ASSOCIATION MAPPING FOR DROUGHT TOLERANCE IN SORGHUM [*Sorghum bicolor (L.)* Moench]

M.Sc. THESIS

Submitted to the School of Graduate Studies,
Jimma University College of Agriculture and Veterinary Medicine

In Partial Fulfillment of the Requirements for the **Degree of Master of Science in Plant Biotechnology**

BY

ALEYE ENDRE ABUSSIE

JULY, 2012

JIMMA UNIVERSITY

APPROVAL SHEET

Jimma University College of Agriculture and Veterinary Medicine Graduate Studies

As thesis research advisor, we hereby certify that we have read and evaluated the thesis prepared by **Aleye Endre Abussie** entitled **Association Mapping for Drought Tolerance in Sorghum [*Sorghum bicolor* (L.) Moench]**. We recommend that it can be submitted as fulfilling thesis requirement.

Kassahun Bantte (Ph.D)	_____	_____
Major Advisor	Signature	Date
_____	_____	_____
Co-Advisor	Signature	Date

As member of Board of Examiners of the M.Sc. Thesis Open Defense Examination, We certify that we have read and evaluated the thesis prepared by **Aleye Endre Abussie** entitled **Association mapping for drought tolerance in sorghum [*Sorghum bicolor* (L.) Moench].** We recommend that it be accepted as fulfilling the thesis requirement for the **Degree of Master of Science in Plant Biotechnology**.

_____	_____	_____
Chair man	Signature	Date
_____	_____	_____
Internal Examiner	Signature	Date
_____	_____	_____
External Examiner	Signature	Date

DEDICATION

This thesis is dedicated to my mother, Zinetu Beshir Yimer, who did a lot for my educational endeavor but not alive today.

STATEMENT OF AUTHOR

I certify that this thesis and the research to which it refers is the product of my own work and that any ideas or quotation from the work of other people, published or otherwise are fully acknowledged in accordance with the standard referring practices of the discipline.

This thesis has been submitted in partial fulfillment of the requirements for the M.Sc. degree at Jimma University, College of Agriculture and Veterinary Medicine and is deposited at the University Library to be made available to borrowers under rules and regulation of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Director of the School of Graduate Studies when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Name: Aleye Endre Abussie
Signature: _____
Place: Jimma University
Date of Defense: **July 27/2012**
Date of Submission: August **05/2012**

BIOGRAPHICAL SKETCH

The author, Aleye Endre Abussie, was born on April 13, 1979 in Amhara Regional State of South Wollo Zone, Woreilu. He attended his elementary and junior secondary school at Kabie Elementary and Junior School. Then, he joined Woreilu Senior Secondary School and completed his secondary education in 1996. Right after the completion of the secondary education, he joined Kotebe College of Teacher's Education (KCTE) in September, 1996 and completed two years diploma program with biology major and chemistry minor in 1998.

After his graduation, he was employed by Ministry of Education and served the organization for five years in Afar Regional State. Then, he joined Jimma University in 2003 and graduated with B.Ed degree in Biology in September, 2007. Thereafter, he continued to serve the organization at Amibara District of Afar Regional State

Finally, after serving for a total of two years with Bachelor of degree (B.Ed.), he joined the School of Graduate studies in Jimma University College of Agriculture and Veterinary Medicine in October, 2009 to pursue his M.Sc. in Plant Biotechnology

ACKNOWLEDGMENTS

It is, indeed, a great pleasure and a feeling of achievement to write the acknowledgements for those who have made significant contributions to the successful completion of this work. I am deeply grateful to Dr. Kassahun Bantte (Assoc.prof.), my major advisor, for giving me an opportunity to work on this fascinating research topic and for his support, insightful critiques and useful suggestions, throughout the course of this study.

I am grateful to Jimma University College of Agriculture and Veterinary Medicine (JUCAVM), my intellectual home, for facilitating partly my research work and financial support. The financial support granted by the national research fund (NARF) is greatly acknowledged, as well. Staff members of Sirinka Agricultural Research Center (SARC), particularly those at Kobo site are appreciated for their cooperation during the field work.

There are many who helped me to arrive at this day. I am very grateful to Haile Yineger (now a PhD student in Australia) who motivated me to join this postgraduate study and to keep myself on the right track. I am grateful to my friend, Yonas Moges for his constructive ideas, and motivation. My special thanks is to Yilkal Beze, who devoted his time making me improve my computer skills, while he was a senior post graduate student at JUCAVM. I am grateful to Dr. Tesfahun Alemu (PhD) for his comments and ideas which were useful for the betterment of my work.

Throughout my entire academic career including the elementary, high school and undergraduate studies that paved the way for today's success, my parents, particularly my father has been a constant source of motivation. To him, I am deeply grateful. I am also thankful to my younger brother, Beshir Endre, for his consistent moral support and encouragement.

Finally, I would be remiss if I did not praise my Almighty God; the everlasting father, and the prince of peace and love for giving me the endurance and patience in accomplishing this piece of work. Amen!!!

ACRONYMS AND ABBREVIATIONS

CTAB	Cetyltrimethyl-ammonium bromide
CSSRI	Central Soil Salinity Research Institute in India
D'	The standardized linkage disequilibrium coefficient
EDTA	Ethylene di amine tetraacetic acid
FAOSTAT	Food and Agriculture Organization *(United Nations)* Statistics Division
GAS	Genome assisted selection
GLM	General linear model
IBC	Institute of Biodiversity Conservation
LD	Linkage disequilibrium
MCMC	Markov chain Monte Carlo algorithm
MLM	Mixed linear model
MAS	Marker assisted selection
QTL	Quantitative trait loci
R^2	The square of the correlation coefficient between alleles at two loci
SAT	Semi arid tropical regions
SSR	Simple sequence repeats
UN	United nation

Contents **TABLE OF CONTENTS** **page**

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF TABLES IN THE APPENDIX

LIST OF FIGURES IN THE APPENDIX

ASSOCIATION MAPPING FOR DROUGHT TOLERANCE IN SORGHUM [*Sorghum bicolor* (L.) Moench]

ABSTRACT

Sorghum [Sorghum bicolor {L.} (Moench] is the third most important cereal crop in Ethiopia next to maize and tef. Drought is the major sorghum production constraint in Ethiopia which requires identification of quantitative trait loci (QTLs) responsible for drought tolerance and thereby development of drought tolerant varieties. The objectives of this study were to identify drought tolerant genotypes, to map chromosomal regions (QTLs) associated with agronomically important traits including drought tolerance and identify SSR markers tightly linked with these QTLs. For this study, one hundred sixty sorghum genotypes (152 land races and 8 released varieties) were genotyped with 39 SSR and evaluated in field at Kobo in the off-season using an alpha lattice design replicated three times. Phenotypic data were collected including days to 50% flowering, plant height, panicle weight, grain weight, grain weight per panicle, panicle harvest index, one thousand grain weight and number of grains per panicle. Analysis of variance showed highly significant ($P<0.0001$) differences among the genotypes for all characters. Grain weight per panicle ranged from 8.68 (210902) to 101.09 (210922) with an average value of 43.52. The first 17 best performers among the entries were the landraces indicating the presence of better performing genotypes in the land races than the existing released varieties and the five best performing genotypes were 210922, 73067, 69183, 69231, and 69241 in that order. Most of the characters showed moderate to high phenotypic and genotypic coefficient of variation. Heritability was high for all of the studied characters. High heritability coupled with high genetic advance as percent of mean was observed for plant height, panicle weight, grain weight per panicle, one thousand grain weight, and number of grains per panicle. Linkage disequilibrium (LD) analysis indicated that in all accessions, 107 loci pairs (32.92%) had a significant ($p< 0.05$) mean LD of 0.19, with an $R^2 > 0.2$ for 33 evaluated loci pairs. Population structure analysis showed that there were four distinct clusters in the studied materials. A total of 10 marker-trait associations were identified using 7 different SSR markers. The percentage of the total variation explained by the markers ranged from 2.6 % (Xtxp114 with THGT) to 17.76 % (Xtxp145 with PHT). The seven SSR markers were localized on chromosomes 1, 2, 3,5,6,7, and 8 harboring one marker each (xcup53, bSbCIR223, Xtxp114, mSbCIR248, Xtxp145, Xtxp278, and gbsp123 respectively). Most of the identified markers were localized with those previously identified as linked to drought tolerance-related traits using conventional QTL mapping supporting the reliability of the present findings. The results of this study can serve as initial effort for the association mapping studies in sorghum particularly in our country as the associated SSR markers are potential candidates for marker-assisted selection to improve drought tolerance in sorghum. However, as this study is the first attempt in the identification of QTLs for drought tolerance using association mapping, the identified QTLs need to be validated in independent or related populations and in different environments before their use in marker-assisted selection.

***Key words**: Association mapping, Drought tolerance, Population structure, Sorghum, SSRs,*

1. INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is a monocot crop in the grass family of *Poaceae*. It is largely self- pollinating (70%-95%) with a diploid set of chromosomes (2n=2x=20) and an estimated genome size of 750 Mb (Doggett, 1988; Yonemaru *et al.*, 2009). Sorghum is an important failsafe crop in the global agro ecosystem as it gives reasonable yield in the face of moisture stress which is not achieved by most other grains (Paterson, 2008). It belongs to the group of C_4 plants with high rates of photosynthesis and efficient use of water and nitrogen resources (Matsuoka *et al*, 2001; Lopes *et al*, 2011).

Sorghum is grown in 105 countries of Africa, Asia, Oceania and the Americas though the largest share (> 70%) of global sorghum area is confined to Africa and India (Kumar *et al*., 2011). It is a staple food for more than 500 million people in 30 countries (Kumar *et al*., 2011). At a global scale, sorghum is the fifth most important cereal crop grown worldwide on the basis of both production and area coverage, after maize, wheat, rice and barley (FAOSTAT, 2010). It is the second major crop (after maize) in almost all agro ecologies in Africa (Bantilan *et al*, 2004). In Ethiopia, sorghum ranks third after maize and tef in total national production (Asfaw, 2007).

Worldwide, 44 million ha of land was allotted for sorghum production in 2009, and the output was 62 million metric tons with an average yield of 1419 kg/ha (FAOSTAT, 2010). The area coverage in Africa was 27.7 million ha in the same year with an output of 27 million metric tons and an average yield of 977.3 kg/ha. In Ethiopia in the same year, 1.6 million ha of land was allotted for sorghum production and 2.8 million metric tons of grain was produced in the country with an average yield of 1,736 kg/ha (FAOSTAT, 2010).

More than 35 percent of world's sorghum production is dedicated to human consumption of which 95% is in developing countries; the rest being used mainly for animal feed, alcohol and industrial products (Dicko *et al*, 2006). Sorghum grain is used to make unfermented and fermented breads, couscous, cakes, muffins, cookies, biscuits, and, alcoholic and non alcoholic beverages (Badi *et al*., 1990). *Injera* and porridge are the popular foods made from sorghum in Ethiopia including local drinks such as *Tella*. Sorghum consists of different kinds of useful nutrients though its major components are

carbohydrates. Dicko *et al*. (2006) reported that sorghum grain contains 65-80% carbohydrates, mainly starch, followed by proteins (7-15%), and fat (1.5-5 %). Sorghum is a good source of minerals such as K, Mg, Fe, Zn, Cu, Ca and P (Hulse *et al*., 1980; Smith and Frederiksen, 2000). It is also rich in vitamins, especially the B vitamins (thiamin, riboflavin, pyridoxine), and the liposoluble vitamins A, D, E and K (Dicko *et al*., 2006). One of the special features of sorghum is that it is inherently gluten-free and is demonstrated to be safe for people with celiac disease (Ciacci *et al*., 2007). As gluten is a component of wheat, oats, barley and rye, sorghum is considered as the best substitute of these crops for people allergic to gluten (Fenster, 2003; Ciacci *et al*., 2007). Moreover, the plant stem and foliage are used for green chop, hay, silage, pasture, source of fuel for cooking, construction of fences and hut making.

