum populations. This is in agreement with what Jeannette *et al* (2002) found. Salinity levels of 10 and 15 dS/m of salt concentrations recorded higher reduction in growth and germination mean values compared to the control, hence the tolerance index pattern.

5.3.2 Salinity tolerance index

Salinity stress induces leaf toxicity thereby reducing photosynthesis per unit leaf area by interference of the photosynthetic apparatus. Mehmood *et al.* (2004) rated this as one of the most important constraints to sorghum productivity in arid and semi arid environments. Serena recorded a sharp decline in terms of tolerance based on wet weight trait whereas high tolerance index in 3018 population was observed. This

indicates that it is difficult to select this wet weight trait. Overall mean values for tolerance index (TI) demonstrated a significant decline in TI with regard to salinity levels suggesting existence of diversity. High values observed at 5 dS/m suggested that substantial growth was promoted at this treatment for the seven sorghum populations.

Ultimately the analysis based on fitness traits separated the sorghum populations into three categories. Population divergence pattern revealed by computation of total percentage variation for individual populations with regard to tolerance index across the three salinity levels showed a substantial level of genetic diversity among the population studied. Seredo, Gadam, 3022 and Kivila were identified as tolerant populations, 3018 and 3008 stood out as moderately tolerant and Serena as a considerably sensitive tolerant population. The observed delineation is ideal for selection of tolerant genotypes. Although Seredo and Gadam which are grain sorghums stood out as tolerant, a general trend of sweet sorghums being salt tolerant was observed. Krishnamurthy *et al.* (2007) found a similar relationship and support the fact that stem sugar plays a role in moderating adverse effects of salts in sorghum plant systems. The result of present study demonstrates how useful this method could be used for selection of salt tolerant populations which could be used for selection of future stock for crop improvement.

5.3.3 Bivariate Analysis

Strong positive correlation as was observed between shoot length and root length as well as other traits in the present study has a survival value of utilization of available moisture in the soil by a large number of seeds during early development stage. Such seeds are endowed with hydrolytic enzymes that break down food material within the seed. Consequently sugars are immobilized down to the developing embryo which constitute the young root and shoot. This could suggest that there is a proportionality property on growth rate between the radical and the young shoot. A radical that responds fast to growth shows low sensitivity to salinity stress and could pump water up the shoot adequately hence cell expansion. However, the impact of sensitivity was

higher in shoot than roots. This agrees with the findings of Asfaw (2011) working with *sorghum bicolor*. It is imperative that breeders target traits showing positive relationship for crop improvement. According to Falconer and Mackay, 1996, positive genetic correlations between traits could share a common biological process or could be components of the same structure. Negative genetic correlations are often found between components of fitness as well. The studied traits are components of fitness because they are interpretable; however they could show negative genetic correlations. The negative effects observed in shoot length trait with root- shoot ratio imply that the latter is undesirable. It indicates that the scope of crop improvement in terms of salinity tolerance for root- shoot ratio of the selected sorghum populations under study could be limited. The traits could be selected separately however. Overall, the strong positively correlated traits suggest that they share a common biological process; therefore such traits could be inherited together hence useful during selection for improvement of sorghum populations.

5.3.4 Principal component analysis

The first three principal components, with eigenvalues greater than unity, explained 79.98 % of the total variation among the seven populations for the tolerance traits studied. The first and second principal component accounted for 45.10 and 20.07% of the total variation, in that order with a cumulative variation of 65.17. PC1 axis was mainly contributed by wet weight, root length and root-shoot ratio. There was no significant variation within all genotypes with regard to these indices, the reason why this was observed. Germination rate and shoot length recorded negative loadings in PC1. This could be explained by the fact that there was considerable variation of genotypes with regard to these two traits. However, the shoot length had the same weight with wet weight in this particular component but in opposite direction. Eigenvector which is the trait coefficient of a particular principal component is considered significant when the value is greater than half divided by the square root of the standard deviation of the eigenvalue (Johnson & Wichern, 1988). Germination percentage and tolerance index had high coefficients and were major contributors of PC2 variation. Germination percentage, root shoot ratio, shoot length tolerance index

and wet weight were identified with the PC3 but with negative loadings. Dry weight showed the highest variation in this component. This could be explained by the fact that increased salt concentration caused simultaneous decrease in shoot and root dry weights in all genotypes. This is in congruence with what (Afzal *et al*., 2005) found. This analysis was highly effective as the first three components accounted for 79.98 % of the total variation. Subsequent PC percentage variation shown by the scree diagram (Figure 4.6) represents the overall seven dimension principal component which constitutes 100% of the total variation.

5.3.5 Scree plot

The scree plot pattern displayed the seven dimension PCA analysis in a linear fashion, thereby presenting 100% variation. The graphical statistic delineated the PC axes from 45.10% to 0.29%. PC 1 recorded the highest variation at 45.10% followed by PC2 with 20.07%. An elbow was detected at Eigenvalue 1.037 scoring 14.81% of the total variation. Thus, the components above this value were considered meaningful. A scree plot (Figure 4.6) of eigenvalues versus their associated component can identify the point at which the eigenvalues decrease in a linear fashion. Since the eigenvalues relate to the variance explained by each component, only retaining those that are above this linear trend results in a parsimonious set that account for a large share of the original variation. Alternatively, Kaiser's Rule suggests that only those variables with eigenvalues of greater than 1 should be retained, because each of the standardized original variables has variance of 1 and thus the new variables should account for more variation.

5.3.6 A two-dimension plot

The biplot (Figure 4.5) statistic differentiated the populations with the studied traits explained by the first two PCA dimensions. The biplot clustered the populations within the four quandrants of the two dimension plot as well as the traits, which were projected by vectors thereby indicating a wide genetic diversity among the populations studied. Dry weight and wet weight are traits projected in different quandrants in the principal component axes suggesting a diverse association between

these two traits. A close relationship was recorded between germination rate and germination percentage. Shoot length diverged away from root length indicating a wide diversity based on tolerance to salinity. Higher negative effects of salt concentration of in shoot dry weight than seedling root dry weight could account for this observation.