Despite its importance, sorghum productivity and food production worldwide is severely limited by drought (House, 1985). Currently, drought is considered, by the Food and Agriculture Organization of the UN, to be the single most common cause of severe food shortage accounting for more than 50% yield loses each year globally (Boyer, 1982; Hao *et a*l., 2011). This loss is estimated to be around USD 500 million per year (Sharma and Lavanya, 2002). The severity increases particularly in developing countries like Ethiopia, where the majority of the people depend on agriculture for their livelihood (Hao *et a*l., 2011). In Ethiopia, sorghum is largely cultivated in moisture stress areas that cover nearly 66% of the total area of the country (Tadesse *et al*., 2008). Rainfall variability and associated droughts are the major causes of food shortages and famines in the country (Woldeamlak, 2009). For example, Senay and Verdin (2002) noted that there was 30% yield loss in 1997 as compared to 1996 cropping season though there was a potential to produce 3.5 tons/ha on farmers' fields in major sorghum growing regions of the country (Geremew *et al*., 2004). This calls for attention and attracts interest to improve efficiency in sorghum production through the development of tolerant crop varieties.

Although conventional breeding in the past has played a great role in developing abiotic stress tolerant varieties in different cereal crops, the selection and testing process takes very long time (Farook and Azam, 2002). Particularly, in quantitatively inherited traits such as drought tolerance that are controlled by many genes each with small effects (QTLs), selection using conventional methods has been neither efficient nor reliable due to

the complex nature of the trait and the complicating effects of the environment (Ejeta and Knoll, 2007).

However, after the invention of molecular markers around 1980s, the limitation of conventional breeding has been compensated by employing DNA markers that are linked to target genes, as bench marks, for selecting the genetic factors that are responsible for the observed variability in quantitative traits (Agrarwissenschaften, 2008). Genome assisted selection (GAS) using molecular markers has been thought to be efficient to manipulate polygenic traits such as drought tolerance (Haussmann *et al,* 2004). Molecular (DNA) markers can be ordered on chromosomes and the entire genome of an individual can be visualized on a genetic linkage map (Varshney and Tuberose, 2007). Using the marker map, genes affecting traits of interest can be detected by testing for statistical associations between marker variants (alleles) and any trait of interest.

However, QTL mapping is a pre requisite for molecular markers to be used as selection criteria for crop improvements (Pennisi, 2008). QTL mapping can be performed by conventional QTL mapping or Association mapping strategy. In the past, using conventional QTL mapping, a large number of genes for various traits (quality traits, resistance to biotic stresses, etc.) have been tagged with markers in sorghum (Agrarwissenschaften, 2008). For example, Xu *et al.* (2000) mapped QTL controlling chlorophyll content in sorghum using a mapping population of 98 RILs derived from the cross B35 xTRx7000 and identified three QTLs for leaf chlorophyll content (ch11, ch12, and ch13) that together explained 25-30% of the observed phenotypic variability.

Reddy *et al*. (2007) identified QTLs associated with stalk rot resistance in sorghum using 93 RILs genotyped with 62 nuclear and 4 genic SSR and 19 RAPD markers. They identified five QTLs on linkage groups A, B, D and I.

Srinivas *et al*. (2009) reported that 49 QTLs were mapped controlling for 12 agronomic traits (plant height, days to anthesis, grain yield, panicle weight, seed weight, etc.) using 149 microsatellites (100 genomic- and 49 genic-microsatellites) and three morphological markers.

However, in conventional QTL mapping, one must grow the plants at least for three years, very large segregating populations are required to achieve high resolution mapping (Tanksley, 1993), and only two alleles at any particular locus can be assessed which limit the utilization of conventional QTL mapping.

In order to overcome these limitations, association mapping has recently been used to map QTLs associated with desired traits, with no need of mapping population development but with a higher level of resolution (Hamblin *et al*, 2005). Recently, association mapping has been reported in crop plants including mapping of flowering time in maize (Thornsberry *et al*. 2001), yield traits in barley (Kraakman *et al.,* 2004), iron deficiency in soybean (Wang *et al*.2008), disease resistance in rice (Gaaris *et al.*2003), potato (Simko *et al*.2004), and maize (Szalma *et a*l., 2005).

In sorghum, Hamblin *et al*. (2005) reported that there is short and medium range (up to100 kb) of linkage disequilibrium (LD) patterns and this makes the crop well suited to association mapping methodologies. They stated that the extent of allelic associations in *S. bicolor*, as assessed by pair wise measures of LD, is higher than in maize but lower than in *Arabidopsis*, in qualitative agreement with expectations based on mating system. Their study suggested that *S. bicolor* is well suited for association studies, since LD typically extends at least several kilo bases but has largely decayed by 15 kb.

Casa *et a*l. (2008) compared different models of association mapping in a panel of 377 sorghum accessions using 47 simple sequence repeat (SSR) loci. The phenotypic and genotypic data were used to evaluate the performance of several association models in controlling for spurious associations. The result of their analysis indicated that association models that accounted for both population structure and kinship performed better than those that did not.

Shehzad *et al*. (2009) conducted genome wide association mapping study with different association models using 107 representative sorghum accessions genotyped with 98 SSR markers and declared the feasibility and resolution of association mapping study in sorghum.

In association mapping the existing germplasm (land races) can be used as mapping population (Sorkheh *et al*., 2008). The possibility of using the existing germplasm resources for gene and allele discovery on the basis of association mapping strategies is of particular interest to sorghum breeders. Ethiopia, where sorghum is believed to have been first domesticated and where the greatest genetic variation for both cultivated and wild sorghum is found, is a rich source of sorghum landraces which are valuable sources of desirable genes such as for drought tolerance (Amsalu *et al*., 2000). However, no study to detect marker-trait association for drought tolerance in sorghum has been previously reported using association mapping in Ethiopia. Therefore, this study was initiated with the following objectives:

Objectives

- To map chromosomal regions associated with drought tolerant and/or agronomically important traits using association mapping
- To identify drought tolerant sorghum genotypes
- To identify SSR markers that are associated with drought tolerant traits

2. LITERATURE REVIEW

2.1. Biology, Origin, and Domestication of Cultivated Sorghum

The genus sorghum belongs to the *Poaceae* (*Gramineae*) family and falls into the tribe *Andropogoneae* along with maize (*Zea mays*), sugarcane (*Saccharum* spp.), and all the millets (*Pennisetum*, *Eleusine*, *Eragrostis*, *Setaria,* etc.) (Paterson *et al*., 1998). This genus (Sorghum) is divided into five sub-genera as *Parasorghum*, *Heterosorghum*, *Chaetosorghum, Stiposorghum and* Sorghum *(Eu Sorghum)* (Paterson *et al*., 1998).

The sub-genera *Eu Sorghum* is further divided in to three species including (1) *Sorghum halepense* (L.) Pers., a rhizomatous, perennial tetra ploid (n=20); (2) *S. propinquum* (Kunth.) Hitchc.; a strongly rhizomatous, perennial diploid (n=10); and (3) *S. bicolor* (L.) Moench, an annual diploid (n=10) (Paterson *et al*., 1998). *Sorghum bicolor* also further grouped in to three sub species: *bicolor* (domesticated grain sorghums), *arundinaceum* (wild progenitors of grain sorghums), and *Drummondi* (derivatives of hybridization between grain sorghum and its closest wild relatives). The ssp. *bicolor* is further divided in to five basic races (*bicolor, caudatum, dura, ginea and kafir*) and ten hybrid (intermediate) races that are two-by-two combinations of the basic races yielding a total of 15 races (Doggett, 1988; Amsalu, 2001).

The origin and early domestication of sorghum is hypothesized to have taken place around 5000 years ago in north east quadrant of Africa in an area including Ethiopia, Sudan and Chad (Mann *et al*. 1983; Doggett, 1988). The cultivated sorghum of today is believed to have been arisen from the wild *sorghum bicolor* sub species a*rundinaceum* (Doggett, 1988). There is no evidence suggesting that presently cultivated sorghums have evolved from the rhizomatous diploid or tetraploid wild species (Doggett, 1988). According to Amsalu (2001), however, the naming of the spontaneous sub species *arundinacium* changed in to *s.bicolor* sub species *verticiliflorum* by De Wet and prasada (1986) and its consistent use is recommended by Doggett (1988). According to Harlan and De Wet (1976), Mann *et al.* (1983), and Doggett (1988), the north east quadrant of Africa, particularly Ethiopia is the center of origin of sorghum. This area is rich in the number of snowdenian species and also contains several varieties of the durra type, which represent

the highly evolved varieties among the cultivated races (Harlan and Dewet, 1976). The great genetic variability of the crop in this geographical area, the wide range of ecological habitats, and the long history of human selection efforts in the region have given sufficient credence to the theory of the origin and early domestication of sorghum in the north east quadrant of Africa mainly in Ethiopian region (Kole, 2006).

Ones grew in the wild before it was domesticated as food and feed and cultivated in Africa for several years, several routes have been recognized for the later movement of sorghum into other parts of Africa and beyond. It is believed that the early movement between 2,000 and 4,000 years ago has taken the crop to west, central, and southern Africa leading to further domestication and appearance of distinct forms (races) in each of these regions (Doggett, 1976). The crop has also moved to other continents through different routes.

Of the five basic races of sorghum, *bicolors* are thought to have been domesticated first and they are now grown across the range of sorghum cultivation in Africa with nonspecific center of diversity (Harlan and De wet, 1972). According to Harlan and De wet (1972), race c*audatum* is widely grown in present Chad, Sudan, northeastern Nigeria and Uganda as its center of diversity; the race guinea, while occasionally cultivated in several places around the world, is uniquely predominant in the West African savannah; the durra race which is dominant in Ethiopia, is also widely grown along the fringes of the southern Sahara, across arid west-Africa; whereas race kafir is distinctly eastern and southern African from Tanzania southward.

Beyond Africa, early domestication of sorghum has been thought to occur in the Indian sub-continent about 3,000 years ago (Doggett, 1976). Later, it is believed that the crop was taken from India to China over the Himalayas where new races of sorghum like Chinese kaoliangs occurred which are considered as the only sorghums in the world that have evolved in the temperate regions and have become an important source of early season cold tolerance in sorghum (Doggett, 1976; Rao, 2009). More recently, around the mid 19th century, the crop was moved from West Africa to the USA and other parts of the world through slave trade (Dogett, 1988). Today, cultivated sorghum is one of the most economically important crops of the world with an annual production of 62 million metric tons (FAOSTAT, 2010).

2.2. Ecological Adaptation

Sorghum is grown in a wide range of latitudes covering the areas between 40°N and 40°S from the equator (Amsalu, 2001; Rao *et al*, 2009) being subjected to a wide variety of temperatures, day- length, and moisture regimes which actually make the crop best suited for breeders to develop improved cultivars that can smoothly fit to any specific environmental condition (Singh *et al*., 1997).

Though sorghum is mostly found at elevations between sea level and 1,500 m, most East African sorghum is grown between the altitudes of 900 to 1,500 m; whereas cold tolerant varieties are grown even between 1,600 and 2,500 m in China (Rao *et al.*, 2009) and up to 3000 m in Ethiopia (Teshome *et al*.,2007).

Grain sorghum is grown successfully in all types of soils with pH ranging from 5.0 to 8.5, but fertile and well drained soils are important to optimize grain yield (House, 1985). Sorghum requires an optimum temperature of about 15 °C for good seed germination and 25 to 30 °C for optimum growth and development (House, 1985; Singh *et al*., 1997). The temperature can, however, be as low as 20 °C, without a dramatic effect on growth and yield. It can with stand maximum temperature up to 37 °C (House, 1985).

Sorghum requires less moisture for growth than some other cereal crops such as Maize, Barley and Wheat (House, 1985). According to House (1985), Sorghum requires 332 kg water per kg of accumulated dry matter; Maize requires 368 kg of water; Barley 434 kg; and Wheat 514 kg. Sorghums become dormant in the absence of adequate water but do not wilt readily and are more efficient than maize in utilizing phosphorus and potassium (Smith and Frederiksen, 2000).

Although many of its varieties are insensitive to photoperiod, sorghum is basically a short day plant. In average, sorghum requires 90 to 140 days to reach maturity (House, 1985).

2.3. Production Constraints

The major factors that affect sorghum production can be grouped in to biotic and abiotic factors. The biotic factors include disease attacks (e.g. charcoal rot, leaf blight, grain molds e.t.c.), insect pests (e.g. shoot fly, sorghum midge, stem borers e.t.c.), and parasitic weeds (e.g. Striga). The major abiotic factor that affects sorghum production is drought (Hao *et al*., 2011).