However a strong and positive association was noted between root length and root-shoot ratio. It is surprising that the biplot demarcated grain from sweet sorghums; however 3018 distanced itself from populations 3008 and 3022. Nevertheless, Kivila a landrace population revealed a close relationship with 3008 and 3022 which are sweet sorghum populations. Gadam and Seredo grain sorghums diverged from Serena showing high diversity, and could present a considerable level of salt tolerance. Such plants presenting high tolerance traits are known to accumulate considerably high quantities of sucrose (Asfaw *et al.,* 2011) and proline under salt stress conditions raising the osmotic potential levels of the cells thereby conferring salinity tolerance to such populations. This agrees with the fact that salinity tolerance is genetically controlled and varies from species to species as well as from one population to another. At molecular level, HKT genes are responsible for conductance of Na+ ions throughout the plant under salt stress hence mediate salinity tolerance mechanisms. Positive correlation between Na+ concentration and HKT gene expression in roots and leaves has been reported from many studies (Moller *et al.* 2009). According to Yao *et al.* (2010) members of the grass family such as rice and barley showed that HKT genes encoded proteins associated with movement of Na+ ions within the cells. Nonetheless, salinity tolerance is a complex trait which is controlled by many genes and so the characters studied are quantitative themselves. Thus Gadam, Seredo, Kivila and 3018 could be a good source of salinity tolerance genes for intercrossing with other populations under study to cope with saline environments of arid and semi arid lands in Kenya. This is of ultimate importance as it offers an opportunity to utilize diverse parents in hybridisation programmes which could provide better hybrids.

5.3.7 Classification of populations

Cluster analysis (Figure 4.7) gathered sorghum populations in three main clusters. Cluster 1 comprised two populations; Serena and Seredo. These two are improved cultivars specifically for the dry areas which are prone to salinity. Both are moderately salt sensitive and showed similarities in growth of shoots and roots. Their shoots and roots declined drastically with increase in salt concentration. Simultaneous decline in shoot and root dry weights occur due to elevation of salt stress in genotypes. Elsewhere Asfaw (2011) reported a similar observation. Cluster 2 grouped Gadam and 3018 (sweet sorghum) at 1.83 coefficient distance. Both populations have juicy stems and this could explain why they show similarities in terms of salinity tolerance. The population Kivila (landrace), 3008 and 3022 were grouped in cluster 3. They are tallvarieties with long internodes and capable of packing sucrose and water in their parenchyma. They have considerably large leaf surface area for elimination of excess salts avoiding accumulation. This explains why they could be grouped together with respect to tolerance to salt stress trait. Minimum coefficient (1.14) of dissimilarity was recorded between Kivila and 3022. This genotype (3022) was placed at a Manhattan genetic distance 2.25 from the other two members of this cluster. The genetic distance (Table 4.2e) between Seredo (2.88) and Serena shows a narrow genetic base. This could plausibly be explained by the fact that a narrow genetic distance shows a close relationship indicating that they could have descended from a common ancestral population.

The high divergence between members of cluster 3 (3.74) and four other populations in cluster 1 and 2 shows a wide genetic diversity suggesting that these populations could be out crossed with these two sweet sorghum populations and a landrace to obtain salinity tolerant genotypes. In general, the high genetic diversity observed among the studied populations indicated that the individual populations descended from divergent ancestral populations and so could be utilized as parental lines for heterosis breeding programmes to improve salinity tolerance in sorghum populations.

5.4 Heritability Values and Population Divergence

5.4.1 Estimates of variance components

Substantially high genotypic and phenotypic variances were observed in shoot length, root length, germination percentage and germination rate at 5dS/m treatment. Coefficient of variance values as well as scores for heritability was higher at control than that of respective selection pressures (5dS/m, 10dS/m and 15dS/m). Generally, the values of PCV were higher than their corresponding GCV values. A number of traits exhibited considerably high genotypic and phenotypic variances at various selection pressures (Table 4.3). These values were lower compared with the control. Generally, the values of PCV were higher than their corresponding GCV values. Based on the PCV and GCV obtained values above 20% are categorized as high, those values which range from 10 % to 20 % as medium and values below 10 % are categorized as low. Following this range at 150mM, germination rate (28.6 %) scored high PCV values. Medium range of PCV values was recognized in wet weight (13.7%) and root shoot ratio (13.7%). Low PCV values were observed in dry weight (9.7%) germination percentage (6.6%), root length (9.6%) and shoot length (3.58%). High genotypic coefficient of variance (GCV) values (15dS/m) was observed in germination rate (21.6%).Conversely, low GCV scores were recorded in germination percentage (5.2%), root length (6.7%), shoot length (2.3%), root shoot ratio (9.3%), dry weight (7.0%) and wet weight (8.5%). Moreover, a wide range of PCV and GCV values was observed in root shoot ratio (13.7%-9.3%) and in wet weight (13.7% - 8.5%) trait. This indicates that there is a high influence of environment on these traits. In comparison the range was considerably low for dry weight (9.6764-7.0%), germination rate (6.6%-5.2%), root length (9.6%-7.0%), germination percentage (6.6%-5.2%) and shoot length (3.6%-2.3 %). Such low difference between PCV and GCV values is an indicator of possibility of selection of these traits for improvement of salinity tolerance parameter in sorghum populations under present study.

5.4.2. Estimation of heritability in broad sense

Heritability in broad sense values ranged from 49.0 % for wet weight to 77.1 % for germination percentage at 10dS/m. At 15dS/m the values were lower at 38.57 % for wet weight and 62.58 % for germination percentage. High heritability of a particular trait above 75 % is attributed with ease of selection. In such a case, small contribution of the environment to phenotype results in close correspondence between genotype and phenotype. It is easier for a breeder to select such traits. Whereas heritability of wet weight (38.57%) was low, dry weight (51.76%) recorded moderate heritability (40 % to 80%). It has been found difficult to select traits which recorded heritability values below 40 % because of high contribution environmental effects to these traits. Moderate heritability (40 % to 80%) for dry weight (51.76%), germination percentage (62.6%), germination rate (57.2%), root length (48.6%), root-shoot ratio (45.8%) and shoot length (42.2%) was observed at 15dS/m selection pressure. Azhar and McNeilly (1989) working with sorghum under NaCl stress found similar broad sense heritability of 73 % (5dS/m) and 38 %(20dS/m). Clearly; heritability could be utilized in predicting breeding of salinity tolerance traits based on the salinity tolerance trait values scored.

5.4.3 Allelic variation for microsatellite loci and allele frequency in studied populations

Five of the SSR markers were polymorphic because of having allele frequencies ≤0.99. According to Cavalli–Sforza and Bodmer (1981) a polymorphic gene is usually one for which the most common allele has a frequency of less than 0.95. The value (0.95 or 0.99), is said to be the limit of allele frequency with the objective of identifying those genes in which allelic variation is common. In contrast, rare alleles are regarded as those with frequencies ≤ 0.05. Therefore no rare allele was detected in the present study. Five of the other loci were monomorphic. This means that, 50% of the markers were polymorphic generating a total of 10 alleles. These findings show consistence with what Ji *et al*, (2011) found. The proportion of polymorphic loci expresses the percentage of variable loci in a population and a value ≤0.5

indicates a substantial discriminatory power of genotypes by SSR markers. Majority of the loci produced 2 alleles with an average of 2.0 which could probably be explained by use of majority di-repeat containing SSR markers and lack of a variety of the SSR series. Furthermore, the average number of alleles per locus refers to a measure that provides complementary information to that of polymorphism.

Polymorphism information content (PIC) values were estimated for the five SSR loci using the allele frequencies. The frequency of the major alleles at different SSR loci ranged from 0.50 (Xtxp 021) to 0.64 (Xtxp 014, Xcup02, Xgap08). Loci Xtxp34 scored 0.57. The allele frequency is important in determining the genetic diversity of studied populations. Nei (1975) indicated that the diversity index depends on the allele frequencies and their respective number of alleles. The PIC value observed over the five SSR markers ranged from 0.3538 to 0.3698 with an average of 0.3612. Bharadwaj *et al*. (2011) reported a considerably high level of PIC values in chick pea which could be attributed to a TAA motif. In essence the PIC of a given SSR marker provides an estimate of the discriminatory power of that SSR marker.

The findings of the present study indicate that SSR markers are highly informative and could detect a number of alleles. This confirms the usefulness of SSR markers in genetic analysis. The population with the lowest PIC value in the loci of Xtxp 014 and Xgap084 were from Kivila and Gadam, which harbored unique alleles of size 200bp and 150bp respectively.

Out of a total of 10 markers, five markers, Xtxp141, Xcup60, Xcup07, Xtxp273, and Xcup37 were monomorphic with regard to the population in the present study. But five markers (Table 4.4a) of Xtxp, Xcup and Xgap SSR series successfully discriminated the seven populations studied thereby revealing the level of genetic diversity reported. Only one SSR marker amplified more than one fragment in the same genotype. Based on the argument forwarded by Agrama and Tuinstra (2003), this finding suggests the existence of a substantial level of residual heterogeneity within the particular population.

Considerably higher PIC values were observed in Xtxp than Xcup and Xgap SSR series. This is attributable to the differences in the SSR origins according to Casa *et al*., (2005). Kong *et al*., (2000) reported that Xtxp and Xgap SSR markers could be generated from small-insert genomic libraries or bacterial artificial chromosome end sequences. According to Schloss *et al*.,(2002) these loci could include non coding regions compared to the Xcup SSRs series that are mainly developed from low-copy RFLP probe sequences and are thought to be located near or in genes. Additionally, a general trend of di-nucleotide repeat containing markers producing higher PIC values than tri-repeat containing markers was observed. This could be explained by their property of undergoing a higher mutation rate hence the observed high variability.

The gene diversity also referred to as expected heterozygosity (H_e) ranged from 0.4592 to 0.5000. Average gene diversity observed 0.4735 was substantially high which could be explained by high level of inbreeding observed in sorghum. The average expected heterozygosity (H_e) is the probability that, at a single locus, any two alleles, chosen at random from the population, are different to each other. Gene diversity could as well be influenced by size homoplasy which occurs when different copies of a locus are identical in state but are not identical by descent. Moreover, high genetic diversity among the seven studied sorghum populations was revealed by the PIC and gene diversity values generated in the analysis. Thus, the relative frequencies of detected alleles accounts for the ability of PIC values to give information about the discriminatory power of a particular marker.

The heterozygosity (H_O) values for each locus ranged from 0.4286 to 0.8942 with an average of 0.6646. The average observed heterozygosity over all loci is an estimate of the extent of genetic variability in the population. Strong differences in heterozygosities recorded between SSR loci used in the present study are useful in delineation of the populations under study. Therefore the high level of heterozygosity clearly indicated that the studied populations are highly heterozygous and could be utilized in improvement of heterosis levels of new progenies from heterotic parents.

5.4.4 Cluster analysis

Cluster analysis based on Jaccard similarity coefficient with Unweighted Neighbor Joining method was performed on the 10 SSR markers for seven populations. The cluster analysis sorted the populations into three main clusters. Cluster 1; consist of two sub clusters Ia and Ib. The pair-wise dissimilarity indices among the sorghum genotypes grouped two sweet sorghums; 3018 and 3008 in sub-cluster Ia at genetic distance 0.286. Cluster Ia members are of Chinese origin with tall thick stems to pack sucrose. This group showed a substantial level of heterozygosity based on banding pattern. Such analysis is indicative of genotypic relatedness at DNA level and the utility of both sorghum populations.

Sub cluster Ib comprises Gadam (grain sorghum) and Kivila (landrace). The members of this cluster showed utmost level of variation and individuality in their banding outline from both the sweet and grain sorghum categories for populations under study. Kivila which is a landrace is of Akamba community origin has an upright semi compact panicle. This cultivar is known for high grain yield, early maturing and for its tolerance to high temperatures and drought. In Kenya Kivila has been predominantly used as a source of genes for enhancing terminal drought tolerance in sorghum due to its stay green attribute. Gadam is a short and an early maturing dry land, high grain sorghum which is presumably a derivative of Kivila (landrace). Both populations have white grains and are highly performing in dry and hot land of Makueni County in Kenya. The observed close relationship (similarity distance=0.333) between these two populations in this classification (Table 4.4c) indicates genetic relatedness.