2.3.1. Drought

Drought can be defined as a deficiency of precipitation of sufficient magnitude for a time period long enough to deplete soil water causing injury to plants (Decker, 1983). Although drought, salinity, and flooding are the major abiotic stresses that limit crop production all over the world, drought is by far the leading environmental stress to crop production in general and to sorghum production in particular (Anjum *et al*., 2011). This might be due to the nature of the environments in which sorghum is commonly grown, where the area is less humid, the temperature is warmer and the rain fall is often low and erratic (Ejeta and Knoll, 2007).

Water stress affects several aspects of smooth functioning in plants including plant's water and nutrients absorption, seed germination, opening and closing of stomata, photo synthetic activity, transpiration rate, enzymatic activity and several other metabolic and physiologic processes leading to reduction of plant size, mass and seed yield (Anjum *et al.*, 2011). One of the impacts of drought is the impairment of nutrient absorption and uptake by plants (Anjum *et al*., 2011). Drought and/or heat stress can decrease nutrient availability including plant nitrogen uptake, and activities of nitrogen assimilatory enzymes (Prasad and Staggenborg, 2008).

Although nutrient and water absorption processes are independent processes, the need of water for absorption and transport makes them highly dependent on each other. Most nutrients are absorbed by plant roots as ions and water is the medium of transport (Anjum *et al*., 2011). Under water stress, roots are unable to take up nutrients from the soil because

of lack of activity of fine roots, water movement, and ionic diffusion of nutrients (Prasad and Staggenborg, 2008).

Water deficit severely hampers plant growth and development. The first and foremost effect of drought is impaired-germination and poor-stand establishment (Prasad and Staggenborg, 2008). Growth is the result of daughter-cell production by meristematic cell divisions and subsequent massive expansion of the young cells. Under severe water deficiency, cell elongation of higher plants can be inhibited by interruption of water flow from the xylem to the surrounding elongating cells (Samarah *et al.,* 2009). Prasad and Staggenborg (2008) pointed out that continued water deficit reduces initiation of new leaves, individual leaf size, the number of leaves per plant, leaf area expansion, and accelerate leaf senescence and lead to death of leaf tissue.

Also, by the intensity and duration of stress, drought influences both the duration of the developmental stage and the transition of one developmental stage to another (Farooq *et al.*, 2009). Drought stress during panicle development inhibits the conversion of vegetative to reproductive phase and plants remain vegetative until the stress is relieved. Craufurd *et al*.(1993) found that panicle initiation in sorghum was delayed by as many as 2 to 25 days and flowering by 1 to 59 days under drought stress, with more severe effects when drought was imposed both at early and late stage of panicle development. Drought and heat stress not only can delay the panicle initiation but also can cause the cessation of panicle development at any stages between panicle initiation and flowering (Prasad and Staggenborg, 2008).

Effect of drought on biomass production

Although in most cases the tolerant genotypes had less reduction in biomass than the susceptible ones in response to drought, the general adverse effect of water stress on crop plants is reflected in the reduction of fresh and/or dry biomass production depending on the developmental stage of the plant and the severity of the stress (Samarah *et al*., 2009). Khan *et al*. (2001) conducted a study comprising of six treatments (six irrigations) namely, control, five, four, three, two and one irrigation in maize and concluded that plant height, stem diameter and leaf area decreased noticeably with increasing water stress. Similarly, Kamara *et al*. (2003) reported that water deficit imposed at various developmental stages

of maize reduced total biomass accumulation at silking by 37%, at grain-filling period by 34% and at maturity by 21%.

Also, water stress can reduce not only biomass production but also accompanied with shifts in photo - assimilates-partitioning within the plants and usually promoting an increased biomass partitioning to below ground organs in contrast to the case when water is available during which greater biomass allocation is promoted to the shoot and leaves thereby increasing the above ground to below ground dry matter ratio (Prasad and Staggenborg, 2008). Kirnak (2001) has justified the reason why root to shoot ratio was 2:1 in water stressed plants (e.g. eggplants) which shows the influence of water stress on the pattern of dry matter distribution favoring roots.

Effect of drought on yield and yield components

Drought stress affects yield and yield components largely through its inhibitory effects on leaf expansion and leaf development resulting in reduced light interception. Drought also initiates stomatal closure in response to low soil water content, which decreases the intake of CO_2 thereby decreasing the rate of photosynthesis which culminates in the reduction of yield and yield components, and may also cause death to the plant if the severity increases (Flexas *et al*., 2004). Rashidi and Seyfi (2007) reported that water stress significantly affected crop yield, grain number, grain weight, number of fruits per plant and fruit weight. Drought at flowering reduces plant growth and development leading to hampered flower production and grain filling and thus smaller and fewer grains and most of the time results in bareness (Prasad and Staggenborg, 2008). A reduction in grain filling occurs due to a reduction in the assimilate partitioning and decreased activities of sucrose and starch synthesis enzymes (Flexas *et al*., 2004).

Manjarrezsandoval *et al*. (1989) reported that grain yield reductions of l7, 34 and l0% were found in sorghum (*Sorghum bicolor* (L.) Moench) when water deficiency occurred before the boot stage, from the boot stage to anthesis, and from the milk dough stage to the soft dough stage, respectively. Inuyama (1978) reported a 61% reduction in grain yield when drought occurred at the boot stage in sorghum. Anjum *et al.* (2011) reported that when maize plants were exposed to drought stress at teaseling stage, it led to substantial

reduction in yield and yield components such as kernel rows/cob, kernel number/row, 100 kernels weight, kernels/cob, grain yield/plant, biological yield/plant and harvest index.

Reductions of sorghum grain yield due to drought stress before anthesis are related to decreases in grain number, while a smaller grain size is responsible for yield losses when water deficits occur after anthesis (Manjarrezsandoval *et al*., 1989). As seed size is the final component of yield, drought stress later during the reproductive development (after fertilization) decreases seed size rather than seed number (Prasad and Staggenborg, 2008). Sinaki *et al*. (2007) stated that water stress during seed development did not have an effect on the sink size (seeds per plant) rather decreased source capacity led to the reduction of seed weight.

2.4. Mechanisms of Drought Tolerance in Plants

Drought tolerance in crops is the ability to grow, flower and display economic yield with minimum loss in a water deficit environment relative to a water-constraint free production environment (Farooq *et al.,* 2009). To improve crop productivity, it is necessary to understand the mechanisms of plant responses to drought conditions with the ultimate goal of improving crop performance in areas where rainfall is limiting (Tuinstra *et al*., 1996). Plants have developed numerous strategies to control water status and to survive under drought which include escape, avoidance and tolerance strategies (Schulze 1986).

2.4.1. Drought escape

Drought escape is the ability of a plant to complete its developmental stage before serious soil and plant water deficits develop (Xu *et al*., 2005). Plants that escape drought exhibit a rapid phenological development and high degree of developmental plasticity, being able to complete their life cycle before physiological water deficit occurs (Xu *et al*., 2005). Escape strategies rely on successful reproduction before the severe stress is perceived.

In environments with terminal drought stress and where physical or chemical barriers inhibit root growth, drought escape through early flowering and/or short growth duration is advantageous (Farooq *et al.,* 2009). On the other hand, later flowering can be beneficial in escaping early season drought if drought is followed by rains (Ludlow and Muchow,

1990). Under non-stress conditions, however, late-flowering varieties tend to yield higher than the early flowering ones (Turner, 1986; Ludlow and Muchow, 1990). This is because the early flowering varieties are likely to leave the yield potential unutilized (Ludlow and Muchow, 1990).

2.4.2. Drought (or dehydration) avoidance

Drought (or dehydration) avoidance is the plant's ability to retain a relatively higher level of hydration under conditions of soil or atmospheric water stress (Hussain, 2006). Plants that avoid dehydration employ either reduced transpiration (water savers) or develop means other than reduced transpiration (water spenders) such as special root features to increase water uptake, leaf and stomata characteristics to reduce water loss and osmotic adjustment to lower the osmotic potential or in combination of all (Farooq *et al.* 2009).

Osmotic adjustment helps some plants to maintain turgor pressure through the active accumulation of solutes (osmoprotectants, or compatible solutes). These molecules, which act as osmotic balancing agents, are accumulated in plant cells in response to drought stress and are subsequently degraded after the alleviation of the stress (Ludlow and Muchow, 1990). Osmoprotectants include amino acids, sugar alcohols, polyols and quaternary ammonium and tertiary sulfonium compounds and help in protecting cell components from the adverse effect of water loss through expression of cell rescue mechanisms and through increased capacity of plants to recover after stress (Ludlow and Muchow, 1990).

2.4.3. Drought (dehydration) tolerance

Dehydration tolerance describes the ability of plants to continue metabolizing and growth at low leaf water potential and to maintain growth despite dehydration of the tissue or to recover after release from stress conditions (Hussain, 2006). Translocation is considered as one of the most dehydration-tolerant processes in plants which can proceed functioning at levels of water deficit sufficient to inhibit photosynthesis (Hsiao, 1973; Boyer, 1976). Hsaio (1973) and Boyer (1976), noted that grain growth in cereals is partially supported by translocated plant reserves stored mainly in the stem during the pre-anthesis growth stage. When water stress occurs and the current photosynthetic source is inhibited, the role of

stem reserves as a source for grain filling increases, and aids the grain filling process in stress affected plants during the grain filling stage.

However, drought escape, by reducing growth duration, might lead to reduced yields. Drought avoidance, by reducing water loss through stomata closure and leaf area reduction, might result in reduced photosynthetic ability ultimately resulting in reduced carbon assimilates and yield. Osmotic adjustment, through increased solute concentration, might have a detrimental effect in addition to energy requirement for active solute absorption. Similarly, accelerated leaf senescence and leaf abscission, to reduce a canopy size by reducing growth and shedding of older leaves to resist drought stress, might reduce the yields of annual crops, with economical loss to farmers (Ludlow and Muchow, 1990). Therefore, adaptation of plants to stress should reflect a reasonable balance in cost for the adaptation mechanisms (escape, avoidance and tolerance) while maintaining adequate productivity.

2.5. Breeding for Drought Tolerance

The ability to tolerate drought and display economic yield under suboptimal water supply differs from plant to plant. Some crops are naturally more drought tolerant than others (susceptible genotypes), and are obviously better suited to drought environments (Tuinstra, 1996; Nguyen, 2004). Conventional breeding works based on the direct phenotypic identification of this genetic variability to drought among crop varieties, or among sexually compatible species, and introducing this tolerance into lines with suitable agronomic characters to develop plants with superior quality (Boyer, 1982).

To this end, conventional plant breeding has been playing a successful role on a global scale in the development of improved cultivars with improved yields, nutritional qualities, and adaptabilities to environmental stresses (Moose and Mumm, 2008). The incorporation of drought tolerance trait in the generation of new maize and wheat germplasm by the International Maize and Wheat Improvement Center (Ribaut *et al.*, 2004), the development of maize hybrids with increased drought tolerance (Bruce *et al.*, 2002), and the introduction of wheat (*Triticum aestivum*) and rice (*Oryza sativa*) varieties that spawned the Green Revolution (Moose and Mumm, 2008) are some examples in the achievements of conventional breeding programs in the past.

However, due to the complex nature of the trait and the complicating effects of the environment, the progress of the classical plant breeding programs to enhance levels of drought tolerance for many crop species is not as rapid as it is for simpler traits (Hussain, 2006). Moreover, only a few of the morphological, biochemical, and physiological mechanisms used for selection criteria have been demonstrated to be casually related to the expression of tolerance under field conditions (Ludlow and Muchow, 1990). As a result, the traditional breeding method by using yield as a selection index and performing multi-environmental yield trials has been costly and slow (Hussain, 2006).

As drought tolerance is a complex trait, breeding for tolerance has been hampered by interactions between genotype and environment as well as by variation and intensity of rainfall from year to year (Tuinstra *et al*., 1996; Pennisi, 2008). As phenotype is the result of genotype and environmental interaction, assessment of desired genotypes is highly dependent on proper environmental conditions (Hussain, 2006). However, getting the appropriate or optimum environment for the assessment and selection of desired genotypes is a problem due to lack of reproducible screening techniques and the inability to routinely create defined and repeatable water stress conditions (Hussain, 2006).