Cluster II consisted of two grain sorghum populations (Serena and Seredo) both cultivated in eastern Kenya. Both are medium in height with a considerably long compact pannicle. Both are improved cultivars from KARLO and their parental ancestors could be traced from Uganda. Serena and seredo clustered together at a genetic distance 0.286 indicating high genetic closeness. It is important to note that

cluster II separated from the main grain sorghums in sub cluster Ib indicating high genetic unrelatedness.

Cluster III sorted 3022(sweet sorghum) a high grain yielding population individually recording the highest genetic distance 0.827 with Kivila. The observed genetic divergence offers an opportunity which could be utilized in selection of heterotic parents. The population singled out from the rest of sweet sorghums showing that these are genetically diverse populations. This is contrary to what Yang *et al* (1996) argued, that sweet sorghums especially from china could be genetically homogeneous due to prolonged isolation. Thus, the analysis in the present study indicated that SSR markers could group sorghum populations based on grain and syrup production.

Cluster analysis of microsatellite data from sorghum populations under study revealed that sub clusters Ia and Ib were strongly bootstrap supported (Figure 4.9a). Similarly, genetic distance data from polymorphic loci in the present study clustered populations regardless of their category (sweet or grain sorghum) and origin. It is evident that sweet sorghums were genetically distant from grain sorghums. However the close grouping between sweet and grain sorghum populations in cluster I could suggest the presence of higher levels of gene flow between sorghum populations due to gene exchange in form of gametes and seeds according to Slatkin (1987).

5.4.5 Principal coordinate analysis (PCoA)

To validate these results obtained using SSR markers, PCoA was performed, which showed that the first two principal coordinates accounted for 78.10 % of total variation, and agreed with the bootstrap values (≥ 40). Coordinate calculated for the two first axes had positive eigenvalues, 0.08517 and 0.06173 for the first and second coordinates in that order. Additionally the first and second coordinates extracted 45.28% and 32.82% of the total variation, respectively. A two dimensional display (Figure 4.9b) revealed that the dispersion of the genotypes of the detected clusters corresponded to the one observed in the dendrogram but with some exceptions. Sweet sorghum populations 3018 and 3008 clustered together in the dendrogram, but

in the biplot 3008 was separated from the group and clustered together with 3022 which is also a sweet sorghum population. This could be attributed to the pedigree relationship with the member (3022) in cluster III of the dendrogram. Grain sorghums Serena and Seredo sourced from KARLO clustered together in the same plane and were very close in the biplot as was the case in the dendrogram. This indicates a narrow genetic distance between the improved sorghum populations. Better separation based on genetic similarity, origin and utility (sweet or grain) displayed by PCoA was evident. Although Kivila and Gadam showed a close relationship in the dendrogram, they dispersed far apart in the PCoA.

This suggests that Kivila, a landrace and Gadam are potential sources of genes for improvement of other grain sorghums. Gower (2005) argued that PCoA seeks to position objects in fewer dimensions while maintaining their distances as accurately as possible. Therefore, the existence of wide genetic distances between, 3018 and 3022, as well as the relationship between 3008 and 3022 as represented by the PCoA depicts accuracy thereby reflecting these populations as future genetic resources. Clearly, PCoA provided a better diversity structure than the dendrogram since PCoA calculated from dissimilarity used two dimensions, compared to one dimension for the dendrogram.

5.4.6 Population equilibrium within seven sorghum populations

The locus Xtxp 021 scored a highly significant p value (0.0472). A probability of 4.7% (p-value of 0.047) means that the observed and expected values differences are due to chance and 95.3 % are not due to chance indicating non - adherence to Hardy-Weinberg equilibrium. Such populations are suspected to be evolving with respect to this locus probably due to natural selection or nonrandom mating or other forces. X^2 computation showed that loci Xtxp34 (p-value=0.6592), Xtxp 014 (p-value=0.1416), Xcup02 (p-value=0.2703), Xgap084 (p-value=0.1416) adhered to Hardy-Weinberg equilibrium indicating that there is no evolution going on in terms of these loci (Table 4.9).

Overall, this study identified great genetic diversity and heterogeneity in sorghum population studied. This information could be a reference for parental selection in hybrid development to maximize the heterosis in sweet and grain sorghum breeding program and is helpful in creating segregating populations to map genes controlling grain and sugar content in sweet and grain sorghum. Thus, the occurrence of distinct groups of sorghum accessions as revealed by SSR marker analysis can be utilized effectively in pre-breeding efforts to overcome yield barriers.

5.5 Conclusion

- The findings under the present study support the continued use of phenotypic markers for genetic diversity studies.
- Statistical analysis strongly differentiated the populations studied into divergent groups demarcating sweet sorghums from grain sorghums using the phenotypic markers.
- The Sorghum populations studied exhibited considerably high variability based on salinity tolerance and were highly divergent providing adequate choices for selection. Seredo and 3022 genotypes emerged highly tolerant in relation to many traits studied.
- The study highlighted multivariate analysis and correlation analysis of measures of fitness for salinity tolerance as a highly useful statistical tool for distinguishing sorghum populations with respect to salinity tolerance - quantitative trait without ambiguity.
- Coefficient values as well as scores for heritability were higher at control than that of respective selection pressures (5dS/m, 10dS/m and 15dS/m). Germination percentage recorded 62.5% heritability at 15dS/m which agrees with what Azhar and McNeilly (1989) found. The findings indicated that all traits studied except fresh weight could be easily improved and incorporated in the breeding programmes. This study justifies the utility of heritability as a useful statistical tool for genetic diversity studies.
- The study of seven sorghum populations revealed a substantial level of the utility of SSR markers in unraveling the genetic similarities and differences

among the set of sorghum populations. Molecular diversity indicators, multivariate analysis, cluster analysis and PCoA statistical tools distinguished genotypes into divergent groups without ambiguity. Cluster analysis sorted out the genotypes into 3 distinct groups. The PCoA in particular, provided a better diversity structure than the dendrogram since PCoA calculated from dissimilarity used two dimensions, compared to one dimension for the dendrogram.

- Examination of genetic indicators, such as polymorphic information content, allele frequency, allele number, gene diversity and heterozygosity generated, revealed the ability of SSR markers to discern genetic variation among genetically diverse sorghum populations.
- Considerably high level of genetic diversity and heterogeneity was evident although the markers could not clearly discriminate sweet and grain sorghums. However, some sugar related markers detected could be useful in marker-aids screen for sugar content in sorghum stems.
- Such a great genetic diversity provided an opportunity for utilization of the existing distinct groups revealed by SSR markers to overcome yield barriers through implementation of effective selection methods.
- The distinct groups are a genetic resource that can provide great genetic gains if the landraces could be incorporated in breeding programmes.

5.6 Recommendations

- The pattern of genetic diversity revealed with microsatellites, phenotypic markers, and multivariate analysis of salinity tolerance measures of fitness and heritability analysis as suitable statistical tools in population genetics offers an opportunity to mine useful genes from sorghum genetic resource. Determination of gene flow mechanism by use of these statistical methods is recommended.
- It is important to develop diverse heterotic parents with high salt tolerance potential through selection from identified populations in this research. Till that time some individuals identified here could serve as candidate parents as

well as cultivars in salt prone areas. Further profound field investigation at whole plant growth stage should be done.

- These methods together with the genetic structure of these populations should be incorporated in the selection criteria for heterotic groups aimed at breaking sorghum performance barrier so as to feed the high population living in arid and semi-arid regions in Eastern Kenya. Gadam, 3018, Kivila and 3022 stood out superior candidates for selection to obtain heterotic parents for crop improvement.

REFERENCES

Abdi, A., Bekele, E., Asfaw, Z. & Teshome, A. (2002). Patterns of morphological variation of sorghum (*Sorghum bicolor* (L.) Moench) landraces in qualitative characters in North Shewa and South Welo, Ethiopia. *Hereditas, 137*, 161-172.

Abu Assar, A.H., Uptmoor, R., Abdelmula, A.A., Salih, M., Ordon, F. & Friedt, W. (2005). Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USAassessed by Simple Sequence Repeats (SSRs). *Crop Science, 45,* 1636-1644.

Afzal, I., Maqsood, S., Basra, A., Ahmad, N. & Farooq, M. (2005). Optimization of hormonal priming techniques for alleviation of salinity stress in wheat (*Triticum aestivum* L.). *Caderno de Pesquisa S'r. Bio., Santa Cruz do Sul, 17*(1), 95-109.

Agarwal, M., Shrivastava., N. & Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell. Rep., 27*, 617-631.

Agrama, H.A. & Tuinstra, M.R. (2003). Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *African Journal of Biotechnology, 2,* 334-340.

Ahnert, D., Lee, M., Austin, D., Livini, C., Woodman, W., Openshaw, S., … & Dalon, G. (1996). Genetic diversity among elite sorghum inbredlines assessed with DNA markers and pedigree information. *Crop Science, 36*, 1385-1392.

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

Allard, R.W. (1999). *Principles of Plant Breeding*. (2nd ed.). New York, USA: John Wiley and Sons,.

Al-Sadi, A.M., Al-Wehaibi, A.N, Al-Shariqi, R.M., Al-Hammadi, M.S., Al-Hosni, I.A., Al-Mahmooli, I.H., & Al-Ghaithi, A.G. (2013). Population genetic analysis reveals diversity in Lasiodiplodia species infecting date palm, Citrus, and mango in Oman and the UAE. *Plant Dis*. *97*, 1363–1369.

Altintas, S., Toklu, F., Kafkas, S., Kilian, B., Brandolini, A. & özkan, H. (2008). Estimating genetic diversity in durum and bread wheat cultivars from Turkey using AFLP and SAMPL markers. *Plant Breeding* 127: 9-14.

Anderson, L. K., Doyle, G. G., Brigham, B., Carter, J., & Hooker, K. D. (2003). High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics, 165*, 849–865.

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

Asfaw, K. G. (2011). Effects of Salinity on Seedling Biomass Production and Relative Water Content of Twenty Sorghum (*Sorghum biolor* L. Moench) Accessions. *Asian Journal of Agricultural Sciences, 3*(3), 242-249.

Ayana, A. & Bekele, E. (1998). Geographical patterns of morphological variation in sorghum [*Sorghum bicolor* (L.) Moench] germplasm from Ethiopia and Eritrea: qualitative characters. *Hereditas, 129,* 195-205.

Azhar, F.M., & McNeilly, T., (1989). Heritability estimates of variation for NaCl tolerance in *Sorghum bicolor* (L.) Moench seedlings. *Euphytica*, *43*, 69–72.

Beyene, Y., Botha, A.M. & Myburg, A.A. (2005). A comparative study of molecular and morphological methods of describing genetic relationships in traditional Ethiopian highland maize. *African Journal of Biotechnology, 4*, 586-595.

Bharadwaj, C., Srivastava, R., Chauhan, S. K., Satyavathi, C.T. & Kumar, J. (2011). Molecular diversity and phylogeny in geographical collection of chickpea (*Cicer* sp.) accessions. *J. Genet., 90,* e94-e100.

Blum, A. (1979). Genetic improvement of drought resistance in crop plants. A case for sorghum. In: Mussel, H. and R.C. Staples (eds.), *Stress Physiology in Crop Plants*, pp: 429–245. New York: John Wiley and Sons, Inc.

Botstein, D., White, R.L., Skolnick, M., & Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Gene, 32*, 314-331.

Carrera, A.D., Poverene, M.M. & Rodriguez, R.H. (1996). Isozyme variability in Helianthus agrophyllus. Its application in crosses with cultivated sunflower. *Helia, 19*, 1928.

Casa, A.M., Mitchell, S.E., Hamblin, M.T., Sun, H., Bowers, J.E., Paterson, A.H., … & Kresovich, S. (2005). Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. *Theoretical and Applied Genetics, 111*, 23–30.

Cavalli–Sforza, L.L. & Bodmer, W.F. (1981). *Genética de las Poblaciones Humanas. Ed.* Barcelona: Omega.

Chakauya, E., Tongoona, P., Matiburi, E.A., & Grum, M. (2006). Genetic diversity assessment of sorghum landraces in Zimbabwe using microsatellites and indigenous local names. *International Journal of Botany,* 2.

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

Craufurd, P.Q., Wheeler, T.R., Ellis, R.H., Summerfield, R.J., & Williams, J.H., (1999). Effect of temperature and water deficit on water-use efficiency, carbon isotope discrimination and specific leaf area in peanut. *Crop Sci. 39*, 136–142.