2.6. Sorghum for Drought Tolerance

Sorghum, an important failsafe crop in the global agro ecosystem, is well known for its capacity to tolerate conditions of limited moisture and be productive during periods of extended drought circumstances that would impede production of most other grains (House, 1985; Paterson, 2008). The evolution of sorghum in semi arid tropical Africa under the pressure of drought and high temperature imparted the ability to tolerate drought, soil toxicities, and extremes of temperature more effectively to cultivated sorghum plants than other cereals (Rao *et al*., 2009). Its waxy leaves, an extensive root system, its ability to temporarily stop growing in periods of drought and recovering again when moisture becomes available make sorghum an important crop in arid and semi arid environments where it may not be productive to grow other cereals (Rao *et al*., 2009).

Also, sorghum belongs to the group of C_4 plants (Wyrich *et al.*, 1998). C_4 photosynthesis is a characteristic feature of C_4 plants conferring them with high rates of photosynthesis

and efficient use of water and nitrogen resources (Matsuoka *et al*, 2001; Lopes *et al*, 2011). Samson and Knopf (1994) indicated that Crops with a four-carbon (C_4) photosynthetic pathways produce 30% more dry matter (DM) per unit of water than three-carbon (C_3) crops and are more adapted to semiarid production regions. As a result, C_4 plants are among the most productive crops in agriculture.

The majority of terrestrial plants, including many important crops such as rice, wheat, barley, soybean, and potato, assimilate atmospheric CO_2 through the C_3 photosynthetic pathway (Calvin cycle) and classified as C_3 plants (Matsuoka *et al*,2001). In this pathway, the enzyme of primary CO_2 fixation, ribulose 1, 5-bisphosphate carboxylase/oxygenase (Rubisco), reacts not only with CO_2 but also with O_2 leading to photorespiration, which essentially wastes assimilated carbon. As a consequence, potential photosynthesis in C_3 plants is suppressed (lost) by oxygen by as much as 40% (Matsuoka *et al*, 2001). The extent of suppression further increases under stress conditions such as drought, high light, and high temperature through a decline of CO_2 concentration inside leaves due to closure of stomata (Matsuoka *et al.*, 2001; Lopes *et al*,2011). C_4 plants such as sorghum, maize, and sugarcane, in contrast, have evolved a novel biochemical mechanism to overcome photorespiration in such a manner that, in addition to the C_3 pathway, they use the C_4 photosynthetic cycle to elevate the CO_2 concentration at the site of Rubisco there by suppressing its oxygenase activity and increasing the efficiency of photosynthesis (Matsuoka *et al*, 2001).

Thus, C_4 photosynthesis is characterized by the presence of a metabolic CO_2 pump that concentrates CO_2 in the vicinity of the main enzyme of carbon dioxide fixation, ribulose-1,5-bisphosphate carboxylase/oxygenase, Rubisco (Matsuoka *et al.*, 2001; Lopes *et al.*, 2011) which increases the rate of photo synthesis and also the water use efficiency (as it allows high rates of photosynthesis to occur even when stomata are closed) while limiting flux through the photo respiratory pathway. This enables C_4 plants such as sorghum to achieve elevated photosynthetic capacity of up to twice as high as that of C_3 plants particularly at higher temperatures (Matsuoka *et al.*, 2001). Also, studies indicated that in C_4 species, the enzymes that comprise the metabolic CO_2 pump are more resistant to water deficits than the enzymes of C_3 photosynthesis (Lopes *et al*., 2011).

In sorghum, two distinct drought responses namely, pre-flowering and post- flowering drought responses, have been described (Sanchez *et al*., 2002). Sorghum lines with distinct phenotypic response to pre-flowering and post-flowering drought stress have been characterized (Kebede *et al.*, 2001). The pre-flowering response occurs when the plants are under significant moisture stress before flowering especially from panicle differentiation to flowering. This type of stress has been demonstrated to affect panicle size, grain number, and grain yield (Tuinstra *et al.,* 1996).

Tuinstra *et al.* (1996) reported that six regions in the sorghum genome were found to be associated with pre-flowering drought tolerance and eight additional regions generally associated with yield or yield components under fully irrigated conditions. Four major QTLs responsible for pre- flowering drought tolerance have also been identified in linkage group C, E, and F using a SC56 × Tx7000 cross in sorghum (Kebede *et al.*, 2001).

Post-flowering drought response in sorghum is expressed when moisture stress occurs during the grain development stage. When water is limited during the grain-filling period, rapid premature leaf death occurs, which in turn leads to charcoal rot, stalk lodging, and significant yield loss (Kebede *et al.*, 2001). Drought stress during the post- flowering period accelerates senescence in many plant species including sorghum (Sanchez *et al*., 2002). Chlorophyll loss and a progressive decline in photosynthetic capacity generally characterize senescence. Early onset of senescence affects assimilation and grain filling in crop plants (Tuinstra *et al.,* 1996).

Ejeta and Knoll (2007) pointed out that drought at any stage of crop development affects growth and production, but drought during the flowering stage causes maximum crop damage. Though sorghum is generally tolerant to many stresses including moisture stress compared to other crops, drought seriously affects this crop at the reproductive stage especially during and post-flowering stage implying that the true potential of sorghum can only be realized through concerted genetic improvement (House 1985; Kebede *et al*., 2001; Ejeta and Knoll, 2007). Besides the direct effect on yield, drought also predisposes the crop to other yield limiting factors such as pests and diseases (Tadesse *et al.,* 2008).

Drought is frequent in large areas of Asia and Africa, where most of the world's poor people live and where they fully depend on rainfall for crop production (Kebede *et al*.,

2001). All over the world, most of the cereal crops are grown under rain fed conditions where irrigation is unavailable or, if available, too costly for marginal farmers (Kebede *et al*., 2001; Ejeta and Knoll, 2007). In Ethiopia, sorghum is largely cultivated in moisture stress areas that cover nearly 66% of the total area of the country (Tadesse *et al*., 2008).

Under water-limited environments, genetic improvement of crops for drought tolerance is a sustainable and economically feasible solution to reduce the impact of drought (Tadesse *et al*, 2008). Wide genetic variations exist among sorghum germ plasm indicating the potential to develop new sorghum varieties for moisture stress environments (Tadesse *et al*, 2008). Basically, sorghum is a warm season, day length-sensitive, C_4-type metabolism plant (Doggette, 1988; Paterson *et al.*, 2008). However, for almost any feature used to describe the plant, diversity rather than homogeneity is a more fitting characterization. This might be linked to the wide environmental adaptation of sorghum. Different sorghum races or cultivars express adaptation to temperate and/or tropical climates, high or low altitudes, water logging, or drought stress conditions, and so on, indicating the existence of wide genetic materials in sorghum for screening and developing desired varieties (Rao *et al.*, 2009).

2.7. Molecular Markers

The amount of genotypic (on the DNA level) variability present in a group of individuals or in a population is generally defined as Genetic diversity, and this diversity gives species the ability to adapt to changing environments, including new pests, diseases and new climates (Parmesan and Yohe, 2003). Knowledge about variability present in germplasm collections and genetic relationships among breeding materials is essential in crop improvement strategies to develop desired traits. A number of methods have been developed over the past decades for detecting genetic variation in germplasm accessions, breeding lines and populations. These methods have relied mainly on the availability of Molecular/genetic markers (Avise, 2004).

A molecular/DNA marker is a constant identifiable-landmark that represents variation at a particular site on the genome, which is inherited in a Mendelian way, is easy to assay and can be followed over generations (Farooq and Azam, 2002).

Before the advent of molecular markers, morphological markers were found to be a source in varietal identification and assessing genetic variability (Farooq and Azam, 2002). Morphological markers are usually easily monitored and visually identifiable phenotypic characters such as flower color, seed shape, and pigmentation. However, These markers were not reliable as they are limited in number, influenced by environment, some appear late in plant development (e.g. flower color) making early scoring impossible, and some morphological markers can also affect other morphological markers or traits of interest in breeding programs because of pleiotropic gene action (Farooq and Azam, 2002). Thus, their utilization was limited by these factors.

Later, protein markers were developed. Protein markers are usually neutral on phenotypes of plants and are also co-dominant making the discrimination possible between homozygous and heterozygous. However, due to limited number of protein (isozyme) markers and because of the requirement of different protocol for each isozyme system, their utilization was very limited (Mohan *et al.*, 1997).

The advent of molecular (DNA) markers in the early 1980's solved the limitation of both morphological and protein (isozyme) markers. Unlike morphological and protein marker systems, molecular (DNA) markers are not limited in number, not influenced by environmental and developmental stage and are also phenotypically neutral (Farooq and Azam, 2002). Genetic variation detected by molecular markers has been useful for improving the breeding efficiency and varieties of molecular markers have been developed during the last two decades which include restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), microsatellites or simple sequence repeat (SSRs), single nucleotide polymorphisms (SNPs) and more recently DNA arrays technology (DArT) markers (Farooq and Azam, 2002).

These DNA markers have been proven useful for a variety of purposes relevant to crop improvement such as for the preparation of saturated molecular maps (genetical and physical), study of genome organization, phylogenetic analysis, germplasm characterization, characterization of transformants, QTL mapping, gene introgression through backcrossing, and marker-assisted selection (MAS) (Varshney and Tuberose, 2007). These DNA marker types differ in a variety of ways-such as technical requirements

(e.g. whether they can be automated or require use of radioactivity); the amount of time, money and labor needed; the number of genetic markers that can be detected throughout the genome; and the amount of genetic variation found at each marker in a given population (Varshney and Tuberose, 2007).

For plant breeding applications in general and for association mapping in sorghum in particular, SSR markers among different classes of the existing markers, are the most commonly used marker systems and have been proven and recommended as markers of choice (Varshney and Tuberose, 2007; Shehzad *et al*., 2009; Wang *et al.*, 2011). Although the more abundant SNP markers have been developed for a number of model crop plants such as rice and maize, due to the high development/detection cost of SNPs, SSRs remain an attractive marker system for sorghum and other crops (Wang *et al*., 2011). Compared to SNP marker system, SSR markers require just the primer sequences which are usually available once published. This is the primary reason why so much effort has been devoted to developing SSR markers (Srinivas *et al*., 2009; Yonemaru *et al*., 2009; Wang *et al*., 2011) and to using SSR markers in genetic mapping (Bhattramakki *et al*. 2000; Haussmann *et al*. 2002; Wu and Huang 2006), diversity studies (Shehzad *et al*. 2009; Wang *et al*., 2011) and molecular breeding (Ejeta and Knoll, 2007) in sorghum.

Microsatellites are tandem repeats of two to six nucleotides found in most nuclear genomes (Jimenz *et al*., 2010) in the form of di nucleotides as (GT)n; (CT)n; (GA)n, tri nucleotides as (CAC)n, tetra nucleotides as (GACA)n and (GATA)n, and so on. Since their discovery over three decades ago and with the rise of polymerase chain reaction (PCR) technology, microsatellite markers have progressively become the current molecular marker of choice for population genetic studies. Their popularity is also based on their ubiquity across most organisms, their co-dominant inheritance, their high genome coverage, and their high variability resulting from high mutation rates (Selkoe and Toonen, 2006).

2.8. Sorghum Linkage Mapping

Mapping is putting markers, genes or QTLs in order, indicating the relative distances among them and assigning them to their linkage groups or chromosomal regions. DNA markers in many fields, including agricultural research, have been used in the construction

of linkage maps for identifying chromosomal regions that contain genes controlling both simple and complex traits (Mohan *et al*., 1997). As constant-identifiable-landmarks that represent variation at a particular site on the genome, DNA markers can be ordered on chromosomes and in this way the entire genome of an individual can be visualized on a genetic linkage map (Varshney and Tuberose, 2007). Using the marker map, putative genes affecting traits of interest can be detected by testing for statistical associations between marker variants (alleles) and any trait of interest.

Molecular breeding becomes more effective if the molecular map is densely populated with markers, in order to provide more choice in the quality and type of marker and to increase the probability of polymorphic markers in important chromosomal intervals (Mace *et al*., 2009). Linkage between a marker and a genomic region (QTL) that influences quantitative trait/s permits detection and mapping of QTLs in relation to the markers. High-density map facilitates MAS especially between closely related types as it provides information on many potentiality polymorphic markers in any genomic region (Chittenden *et al.,* 1994). Hence, construction of linkage map is the most fundamental step required for a detailed genetic study of quantitative characters in any crop, to exercise indirect selection for several agronomic traits, and to isolate the genes involved based on their map position.