Dillon, S.L., Lawrence, P.K., Henry, R.J. & Price, H.J. (2007). Sorghum resolved as a distinct genus based on combined ITS1, ndhF and Adh1 analyses. *Plant Systematics and Evolution,* 268, 29-43.

Duan, D., li, W., Ouyang, H., & An, P. (2007). Seed germination and seedling growth of *Suaeda salsa* under salt stress. *Ann. Bot*. *Fennici, 44*, 161- 169.

Evan, M. (2009). *Crop-to-wild geneflow: environmental risk assessment for the release of genetically modified sorghum in Kenya*. Unpublished PhD Thesis, Bloemfontein, South Africa: University of the Free State.

Falconer, D.S., & Mackay, T.F. (1996). *Introduction to Quantitative Genetics,* Longman: Harlow.

FAO, (2011). FAOSTAT. Retrieved from: http://faostat.fao.org.

Folkertsma, R.T., Frederick, H., Rattunde, W., Chandra, S., Raju, G.S. & Hash, C.T. (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.

Gascuel, O. (1997). Concerning the NJ algorithm and its unweighted version, UNJ. In: Mathematical Hierarchies and Biology. DIMACS workshop, Series in Discrete Mathematics and Theoretical Computer Science. Am. *Mathematical Soc. 37,* 149-170.

Gaston, K.J. (1998). Species–range size distributions: products of speciation, extinction and transformation. *Philosophical Transactions of the Royal Society, London. 353*, 219–230.

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

GenStat Discovery, (2004). Version number 12.1.0.3338., 12th Edition Program

Geressu, K. & Gezahagne, M. (2008). Response of some lowland growing sorghum (*Sorghum bicolor* L. Moench) accessions to salt stress during germination and seedling growth. *Afr. J. Agric. Res., 3*(1), 044-048.

Ghebru, B., Schmidt, R.J., & Bennetzen, J.L. (2002). Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. *Theor Appl Genet. 105*, 229–236.

Gomez, K.A. & Gomez, A.A. (1984). *Statistical procedures for agricultural research,* (2^{nd} ed.), pp. 322. New York:A Wiley Inter-science Publication,

Gower, J.C. (2005). Principal Coordinates Analysis *Encyclopedia of Biostatistics*. London: John Wiley and Sons, Ltd.

Grenier, C., Bramel, P.J., Dahlberg, J.A., El-Ahmadi, A., Mahmoud, M., Peterson, G.C., Rosenow, D.T. & Ejeta, G. (2004). Sorghums of the Sudan: analysis of regional diversity and distribution. *Genetic Resources and Crop Evolution, 51,* 489-500.

Guo, Y., Chen, S., Li, Z., & Cowling, W.A. (2014). Center of origin and centers of diversity in an ancient crop, Brassica rapa (turnip rape). *J Heridity, 105,* 555–556

Harlan, J.R. & de Wet, J.M.J. (1972). A simplified classification of cultivated sorghum. *Crop Science, 12*, 172-176.

Hartl, D.L. and Clark, A.G. (1989). *Principles of Population Genetics*, (2nd edn.). Sunderland, Massachusetts: Sinauer and Associates.

Hedrick, P.W. (2000). *Using Evolutionary Genetic Principles in Conservation Biology Groningen*, Netherlands: Festshrift for Wilke Van Delden

Howell, E.C., Barker, G.C., Jones, G.H., Kearsey, M., J. & King, G.J. (2002). Integration of the cytogenetic and genetic linkage maps of *Brassica oleracea*. *Genetics, 161***,** 1225–1234.

IBPGR & ICRISAT, (1993). Descriptors for sorghum [*Sorghum bicolor* (L.) Moench]. In: International Board for Plant Genetic Resources Rome Italy and International Crops Research Institute for the Semi-Arid Tropics Patancheru India (Eds.)*:* IBPGR & ICRISAT.

ICRISAT, (1996). *The World Sorghum and Millet Economies: Facts, Trends and Outlook,* Patancheru, India/Rome, Italy: ICRISAT/FAO.

IPGRI, (1993). *Diversity and development.* The strategy of International Plant Genetic Resources Institute, Rome, Italy: IPGRI,

Iqbal, A., Sadia, B., Khan, A.I., Awan, F.S., Kainth, R.A. & Sadaqat, H.A. (2010). Biodiversity in the sorghum (*Sorghum bicolor* (L.) Moench germplasm ofPakistan. *Genetics and Molecular Research, 9,* 756-764.

Ito, M., Ichinose, Y., Kato, H., Shiraishi, T., & Yamada, T. (1997) Molecular evolution and functional relevance of the chalcone synthase genes of pea. *Mol Gen Genet*, *255*, 28–37.

Jaleel, C.A., Sankar, B., Sridharan, R. & Panneerselvam, R. (2008). Soil salinity alters growth, chlorophyll content and secondary metabolite accumulation in *Catharanthus roseus*. *Turk. J. Biol., 32,* 79-83.

Jamil, M., Lee, D.B., Jung, K.Y., Ashraf, M., Lee, S.C. & Rha, S.E., (2006). Effect of salt (NaCl) stress on germination and early seedling growth of four Vegetable species. *J. Central Eur. Agric., 7*(2), 273-282.

Jeannette, S., Carig, R. & Lynch, J.P. (2002). Salinity tolerance of *Phaseolus* species during germination and early seedling growth. *Crop Sci., 42*, 1584-1594.

Ji, G.S., Song, Y.F., Liu, G.Q., Du, R.H. & Hao, F.W. (2011). Genetic analysis of sorghum resources from China using SSRs. *Journal of SAT Agricultural Research, 9.*

Kaya, C., Tuna, A.L., Ashraf, M., & Altunlu, H. (2007). Improved salt tolerance of melón (*Cucumis melo* L.) by addition of proline and potassium nitrate. Environ. *Exp. Bot., 60*, 397-403.

Keshavarzi, M. H. B. (2011). Effect of Salt Stress on Germination and Early Seedling Growth of Savory (*Satureja hortensis*). *Australian Journal of Basic and Applied Sciences*, *5*(12), 3274-3279.

Kisua, J., Mwikamba, K., Makobe, M., & Muigai, A. (2015). Genetic diversity of sweet and grain sorghum populations using phenotypic markers. *IJB. 6*(9), 34-46.

Kong, L., Dong, J. & Hart, G.E. (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.

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

Krishnamurthy, L., Serraj, R., Hash, C.T., Abdullah, J., Dakheel, L., & Reddy, B.V.S. (2007). Screening sorghum genotypes for salinity tolerant biomass production. *Euphytica,* 156, 15–24.

Kumar, V., Sahrma, S., Kero, S., Sharma, S., Sharma, A.K., Kumar, M. & Bhat, K.V. (2008). Assessment of genetic diversity in common bean (*Phaseolus vulgaris*L.) germplasm using amplified fragment length polymorphism (AFLP). *Scientia Horticulture, 116*, 138-143.

Li, R., Zhang, H., Zhou, X., Guan, Y., Yao, F., Song, G., Wang, J. & Zhang, C. (2010). Genetic diversity in Chinese sorghum landraces revealed by chloroplast simple sequence repeats. *Genetic Resources and Crop Evolution, 57*, 1-15.

Liu, K. (2003). Power Marker: New Genetic Data Analysis Software, Version 3.25. Retrieved from; http://www.powermarker.net

Lokko, Y., Danquah, E.Y., Offei, S.K., Dixon, A.G.O. & Gedil, M.A. (2005). Molecular markers associated with a new source of resistance to the cassava mosaic disease. *African Journal of Biotechnology, 4*, 873-881.

Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer: Sinauer Associates, Inc.

Makobe, M.N., Kahangi, E.M., Misra, A.K., & Imbuga M.O. (2006). Development of Hardy Sorghum Cultivars for the Arid and Semi arid Regions. *African Crop Science Journal, 14*(4), 297-309.

Mead, R., Curnow, R.N., & Hasted, A.M. (2002). Statistical methods in agriculture and experimental biology (3rd ed.). New York: Chapman and Hall/CRC Press Company.

Mehmood, S., Bashir, A., Amad, A., & Akram, Z., (2008). Molecular characterization of regional *Sorghum bicolor* varieties from Pakistan. *Pak. J. Bot. 40*, 2015-2021.

Menz, M. A., Klein, R. R., Mullet, J. E., Obert, J. A., & Unruh, s., (2002). A high-density genetic map of *Sorghum bicolor* (L.)Moench based on 2926 AFLP, RFLP and SSR markers. *Plant Mol. Biol.48,* 483–499.

Moller, I.S., Gilliham, M., Jha, D., Mayo, G., Roy, S.J., Coates, J.C., Haseloff, J., & Tester, M., (2009). Shoot Na^+ exclusion and increased salinity tolerance engineered by cell type-specific alteration of Na^+ transport in *Arabidopsis. The Plant Cell; 21,* 2163-2178.

Mukuru, S.Z., (1993). Sorghum and millets in the Eastern Africa. Sorghum and millets commodity research environments, *ICRISAT,* 57-62.

Munns, R., James, R. A., & Läuchli, A. (2006). Approaches to increasing the salt tolerance of wheat and other cereals. *Journal of experimental botany*, *57*(5), 1025-1043.

Munsell Color Film, (1990). The Munsell book of color: Matte collection. New Windsor, New York: Munsell Color Film.

Murray, S.C., Rooney W.L., Hamblin, M.T., Mitchell, S.E. & Kresovich, S. (2009). Sweet sorghum genetic diversity and association mapping for brix and height. *Plant Genome, 2*(1), 48–62.

Mutegi, E., Sagnard, F., Semagn, K., Deu, M., Muraya, M., Kanyenji, B., …& Labuschagne, M. (2011). Genetic structure and relationships within and between cultivated and wild sorghum (*Sorghum bicolor* (L.) Moench) in Kenya as revealed by microsatellite markers. *Theoretical and Applied Genetics, 122,* 989-1004.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution, 29,* 1-10.

Netondo, G.W., Onyango, J.C., & Beck, E. (2004). Sorghum and salinity: I. Response of growth, water relations, and ion accumulation to NaCl salinity. *Crop Science, 44,* 797-805.

Ng'uni, D., Geleta, M., Hofvander, P., Fatih, M., & Bryngelsson, T., (2012). Comparative genetic diversity and nutritional quality variation among some important Southern African sorghum accessions [*Sorghum bicolor* (L.) Moench].*Crop science*, *AJCS*, *6*(1), 56-64.

Noreen, S. & Ashraf, M. (2008). Alleviation of adverse effects of salt stress on sunflower (*Helianthus annuus* L.) by exogenous application of salicylic acid: Growth and photo-synthesis. *Pak. J. Bot., 40*(4), 1657-1663.

Paterson, A.H., Bowers, J.E., & Chapman, B.A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proceedings of the National Academy of Sciences of the United States of America, vol.* 101, no. 26, pp. 9903–9908.

Perry, M.C. & McIntosh, M.S., (1991). Geographical patterns of variation in the USDA soybean germplasm collection. I. Morphological traits. *Crop Sci. 31*, 1350–1355.

Price, H. J., Dillon, S. L., Hodnett, G., Rooney, W. L., Ross, L. & Johnston, J. S. (2005). Genome evolution in the genus *Sorghum* (Poaceae). *Annals of Botany*, *95*(1), 219–227.

Ramesh, S., Reddy, B.V.S., Reddy, P.S., Hebbar, M. & Ibrahim, M. (2005). Response of selected sorghum lines to soil salinity-stress under field conditions. *International Sorghum and Millets Newsletter, 46*, 14–17.

Rao, S.A. & McNeilly, T. (1999). Genetic basis of variation for salt tolerance in maize (*Zea mays* L). *Euphytica.*, *108*, 145–50.

Reddy, B.V.S., Ramesh, S., Sanjana Reddy, P., & Kumar, A. (2009). Genetic enhancement for drought tolerance in sorghum. *Plant Breeding Reviews, **31***, 189–222.

Ren, H.Z., Goa, J.P., Li, L.G., Cai, X.L., Huang, W., Chao, D.Y., … & Lin, H.X. (2005). A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nature Genetics, 37*, 1141–1146.