Sorghum linkage mapping based on DNA markers began in early 1990s and since then, several genetic maps have been developed with a large number of DNA-based markers including RFLPs, AFLPs and SSRs (Litt and Luty, 1989). The genome size of sorghum is relatively small as compared to other cereals and is thought to be used for the study of the evolution and function of other cereal genomes, as well. Among the five major cereal crops, the genome of sorghum is the second smallest (750 Mb) after that of rice (440 Mb) and is between one-third and one quarter the size of the maize genome (2500 Mb) (Yonemaru *et al*., 2009). Sorghum is more closely related to sugarcane and maize than to rice, and it shared a common ancestor as recently as10 million years ago with sugarcane (Yonemaru *et al*., 2009), 18–25 million years ago with maize (Yonemaru *et al.*, 2009) and about 50 million years ago with rice (Chen *et al*., 1998). However, despite the separation of sorghum from maize and rice approximately 18-25 million years ago and 50 million years ago respectively, significant conservation of gene order exists among the genomes of

these plants, which facilitates comparative genome mapping approaches (Yonemaru *et al.*, 2009).

Its small genome size with lower level of gene duplication than many other tropical cereals and its representativeness of tropical grasses because of its C_4 photosynthesis coupled with its adaptation to harsh environments, diverse germplasm, and close degree of relatedness to other economically important crops, make sorghum an attractive model plant among the panicoid grasses for the study and better understanding of the structure, function, and evolution of cereal genomes including the genomes of the complex grasses such as maize and sugar cane (Paterson *et al*., 2008; Mace *et al.*, 2009; Yonemaru *et al*., 2009). Linkage mapping in sorghum also takes advantage of the plant's straightforward diploid genetics, amenability to inbreeding, and high levels of polymorphism between sorghum species and adequate levels within *S. bicolor* (Paterson *et al*., 2008).

Recently, the whole genome of sorghum was sequenced and annotated and the draft of sorghum genome sequence was released by the US Department of Energy's Joint Genome Institute and Center for Integrative Genomics in 2007 (Li *et al*., and Paterson *et al*., 2009). This whole genome shotgun (WGS) sequencing by means of methylation filtration has tagged 96% of the sorghum genes. In the sorghum genome, a total of 109,039 tandem repeats were detected, of which 15,194 were microsatellites (SSR) (Li *et al.,* and Paterson *et al*., 2009). This has provided a valuable resource to develop large numbers of SSR markers in sorghum.

Molecular breeding strategies are increasingly being adopted to develop genetic linkage maps and to identify genomic regions influencing traits of importance in sorghum such as stay-green (Harris *et al*., 2007), fertility restoration (Klein *et al*., 2005), ergot resistance (Parh *et al*., 2006), midge resistance (Tao *et al*., 2003), and photoperiod sensitivity (Chantereau *et al*., 2001). In the past several linkage maps have been constructed in sorghum.

For example, Bhattramakki *et al*. (2000) constructed a linkage map composed of 147 SSR loci and 323 RFLP loci and reported the genetic map locations of 113 novel SSR loci in sorghum using a mapping population of 137 F8 recombinant inbred lines (RILs) derived

from the cross between BTx623 and IS 3620C. Most of the SSR primer sequences reported were developed from clones isolated from two sorghum BAC libraries and three enriched sorghum genomic DNA (gDNA) libraries. Very few of the sorghum SSRs primer sequences reported were developed from the sorghum DNA sequences present in the public data bases. Loci detected by 323 RFLP probe-enzyme combinations and 313 SSR primer pairs were mapped (LOD score _3.0). Of the SSR primers developed, 165 (53%) were found to detect polymorphism in a population composed of 18 diverse sorghum lines.

Hausmann *et al*. (2002) reported the construction of linkage map using two recombinant inbred populations (RIP-1, -2) of F3:F5 lines developed from the crosses IS9830 x E36-1 (1) and N13 x E36-1 (2). The genetic maps of RIP-1 and RIP-2 spanned 1,498 cM and 1,599 cM, respectively, with 137 and 157 markers distributed over 11 linkage groups.

Nagaraji *et al*. (2005) constructed thirteen linkage groups containing 60 SSR loci by using a set of sorghum recombinant inbred lines (RILs) obtained from the cross 96-4121 (greenbug-tolerant) x Redlan (greenbug susceptible). The linkage group (LG) spanned a distance of 603.5 cM, with the number of loci per LG varying from 2 to 14.

2.9. Conventional QTL Mapping

Conventional QTL mapping requires the construction of a linkage map using segregating populations derived from a cross between phenotypically divergent accessions where recombination frequencies between markers and the genes of interest are estimated from their patterns of co segregation (Ribaut & Hoisington, 1998). In the offspring of such a cross, association between a trait and marker alleles arises from a linkage between a marker loci and a trait. By identifying these associations and by statistically testing their significance, the method allows the location of genomic regions on a marker linkage map that most likely contain genes involved in controlling the trait of interest(Mohan *et al.*, 1997).

Therefore, in conventional quantitative trait loci (QTL) mapping in plants, the first step is generating a mapping population (F_2, backcross, recombinant inbred, etc.) from a bi parental cross, then genotyping the individuals with genetic markers across the genome, the third step is phenotyping the individuals for the trait of interest, and finally analyzing

the results via linkage mapping (Flint-Gracia *et al*., 2005). Once molecular markers are identified as being significantly linked to traits of interest, plant breeders can utilize this information to aid in cultivar development by genotypic selection. The ultimate goal of QTL mapping is, therefore, to employ molecular markers that are tightly linked to agronomically important genes as "land marks "on chromosomes for marker-assisted selection (MAS) and gene introgression (Ribaut & Hoisington, 1998).

The basic concept of associating genetic markers with quantitative traits (QTL mapping) was first proposed in 1923 by Sax where he noted that seed weight in bean (a complex trait) was associated with seed coat color (a simple, monogenic trait) in an F_2 cross of bean (Young, 1996). The potential use of linkages between qualitative genes and QTLs for studying the nature of quantitative genetic variation was largely recognized by geneticists of that time in spite of the fact that the availability of relatively small number and sometimes-deleterious nature of qualitative marker genes limited their utility for QTL mapping (Brubeck *et al.,* 1993). This concept of using simple monogenic traits for studying quantitative genetic variation was further elaborated by Thoday in 1961 who suggested that if the segregation of simply inherited monogenes could be used to detect linked QTLs, then it should eventually be possible to map and characterize all the QTLs involved in complex traits (Young, 1996).

Modern QTL mapping is the fulfillment of this idea, with the key innovation being that defined sequences of DNA act as the linked monogenic markers (Young, 1996). With the development of comprehensive DNA marker maps, it is now possible to search for QTLs throughout the genomes of most crop species (Myles *et al*., 2009). QTLs can be mapped either using conventional QTL mapping method or using association mapping strategies. In the past, using conventional QTL mapping, a large number of genes for various traits (quality traits, resistance to biotic stresses, etc.) have been tagged with markers in sorghum and other crops (Agrarwissenschaften, 2008).

2.9.1. Conventional QTL mapping in sorghum

Drought tolerance is one of the difficult traits that are genetically complex and highly influenced by environmental factors where selection for this trait is often difficult, and gains from selection are usually quite low (Ejeta and Knoll, 2007). Controlled by many

genes and dependent on the timing and severity of moisture stress, drought tolerance has been difficult to study and characterize (Ejeta and Knoll, 2007). Identification of QTLs and associated molecular markers is expected to facilitate improvement of drought tolerance in sorghum and other crops (Ejeta and Knoll, 2007).

Selection based on linked DNA markers to phenotypic variations is thought to facilitate manipulation of polygenic traits such as drought tolerance without affecting other traits (Tuinstra *et al.,* 1997). In sorghum, many efforts have been made to map QTLs or link them with molecular markers.

Xu *et al.* (2000) mapped QTLS controlling chlorophyll content in sorghum using a mapping population of 98 RILS derived from the cross B35 xTRx7000. They identified three QTLs for leaf chlorophyll content (ch11, ch12, and ch13) that together explained 25 - 30% of the observed phenotypic variability.

Reddy *et al*. (2007) identified QTLs associated with stalk rot resistance in sorghum using 93 RILs derived from a cross between inbred lines IS22380 (susceptible) x E36-1 (resistance). This population was genotyped with 62 nuclear and 4 genic SSR and 19 RAPD markers and linkage group was constructed using mapmaker/QTL software. They identified five QTLs on linkage groups A, B, D and I.

Srinivas *et al*. (2009) reported that 49 QTLs were mapped controlling for 12 agronomic traits including plant height , days to anthesis , days to maturity , total number of leaves , green leaf area at anthesis , green leaf area at maturity , % green leaf area preserved at maturity , grain yield , panicle length , panicle weight , number of primary branches , and seed weight using a mapping population of 168 F_7 RILs derived from the cross 296B × IS18551. The linkage map was constructed using 152 marker loci comprising 149 microsatellites (100 genomic- and 49 genic-microsatellites) and three morphological markers. They reported that several genomic regions affected multiple traits, suggesting the phenomenon of pleiotropy or tight linkage. Stable QTLs were identified for the studied traits across different environments, and genetic backgrounds by comparing the QTLs in the study with previously reported QTLs in sorghum. Of the 49 mapped genic-markers, 18 were detected associating either closely or exactly as the QTL positions of agronomic traits.

Shiringani *et al.* (2010) identified QTLs associated with sugar components (Brix, glucose, sucrose, and total sugar content) and sugar-related agronomic traits (flowering date, plant height, stem diameter, tiller number per plant, fresh panicle weight, and estimated juice weight) using a population of 188 recombinant inbred lines (RILs) from a cross between grain (M71) and sweet sorghum (SS79) grown in four different environments (two locations). A genetic map with a total of 157 markers (102 AFLP, 49 SSR, and 6 EST-SSR markers) was constructed, and several QTLs (seventeen QTLs for controlling sugar content, fourteen QTLs for stem diameter, 11 QTLS for glucose content, seven QTLs for plant height, seven QTLs for sucrose content, six QTLs for number of tillers per plant, four QTLs for brix, five QTLs for flowering date, one major QTL for fresh panicle weight, and one QTL for stem juice weight) were detected.

2.10. Association Mapping

While conventional QTL mapping in the past has served plant geneticists and breeders well, it has some key limitations. In conventional QTL mapping, one must grow the plants for at least three years which makes it very expensive both in terms of time and man power; only two alleles at any particular locus can be assessed which lowers the diversity ; and very large segregating populations are required to achieve high resolution mapping and even difficult (low power of resolution) due to the modest degree of recombination within the population, where resolution is often in the range of 10–30 cM, which corresponds to several millions of bases and hundreds of genes that in turn limits the precision of QTL localization (Flint-Gracia *et al*., 2003).

Association mapping is recently emerged as a powerful tool to identify QTLs in plants (Myles *et al.,* 2009). Unlike conventional QTL mapping which is a highly controlled experiment where individuals are crossed to generate a mapping population (Myles *et al,* 2009; Yu *et al.,* 2006), association mapping is a natural experiment where genotype and phenotype data are collected from a population in which relatedness is not controlled by the experimenter, and correlations between genetic markers and phenotypes are sought within that population (Myles *et al*., 2009).

In association mapping studies, a collection of cultivars, lines, and/or landraces, genotyped with densely spaced markers, can be used as mapping population which makes it different from conventional QTL mapping where typically bi parental crosses with contrasting genotypes are used (Sorkheh *et al*., 2008). Sorkheh *et al*. (2008) stated that in plant genetics, using a collection of cultivars has a number of advantages over the use of a bi parental cross. Firstly, in the population a broader genetic variation in a more representative genetic background will be available. This implies that one is not limited to the marker and trait loci that happen to differ between two parents. Secondly, LD mapping can attain a higher resolution because of the use of all meioses accumulated in the breeding history. Thirdly, historic phenotypic data on cultivars can be used to link markers to traits, without the need for new trials with special mapping populations (Sorkheh *et al*., 2008).