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

Sankarapandian, R., Krishnadoss, D., & Devarathinam, A.A., (1996). Genetic parameters, correlations and path analysis among yield and yield characters in grain sorghum. *Journal of Madras Agriculture, 83*, 625-628.

Schloss, S.J., Mitchell, S.E., White, G.M., Kukatla, R, Bowers, J.E., Paterson, A.H. & Kresovich, S., (2002). Characterization of RFLP clone sequences for gene discovery and SSR development in *Sorghum bicolour* (L).Moench. *Theor. Applied Genet, 105*, 912-920.

Schnable, P.S., Ware, D., & Fulton, R.S. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science, 326,* 1112–1115.

Schober, T.J., Messerschmidt, M., Bean, S.R., Park, S.-H., & Arendt, E.K. (2005). Gluten-free bread from sorghum: quality differences among hybrids. *Cereal Chemistry, 82*, 394–404.

Shannon, MC. (1997). Adaptation of plants to salinity. *Advances in Agronomy 60*, 75–120.

Shewayrga, H., Jordan, D.R. & Godwin, D. (2006). Genetic erosion and changes in distribution of sorghum (*Sorghum bicolour* (L.) Moench) landraces in northeastern Ethiopia. *Plant Genetic Resources: characterisation and utilization, 6*, 1-10.

Shiringani, A. L., Frisch, M. & Friedt, W. (2010). Genetic mapping of QTLs for sugar-related traits in a RIL population of Sorghum bicolor L. Moench. *Theor. Appl. Genet. 121*, 323–336.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science, 236*, 787-792.

Slewinski, T.L. (2012). Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. *J. Exp. Bot*., *63*, 4647–4670.

Smith, J.S.C., Kresovich, S., Hopkins, M.S., Mitchell. S.E., Dean. RE., Woodman, W.L., Lee. M., & Porter, K. (1987). Genetic diversity among elite sorghum inbred lines assessed with Simple sequence repeats. *Crop Sci*., *40,* 226-232.

Sneath, P.H.A. & Sokal, R.R. (1973). *Numerical Taxonomy*. London: Freeman.

Sobral, B.W.S., Braga, D.P.V., LaHood, E.S. & Keim, P. (1994). Phylogenetic analysis of chloroplast restriction enzyme site mutations in the *Saccharinae* griseb. subtribe of the *Andropogoneae* Dumort. Tribe. *Theoretical and Applied Genetics*, *87*(7), 843–853.

Stenhouse, J.W., Prasada Rao, K.E., Gopal Reddy, V. & Appa Rao, S. (1997). Sorghum. In: Fuccillo, D.*, et al.* (Eds.) *Biodiversity in Trust, Conservation and Use of Plant Genetic Resources in CGIAR Centres*. pp. 292-308. Cambridge: Cambridge University Press.

Stickler, F.C., Weaden, S., & Pauli, A.W. (1961). Leaf area determination in grain sorghum. *Journal of Agronomy, 53*, 187-188.

Tazeen, M., Nadia, K., & Farzana, N.N. (2009). Heritability, phenotypic correlation and path coefficient studies for some agronomic characters in synthetic elite lines of wheat. *J. Food. Agric. Environ. 7*(3&4), 278 - 282.

Tester, M., & Davenport, R. (2003). Na+ tolerance and Na+ transport in higher plants, *Ann. Bot. 91*, 503-527.

Tóth, G., Gáspari, Z., & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research, 10*, 967-981.

Ulanch, P.E., Childs, K.L., Morgan, P.W. & Mullet, J.E. (1996). Molecular markers linked to Ma (1) in sorghum. *Plant Physiology*, *111*, 709.

Uozumi, N., Kim, E.J, Rubio, F., Yamaguchi, T., & Muto, S. (2000). The *Arabidopsis* HKT1 gene homolog mediates inward Na+ currents in *Xenopus laevis* oocytes and Na+ uptake in *Saccharomyces cerevisiae*. *Plant Physiol, 122*, 1249–59.

Volkov, V. & Amtmann, A. (2006). *Thellungiella halophila*, a salt-tolerant relative of *Arabidopsis thaliana*, has specific root ion-channel features supporting K+/Na+ homeostasis undersalinity stress. *Plant J.* 48, 342–53.

Weimberg, R., Lerner, H.R., & Poljakoff-Mayber, A. (1984). Changes in growth and water soluble solute concentrations in *Sorghum bicolor* stressed with sodium and potassium, *Physiol. Plant*. *62*, 472–480.

Weir, B.S. (1996). Intraspecific differentiation P. 385-403. in D.M. Hillis *et al.* (ed). *Molecular systematics* (2nd ed.). M.A: sunderlands.

Willis, J. H. (1996). Measures of phenotypic selection are biased by partial inbreeding. *Evolution, 50,* 1501–1511.

World Hunger Education Service, (2012). World Hunger Organization. *2012*. Retrieved from: www .worldhunger.org

Yang, M., Liu, L., Li, L., Wang, B., Chang, J., & Liu, G. (2009). Correlation analysis between sucrose synthaseexpression and sucrose accumulation in sweet sorghum [*Sorghum bicolor* (L.)Moench]. *Acta, Agronomic Sinca 35*(1),185–189.

Yao, X., Horie,T., Xue,S., Leung, H.Y., Katsuhara, M., Brodsky, D.E., Wu, Y., & Schroeder, J.I. (2010). Differential sodium and potassium transport selectivities of the rice OsHKT2; 1 and OsHKT2; 2 transporters in plant cells. *Plant Physiology,* 152*,* 341-355.

Yeh, F.C., & Boyle, T.J.B. (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany 129,* 157.

Yilmaz, K., Akinci, I.E., & Akinci, S. (2004). Effect of salt stress on growth and Na, K contents of pepper (*Capsicum annuum* L.) in germination and seedling stages. *Pak. J. Biol. Sci., 7*, 606-610.

Zhang, J.L., Flowers T.J., & Wang, S.M. (2010). Mechanisms of sodium uptake by roots of higher plants. *Plant Soil, 326*, 45-60.

Zheng, L.Y., Guo, X.S., He, B., Sun, L.Y, Peng, Y., Dong, S.S., … & Jing, H.C. (2011). Genome data from sweet and grain sorghum (Sorghum bicolour).*GigaScience* .