Association mapping is based on linkage disequilibrium. The terms linkage disequilibrium and association mapping have been often used interchangeably in literature (Flint-Gracia *et al*., 2003; Gupta *et al*., 2003). However, association mapping refers to a significant linkage of a molecular marker with a phenotypic trait, whereas Linkage disequilibrium (gametic phase disequilibrium or allelic association) is the non-random association of alleles at different loci (Zhao *et al*., 2005).

From the view point of association mapping, the term Linkage disequilibrium (LD) does not reveal its appropriate meaning and its usage is considered more of a barrier than an aid to understanding (Gupta *et al*.,2005). This is due to the fact that non random association of alleles can occur due to different factors such as, population structure, genetic drift, inbreeding, admixture, physical linkage of alleles, and so on. Among these factors, only the correlated inheritance of alleles at two or more loci due to the physical proximity (linkage) on a chromosome serves the purpose for association mapping. Hence, association mapping is one of the several manifestations of LD (Gupta *et al*., 2003).

Generally, factors such as physical linkage of alleles at two or more loci, migration, selection, inbreeding, population subdivision, low recombination rate, population admixture, genetic drift, and bottle neck tends to increase LD in a given population, while factors like out crossing, high recombination rate, high mutation rate, gene conversion,

and so on., lead to a decrease/disruption in LD (Flint-Gracia *et al*., 2003; Gupta *et al*., 2003).

In the absence of genetic drift, population admixture, migration, selection e.t.c., that confound the linkage disequilibrium patterns, knowing the level of LD in the germplasm is important in association mapping (Zhao *et al.*, 2005). A variety of statistics have been used to measure LD (Ardrie *et al.*, 2002, Zhao *et al.,* 2005). Of those LD measures, D' (the standardized LD coefficient) and R^2 (the square of the correlation coefficient between alleles at two loci) have been most commonly used in literature and are the preferred measures of LD (Farnir *et al*., 2000; Ardrie *et al.,* 2002).

However, *D*' is strongly affected by small sample sizes, and low allele frequencies (Flint-Gracia *et al*., 2003). For the purpose of examining the resolution of association studies, the R^2 statistic is generally favored, as it is indicative of how markers might correlate with the QTL of interest (Flint-Gracia *et al*., 2003). When R^2 is zero, alleles at two loci do not co-occur more frequently than would be expected under random sampling. R^2 approaches its maximum of 1 as alleles at two loci show more frequent co-occurrence within the population sample examined. With respect to association mapping, the most significant aspect of LD statistics is its ability to infer the predictive value of a marker locus for the association of the chromosomal region it resides with the phenotype.

However, the extent of LD (in base pairs) within species and even within individual genomes is highly variable (Long and Langley, 1999 cited in Varshney and Tuberosa, 2007). If LD decays too fast within a region, large number of markers would be required to scan target regions of a genome (Flint-Gracia *et al*., 2003). On the other hand, if LD decays too slowly, the size of the haplotype blocks would be too large to unambiguously reveal underlying causative locus. In other words, the decay of LD over physical distance in the study population determines the marker density required and the level of resolution that may be obtained in an association study (Flint-Gracia *et al.*, 2003). Also, the statistical power of associations is determined by the sample size used for the study (Wang and Rannala, 2005).

Although association mapping is based on LD, detecting LD in a given population does not ensure physical linkage of alleles on the same chromosome unless the cause that

makes it to occur is carefully studied, as mentioned earlier. Thus, one of the major limitations in applying association mapping to crop species is that the complex breeding histories of many important crops have created complex population structures within the germplasm (Flint-Garcia *et al*., 2003) and this uncontrolled relatedness among individuals can result in spurious signals of associations in downstream analyses due to the presence of unequal distribution of alleles within subpopulations which causes linkage disequilibrium to occur due to factor/s other than physical linkage (Myles *et al.*, 2009; Zhao *et al.*, 2005).

In reality, most populations have some degree of structure or subdivision and unless the amount of its effect is statistically considered, the simple relationship between strength of correlation and meiotic distance does not apply as correlations between unlinked loci often occur (Myles *et al*., 2009).

In such cases of genotype-phenotype covariance (i.e., when trying to map a phenotype whose variation is correlated with genetic relatedness), many genetic markers across the genome will appear to be associated with the phenotype when in fact these genetic markers simply capture the genetic relatedness among individuals (Myles *et al.*, 2009; Zhao *et al.*, 2005). According to Aranzana *et al*. (2005), this problem is particularly apparent when trying to map traits that have been subject to local adaptation, like flowering time because variation in these phenotypes between populations is highly correlated with allele frequency differences between populations. Flint-Garcia *et al*. (2005) reported that even for a set of common traits of agronomic interest in maize, such allele frequency differences account for an average of 9.3% of the phenotypic variation across all traits.

Recently, methods of association mapping (LD mapping) with an adjustment of marker-trait associations for these spurious associations have been introduced (Myles *et al*., 2009). The common approaches used to correct for population structure in association mapping include structured association approach (Pritchard *et al*., 2000b), K-model, and Q + K model. The structured association approach is used to identify populations and then estimate the proportion of each individual's variation that came from a particular population (Myles *et al*., 2009). The matrix of these estimates is known to be Q-matrix, and the estimates are used as covariates in the model to control for population structure in

association mapping (Pritchard *et al*., 2000b). The Q-matrix method is implemented in TASSEL as a general linear model. In this model, population membership estimates (covariates) can be derived using programs such as, STRUCTURE (Pritchard *et al*., 2000b). However, the problem with this approach is that individuals can only vary along a few axes of differentiation that may or may not be well captured by the general linear model function alone in TASSEL. This is because; extreme non linear type of relatedness may exist, for example, where many individuals have close relationships that cannot be described by a single vector of relatedness (Myles *et al*, 2009; Zhao *et al.*, 2007).

An alternative approach to capture this complex differentiation is to estimate the pair wise relatedness between all individuals in the sample; k-matrix (Myles *et al*., 2009). The statistical approach used to relate the pair wise relatedness matrix to a phenotype is the mixed linear model in TASSEL (K-model), where the variance explained by pair wise relatedness is fit to the vector of phenotypes (Yu *et al*., 2006; Zhao *et al*., 2007). Several authors (Yu *et al*., 2006; Malosetti *et al.*, 2007; Zhao *et al*., 2007; Kang *et al*., 2008) indicated that the application of mixed model methods using the K model in maize, human, mouse, Arabidopsis, and potato (*Solanum tuberosum*) demonstrated that the correction for pair wise relatedness using K-model significantly decreases false positives and false negatives much more than corrections involving the Q matrix. While Q-matrix takes only a few axes of variation into account, the K matrix captures the relatedness between each possible pair of individuals in a sample. However, although the k-model is far superior to the Q-matrix approaches (Q), in many cases a combination of these two approaches by including both Q-matrix and Kinship matrix in the analysis (Q+K) appears to be most powerful (Kang *et al.*, 2008).

Currently, the mixed linear model approach (which employs Q + K) is the method of choice for association mapping studies (Zhao *et al*., 2007; Myles *et al.*, 2009). On top of these methods, in order to further control spurious associations, rare alleles (with frequency<5%) in the population can be treated (filtered out) as missing data during population structure, linkage disequilibrium and association mapping analysis (Yu *et al.*, 2006). Successful use of these methods is expected to meet the challenges of connecting sequence diversity with heritable phenotypic differences (Bodmer, 1986).

In crop plants, the potential of exploiting association mapping to detect marker-trait associations was recently investigated for different crop species. According to Abdurakhmonov and Abdukarimov (2008), the pioneer association mapping studies in plants were performed by Virk *et al.* (1996) in rice and by Beer *et al.* (1997) in oat.

Virk *et al.* (1996) conducted their experiment using RAPD markers genotyped on 48 Asian rice accessions where they identified six traits (out of twelve) significantly associated with RAPD markers. Beer *et al.* (1997) detected significant associations of 13 QTL with RFLP loci using a total of 64 oat varieties including some land races, yet without considering the population structure that resulted in more increased associations than what were obtained in separate analysis of subpopulations (Jannink *et al*., 2002). Later, association mapping was extended to different crops by considering population structure and some of these works for different crop species are introduced as follows.

Thornsberry *et al*. (2001) conducted association mapping to evaluate *Dwarf8* sequence polymorphisms for flowering time in maize using a total of 92 inbred lines and reported that the sequencing of *Dwarf8* from these inbred lines led to the discovery of 41 distinct haplotypes with a variety of polymorphisms associated with differences in flowering time.

Kraakman *et al*. (2004) reported that associations between markers and complex quantitative traits were investigated using association mapping in a collection of 146 modern barley cultivars that were genotyped using 236 AFLP markers. A total of 24 significant marker trait associations were detected for yield and yield related traits. In addition, they reported that many of the associated markers were located in regions where earlier QTL were found for yield and yield components.

Simko *et al*. (2004) conducted association mapping in potato (*Solanum tuberosum* L.) for wilt disease resistance using 30 North American potato cultivars and allele-specific SNP markers. QTLs for resistance against wilt disease were detected on four chromosomes (2, 6, 9, and 12) with one QTL on chromosome 2 contributing over 40% to the total phenotypic variation of the trait.

Jun *et al.* (2008) reported that a genome-wide scan was performed using 150 simple sequence repeat (SSR) markers to identify QTL associated with seed protein content in

soybean. A total of 96 accessions obtained from Korea, China, and Japan germplasm collections were used for the study. Their report indicated that a total of 11 QTLs associated with protein content were identified with R^2 ranging from 6.3% to 65%.

Stewart and Kantartzi (2008) reported that fifty-six accessions of *G. arboreum* from nine regions of Africa, Asia, and Europe were evaluated for fiber related traits using 98 SSR markers and a total of 30 marker-trait associations were identified using 19 SSR markers located on 11 chromosomes with R^2 ranging from 10% to 18%.

Zhao *et al*. (2009) conducted association mapping for amino acid content in rice and reported a total of 42 marker-trait associations ranging from one to seven associations which were identified by 15 significantly linked SSR markers.

Liu *et al.* (2010) reported that association mapping was conducted using 103 wheat germ plasm accessions from China genotyped with 76 SSR markers and 40 EST-SSR markers and a total of 10 SSR markers on chromosome 4A were significantly associated with six agronomic traits.

Zhou *et al.* (2012) using 128 rice varieties genotyped with 152 microsatellite markers identified a total of 16 marker-trait associations using 12 significant SSR markers with R^2 ranging from 1.99 to 21.58%., in a genome wide association mapping study.

Reports indicated that sorghum is well suited to association mapping methodologies because of its medium-range patterns of linkage disequilibrium and its self pollinating mating system (Hamblin *et al*., 2005).

Among some recently published works regarding the feasibility of association mapping in sorghum, Hamblin *et al.* (2004) reported that population genetic parameters in sorghum were estimated by surveying 27 diverse *S. bicolor* accessions for sequence variations in 29,186 bp from 95 short regions derived from RFLP map. They reported that, consistent with its higher level of inbreeding, the extent of LD was several folds greater in sorghum than in maize though the total sequence variation in sorghum was about four fold lower than that in maize.

Hamblin *et al*. (2005) also detected the presence of short and medium range (up to100 kb) patterns of LD in sorghum. They reported that the extent of allelic associations in *S. bicolor*, as assessed by pair wise measures of LD, is higher than in maize but lower than in *Arabidopsis*, in qualitative agreement with expectations based on mating system. They indicated that, from a practical standpoint, these results suggest that *S. bicolor* is well suited for association studies since LD typically extends at least several kilo bases but has largely decayed by 15 kb (Hamblin *et al*., 2005).

Casa *et al.* (2008) compared different models of association mapping in a panel of 377 sorghum accessions representing all major cultivated races from tropical lines of diverse geographic and climatic regions and important U.S. breeding lines and their progenitors. Accessions were phenotyped for eight traits, and levels of population structure and familial relatedness were assessed with 47 simple sequence repeat (SSR) loci. The result of their analysis indicated that association models that accounted for both population structure and kinship performed better than those that did not.

Shehzad *et al*. (2009) reported that genome wide association mapping study was conducted using different association models to understand the feasibility and resolution of association mapping study in sorghum. In that study, 107 representative sorghum accessions collected from 27 countries of Africa and Asia and 98 SSR markers which were selected from three previously published linkage maps were used. Phenotypic data was recorded for 26 morphological traits. A total of 14 significant SSR loci were found to be associated with 12 different morphological traits including days to heading, days to flowering, Culm length, number of tillers, number of panicles and panicle length.

Wang *et al*. (2011) conducted pool based association mapping by screening the tall and short pools of sorghum accessions from the sorghum Mini Core collection developed at the International Crops Research Institute for the Semi-Arid Tropics with 703 SSR markers, and identified four markers that were closely associated with sorghum height on chromosomes 2, 6, and 9. They indicated that comparison with published maps showed that all four markers were clustered with markers previously mapped to height or height-related traits and with candidate genes involved in regulating plant height such as FtsZ, Ugt, and GA 2-oxidase.

3. MATERIALS AND METHODS

3.1. Description of the Study Area

The field experiment was conducted at Kobo (12°9′ N latitude 39°38′E longitude) which is located 581 kms north of Addis Ababa, in North Wollo, Ethiopia at an altitude of 1468 masl. The area receives an average annual rainfall of 718.6 mm with average minimum and maximum temperatures of 15 °C and 35 °C, respectively. The dominant soil type is acrisol with pH 8.5. Mean monthly temperature (^{o}C) and monthly total rainfall (mm) at Kobo during the experimental period is given in appendix Table 2.

3.2. Plant Materials

The entries for this study consisted of 160 sorghum (*Sorghum bicolor)* accessions (152 land races and 8 released lines). These accessions were sampled from sorghum accessions collected from all sorghum growing woredas of the country. The seeds of the land races were obtained from the Institute of Biodiversity Conservation (IBC) of Ethiopia where as the seeds of released varieties were provided by Melkasa and Sirinka Agricultural Research Centers.

The accessions were grown from May, 2010 to December, 2010 at Eladale, Jimma University College of Agriculture and Veterinary Medicine research field for seed multiplication and preliminary screening. After obtaining enough seeds from seed multiplication and preliminary screening conducted at Eladale, the seeds (the experimental materials) for the actual study were sown at kobo during the off season of 2011 using irrigation (from January- June). The list of the experimental materials is given in (Tables 1 and 2).

Table 1. List of sorghum land races used in the study

Entry number	Accesson name	Collection area			Altitude
		Region	Zone	Woreda	
1	69046	Oromia	Bale	Ginir	1630
2	69057	Oromia	Bale	Ginir	1630
3	69094	SNNP	Debub Omo	Bako Gazer	1410
4	69096	SNNP	Debub Omo	Bako Gazer	1410
5	69105	SNNP	Debub Omo	Hamer Bena	500
6	69183	Oromia	Mirab Harerge	Mieso	1460
7	69192	Oromia	Mirab Harerge	Mieso	1530
8	69210	Amhara	Oromia	Bati	1640
9	69227	Oromia	Mirab Harerge	Chiro	1750
10	69231	Oromia	Mirab Harerge	Chiro	1850
11	69238	Oromia	Mirab Harerge	Doba	1800
12	69241	Oromia	Mirab Harerge	Doba	NA
13	69286	Amhara	Debub wollo	Sayint	NA
14	69306	Amhara	Semen Gonder	Debark	1470
15	69371	Gambela	Zone 1	Itang	550
16	69391	Gambela	Zone 1	Itang	550
17	69392	Gambela	Zone 1	Itang	550
18	69442	Gambela	Zone 1	Gambela	630
19	69452	Gambela	Zone 1	Gambela	630
20	69468	Gambela	Zone 2	Abobo	530
21	69492	SNNP	Semen Omo	Bonke	1150
22	69494	SNNP	Semen Omo	Bonke	1150
23	69516	Amhara	Semen Gonder	Gonder Zuria	1580
24	70068	Oromia	Misrak Shewa	Boset	1450
25	70069	Oromia	Misrak Shewa	Boset	1450
26	70075	Oromia	Misrak Shewa	Adama	1600
27	70189	Oromia	Mirab Harerge	Mieso	1460
28	70216	Oromia	Misrak Shewa	Adama	1600
29	70301	Tigray	Mehakelegnaw	Laelay Maychew	NA
30	70306	Oromia	Misrak Harerge	Babile	1720
31	70537	SNNP	Gurage	Cheha	NA
32	71021	Oromia	Misrak Harerge	Meta	NA
33	71370	Oromia	Misrak Harerge	Kersa	NA
34	71418	Tigray	Mirabawi	Kafta humera	710
35	71421	Tigray	Mirabawi	Kafta humera	710
36	71422	Tigray	Mirabawi	Kafta humera	710
37	71425	Tigray	Mirabawi	Kafta humera	710
38	71522	Oromia	Mirab Harerge	Habro	NA
39	71536	Oromia	Mirab Harerge	Habro	NA

Table 1. List of sorghum land races (Continued)

40	71562	Oromia	Illubabor	Ale	NA
41	71565	Oromia	Illubabor	Ale	NA
42	71570	Gambela	Zone 1	Gambela	630
43	71590	Tigray	Mehakelegnaw	Laelay Maychew	NA
44	71625	Gambela	Zone 1	Gambela	630
45	71657	Gambela	Zone 1	Gambela	630
46	71739	Amhara	Semen Gonder	Chilga	NA
47	71748	Amhara	Semen Gonder	Chilga	1000
48	71788	Amhara	Misrak Gojam	Guzamin	NA
49	71795	Amhara	Misrak Gojam	Guzamin	NA
50	71810	oromia	NA	NA	NA
51	72437	Amhara	Debub Wollo	Ambasel	NA
52	72438	Amhara	Debub Wollo	Ambasel	NA
53	72470	Amhara	Debub Wollo	Dessie zuria	NA
54	72506	NA	NA	NA	NA
55	72588	Amhara	Debub Wollo	Kalu	NA
56	72611	NA	NA	NA	NA
57	72997	Affar	Zone 1	Ayisayita	1450
58	73067	Amhara	Semen Wollo	Weldiya	NA
59	73068	Amhara	Semen Wollo	Weldiya	NA
60	73339	Tigray	Mehakelegnaw	Adwa	1400
61	73341	Tigray	Mehakelegnaw	Adwa	1400
62	73637	Amhara	Misrak Gojam	Enemay	NA
63	73641	Affar	Zone 1	Ayisaita	NA
64	73646	Amhara	Misrak Gojam	Enemay	NA
65	73762	Tigray	Debubawi	Enderta	NA
66	73797	Tigray	Debubawi	Enderta	NA
67	73825	NA	NA	NA	NA
68	73940	NA	NA	NA	NA
69	73995	Tigray	Mirabawi	Tahtay koraro	NA
70	74097	Amhara	Semen Wollo	Gubalafto	1470
71	74108	Tigray	Debubawi	Rayaazebo	NA
72	74115	Tigray	Debubawi	Rayaazebo	NA
73	74262	Amhara	Semen Shewa	Efratanagidim	1420
74	74268	Amhara	Semen Shewa	Efratanagidim	1420
75	74698	SNNP	Bench Maji	Derashe special	1200
76	74761	Oromia	Semen Shewa	Gerarjarso	1600
77	74766	Oromia	Semen Shewa	Gerarjarso	1600
78	74935	Tigray	Mirabawi	Tahtay koraro	NA
79	74999	Oromia	Mirab Shewa	Ejerie (Addisalem)	NA
80	75000	Oromia	Mirab Shewa	Ejerie (Addisalem)	NA

Table 1. List of sorghum land races (continued)

81	75033	Amhara	Semen Gonder	Gonder zuria	NA
82	75065	Oromia	Arssi	Merti	1000
83	75066	Oromia	Arssi	Merti	1000
84	75129	Oromia	Jimma	Mana	NA
85	75216	Amhara	Semen Shewa	Debreberhan zuria	NA
86	75220	Amhara	Semen Shewa	Debreberhan zuria	NA
87	75353	Amhara	Misrak Gojam	Hulet ejenese	NA
88	75455	Amhara	Semen Wollo	Bugna	NA
89	200546	NA	NA	NA	NA
90	200547	NA	NA	NA	NA
91	201299	Oromia	Misrak Harerge	Kombolcha	NA
92	201302	Oromia	Misrak Harerge	Kombolcha	NA
93	201318	Amhara	Semen Wollo	Gubalafto	1470
94	201349	SNNP	Bench Maji	Konso special	1100
95	201491	Oromia	Mirab Harerge	Tulo	NA
96	201499	Oromia	Misrak Harerge	Kersa	NA
97	201502	Oromia	Mirab Shewa	Chelia	NA
98	204776	Eritrea			1426
99	206910	Amhara	Debub Gonder	Laygayint	NA
100	206921	Amhara	Semen Gonder	Wegera	NA
101	206923	Amhara	Semen Gonder	Belesa	NA
102	206936	Amhara	Semen Gonder	Belesa	NA
103	206943	Amhara	Semen Gonder	Wegera	NA
104	206952	Amhara	Debub Gonder	Laygayint	NA
105	206962	Amhara	Debub Gonder	Simada	NA
106	206963	Amhara	Debub Gonder	Simada	NA
107	207573	Tigray	Mehakelegnaw	Laelay Maychew	NA
108	210902	SNNP	Semen Omo	Bonke	1230
109	210922	SNNP	Semen Omo	Bonke	1150
110	213008	SNNP	Bench Maji	Derashe special	1450
111	213386	Amhara	Semen Gonder	Dembia	NA
112	213393	Amhara	Semen Gonder	Addi arkay	NA
113	214046	SNNP	Bench Maji	Sheko	NA
114	214067	SNNP	Bench Maji	Sheko	NA
115	215053	Oromia	Borena	Teltele	NA
116	215054	Oromia	Borena	Teltele	NA
117	215330	Amhara	Misrak Gojam	Hulet ejenese	NA
118	216739	Gambela	Zone 1	Itang	550
119	216742	Gambela	Zone 2	Abobo	600
120	216824	Oromia	Misrak Harerge	Babile	1670
121	217685	SNNP	Debub Omo	Bako gazer	1410
122	217705	SNNP	Bench Maji	Derashe special	1200

Table 1. List of sorghum land races (continued)

123	220251	Eritrea			1426
124	220252	Eritrea			1426
125	220267	Eritrea			1426
126	222882	Gambela	Zone 2	Gog	530
127	222884	Gambela	Zone 2	Gog	530
128	226080	Amhara	Semen Gonder	Dembia	1760
129	229230	Amhara	Semen Shewa	Siyadebrinawayuens	1440
130	229232	Amhara	Semen Shewa	Laybet na tachbet	NA
131	229238	Amhara	Semen Shewa	Laybet na tachbet	NA
132	229843	Amhara	Mirab Gojam	Burewemberma	1850
133	229844	Amhara	Misrak Gojam	Bibugn	1850
134	231191	Oromia	Misrak Harerge	Babile	1720
135	231200	Oromia	Misrak Harerge	Babile	1720
136	231230	Oromia	Arssi	Sherka	1740
137	235459	Tigray	Mehakelegnaw	Kolatemben	1750
138	235469	Tigray	Mehakelegnaw	Kolatemben	1700
139	235820	SNNP	Debub Omo	Hamer bena	500
140	235913	Amhara	Semen Gonder	Addiarkay	1640
141	235921	Amhara	Semen Gonder	Layarmacho	1150
142	235922	Amhara	Semen Gonder	Layarmacho	1000
143	237274	Tigray	Mehakelegnaw	Abergele	1450
144	237287	Tigray	Mehakelegnaw	Mereblehe	1420
145	237289	Tigray	Mehakelegnaw	Mereblehe	1350
146	239210	Amhara	Semen Shewa	Efratana gidim	1420
147	239232	Amhara	Semen Wollo	Guba lafto	1470
148	241183	Oromia	Mirab Harerge	Mieso	1320
149	241227	Oromia	Mirab Harerge	Mesela	1440
150	241705	SNNP	Bench Maji	Derashe special	1200
151	241728	SNNP	Bench Maji	Konso special	1600
152	243681	Tigray	Mirabawi	Tselemti	1200

RV= Released varieties, MARC=Melkasa agricultural research center,EIAR=Ethiopian Institute of Agricultural research, SARC= Sirinka Agricultural research center, ARARI=Amhara regional agricultural research institute,SNNP,Southern nations and nationalities people, NA=Information not available.

Table 2. List of the released varieties used

Entries	Released variety	Releasing center	Year of release	Pedigree/ Source	Adaptation	Special merit
1	76T1-23	MARC/EIAR	1979	954062Xpp9	Low land	Early maturing
2	B35	NA	NA	ICRISAT	Low land	Stay green
3	Baji	MARC/EIAR	1996	85MW5334	Mid altitude with high rain fall	High yield
4	Birmash	MARC/EIAR	1989	80LPYT-1	Mid altitude with high rain fall	High yield
5	E36-1	NA	NA	ICRISAT	Low land	Stay green
6	Gambela	NA	NA		Low land	Mdium to early maturing
7	Meco-1	MARC/EIAR	1997	ICRISAT	Low land	Early maturing
8	Teshale	SRARC,ARARI & MARC/EIAR	2002	ICRISAT	Low land	Early maturing

RV= Released varieties, MARC=Melkasa agricultural research center,EIAR=Ethiopian Institute of Agricultural research, SARC= Sirinka Agricultural research center, ARARI=Amhara regional agricultural research institute,SNNP,Southern nations and nationalities people, NA=Information not available.

3.3. Experimental Design and Trial Management

The field experiment was laid down in alpha lattice design with three replications having 16 blocks per replication and 10 plots per block and a spacing of 75 cm and 20 cm between rows and plants, respectively. Forty plants were planted per plot in two rows of 4 m long. The experimental plots were watered immediately after sowing to ensure uniform germination. Weekly interval irrigation was applied for the first three weeks. Starting from the fourth irrigation, watering was applied with 12 days interval till 50% flowering. Irrigation was withheld when the majority of the entries reached 50% flowering to allow moisture stress to develop during the grain-development stage. The recommended fertilizer rate of 100 kg Di ammonium phosphate (DAP) was applied by incorporating in to the soil during sowing the seeds followed by 25 kg urea ha $^{-1}$ by side dressing 55 days after the seeds were sown. Thinning was conducted after three weeks of sowing to maintain the plant distance and to balance the plant density. Karate was applied two times with a rate of 1mm1^{-} of water 30 and 45 days after emergence to protect against shoot fly. Other agronomic practices such as weeding were followed uniformly to all plots according to the recommendation of the location. No herbicide was applied to control weeds. Bird damage was protected by covering the heads of eight randomly chosen plants from each plot to obtain the average grain yield per panicle.

3.4. Data Collected

The following characters were measured on plants sampled at random from each plot at maturity. The procedure outlined in the sorghum descriptor was used to measure each trait (IBGR and ICRISAT, 1993).

Number of days to 50% flowering (days)

Determined by counting the number of days from emergence to when 50% of plants have started flowering in each plot.

Plant height (cm)

The heights of five randomly taken plants were measured in centimeters from the ground level to the tip of the main stalk.

Panicle weight (g)

The panicles from the main stem of five randomly chosen plants were harvested; air and oven dried, and used to calculate panicle weight. Panicle weight was calculated as the average of the weights of the five panicles measured.

Grain weight per panicle (g)

After obtaining the panicle weight of the five sampled panicles, threshing was done and grain weight per panicle was determined by weighing the seeds of the five sampled panicles and taking their average.

Thousand seed weight (g)

One thousand seed weight was determined by weighing 1000 seeds from the sampled panicles. An electrical seed counter and a sensitive balance were used, respectively, to count and weigh the samples.

Number of seeds per panicle (number)

The number of seeds per panicle was obtained by dividing the seed yield per panicle with one thousand seed weight and then multiplying by 1000 as follows:

$$\text{Number of seeds per panicle} = \frac{\text{Grain weight per panicle}}{\text{One thousand seed weight}} \times 1000$$

Panicle harvest index (PHI)

It was estimated as a ratio of seed weight to panicle weight. It was computed from five randomly selected plants per plot for panicle weight and seed weight per panicle.

3.5. Genotyping

SSR markers used

A total of 39 Simple Sequence Repeat (SSRs) markers, including 22 di, 9 tri, and 4 tetra nucleotide or longer motifs, and 4 compound repeats were used. These SSR markers were selected based on their uniform distribution in the sorghum genome. Four of them from chromosome SBI-01, five of them from chromosome SBI-02, four of them from chromosome SBI-03, two of them from chromosome SBI-04, four of them from chromosome SBI-05 and chromosome SBI-06 each, five of them from chromosome SBI-07 and chromosome SBI- 08 each and three of them from chromosome SBI-09 and chromosome SBI-10 each. These are the same set of markers that are selected and being used by the Generation Challenge Programme for genetic diversity assessment of global sorghum germplasm. The list of the SSRs markers, including primer sequences, information on repeat motif and length, and chromosome location are given in Table 3.

Table 3. List of sorghum SSR markers used for the study

Marker code	Marker name	Forward and reverse primer sequences(5'-3')	Motif type	Anealing temperture
1	gpsb067	F TAGTCCATACACCTTTCA	(GT)10	49
		R TCTCTCACACACATTCTTC		
2	gpsb123	F ATAGATGTTGACGAAGCA	(CA)7+(GA)5	50
		R GTGGTATGGGACTGGA		
3	mSbCIR223	F CGTTCCAATGACTTTTCTTC	(AC)6	55
		R GCCAATGTGGTGTGATAAAT		
4	mSbCIR238	F AGAAGAAAAGGGGTAAGAGC	(AC)26	55
		R CGAGAAACAATTACATGAACC		
5	mSbCIR240	F GTTCTTGGCCCTACTGAAT	(TG)9	55
		R TCACCTGTAACCCTGTCTTC		
6	mSbCIR246	F TTTTGTTGCACTTTTGAGC	(CA)7.5	55
		R GATGATAGCGACCACAAATC		
7	mSbCIR248	F GTTGGTCAGTGGTGGATAAA	(GT)7.5	56
		R ACTCCCATGTGCTGAATCT		
8	mSbCIR262	F GCACCAAAATCAGCGTCT	(CATG)3.25	57
		R CCATTTACCCGTGGATTAGT		
9	mSbCIR276	F CCCCAATCTAACTATTTGGT	(AC)9	53
		R GAGGCTGAGATGCTCTGT		
10	mSbCIR283	F TCCCTTCTGAGCTTGTAAAT	(CT)8 (GT)8.5	54
		R CAAGTCACTACCAAATGCAC		
11	mSbCIR286	F GCTTCTATACTCCCCTCCAC	(AC)9	55
		R TTTATGGTAGGATGCTCTGC		
12	mSbCIR300	F TTGAGAGCGGCGAGGTAA	(GT)9	61
		R AAAAGCCCAAGTCTCAGTGCTA		
13	mSbCIR306	F ATACTCTCGTACTCGGCTCA	(GT)7	56
		R GCCACTCTTTACTTTTCTTCTG		
14	mSbCIR329	F GCAGAACATCACTCAAAGAA	(AC)8.5	55
		R TACCTAAGGCAGGGATTG		
15	SbAGB02	F CTCTGATATGTCGTTGTGCT	(AG)35	55
		R ATAGAGAGGATAGCTTATAGCTCA		

Table 3 (continued)

16	Xcup02	F	GACGCAGCTTTGCTCCTATC	(GCA)6	54
		R	GTCCAACCAACCCACGTATC		
17	Xcup14	F	TACATCACAGCAGGGACAGG	(AG)10	54
		R	CTGGAAAGCCGAGCAGTATG		
18	Xcup53	F	GCAGGAGTATAGGCAGAGGC	(TTTA)5	54
		R	CGACATGACAAGCTCAAACG		
19	Xcup61	F	TTAGCATGTCCACCACAACC	(CAG)7	54
		R	AAAGCAACTCGTCTGATCCC		
20	Xcup63	F	GTAAAGGGCAAGGCAACAAG	(GGATGC)4	54
		R	GCCCTACAAAATCTGCAAGC		
21	Xgap72	F	TGCCACCACTCTGGAAAAGGCTA	(AG)16	55
		R	CTGAGGACTGCCCCAAATGTAGG		
22	Xgap206	F	ATTCATCATCCTCATCCTCGTAGAA	(AC)13/(AG)20	55
		R	AAAAACCAACCCGACCCACTC		
23	Xgap84	F	CGCTCTCGGGATGAATGA	(AG)14	55
		R	TAACGGACCACTAACAAATGATT		
24	Xisep0310	F	TGCCTTGTGCCTTGTTTATCT	(CCAAT)4	60
		R	GGATCGATGCCTATCTCGTC		
25	Xtxp010	F	ATACTATCAAGAGGGGAGC	(CT)14	50
		R	AGTACTAGCCACACGTCAC		
26	Xtxp012	F	AGATCTGGCGGCAACG	(CT)22	55
		R	AGTCACCCATCGATCATC		
27	Xtxp015	F	CACAAACACTAGTGCCTTATC	(TC)16	55
		R	CATAGACACCTAGGCCATC		
28	Xtxp021	F	GAGCTGCCATAGATTTGGTCG	(AG)18	60
		R	ACCTCGTCCCACCTTTGTTG		
29	Xtxp040	F	CAGCAACTTGCACTTGTC	(GGA)7	55
		R	GGGAGCAATTTGGCACTAG		
30	Xtxp057	F	GGAACTTTTGACGGGTAGTGC	(GT)21	55
		R	CGATCGTGATGTCCCAATC		
31	Xtxp114	F	CGTCTTCTACCGCGTCCT	(AGG)8	50
		R	CATAATCCCACTCAACAATCC		

Table 3 (continued)

32	Xtxp136	F	GCGAATAGCATCTTACAACA	(GCA)5	55
		R	ACTGATCATTGGCAGGAC		
33	Xtxp141	F	TGTATGGCCTAGCTTATCT	(GA)23	55
		R	CAACAAGCCAACCTAAA		
34	Xtxp145	F	GTTCCTCCTGCCATTACT	(AG)22	55
		R	CTTCCGCACATCCAC		
35	Xtxp265	F	GTCTACAGGCGTGCAAATAAAA	(GAA)19	55
		R	TTACCATGCTACCCCTAAAAGTGG		
36	Xtxp273	F	GTACCCATTTAAATTGTTTGCAGTAG	(TTG)20	55
		R	CAGAGGAGGAGGAAGAGAAGG		
37	Xtxp278	F	GGGTTTCAACTCTAGCCTACCGAACTTCCT	(TTG)12	50
		R	ATGCCTCATCATGGTTCGTTTTGCTT		
38	Xtxp320	F	TAAACTAGACCATATACTGCCATGATAA	(AAG)20	54
		R	GTGCAAATAAGGGCTAGAGTGTT		
39	Xtxp321	F	TAACCCAAGCCTGAGCATAAGA	(GT)4+(AT)6+(CT)2	55
		R	CCCATTCACACATGAGACGAG		

DNA Extraction

Ten seeds were grown in pots in the laboratory of Jimma University College of Agriculture and Veterinary Medicine. For each accession, genomic DNA was extracted from the fresh leaves of ten plants which were harvested in bulk from 14 days old seedlings. The fresh leaves were dried with silica gel. DNA was extracted following a modified CTAB (cetyl trimethyl ammonium bromide) extraction protocol (Mace *et al*., 2003).

The quality and quantity of the isolated DNA was determined by comparing DNA samples with a known concentration of λ-DNA after running them on a 0.8% agarose gel (0.8 gm agarose dissolved in 100 ml 1X TBE buffer) that contained 0.3 µg/ml ethidium bromide solution. At the end of electrophoresis, the gel was visualized using UV light and photographed using a video capture (Flowgen IS 1000). All samples were then normalized to the same concentration level (50 ng) and used for PCR.

Polymerase Chain Reaction Conditions and Amplifications

The PCR was performed in Nirobi (Kenya) using Gene-Amp PCR System 9600 (PE-Applied Biosystems) in 384- wells plates (ABGene, Rochester, New York.) in a total reaction volume of 10 µl that consisted of 1 µl DNA (50ng), 1 µl 10X PCR buffer, 1.5 µl MgCl2 (10 mM), 1 µl reverse primer (2 pmoles), 1 µl forward primer (2 pmoles), which were 5'-labelled with one of the 6-FAM, VIC, NED, PET fluorescent dyes (PE-Applied Biosystems), 0.5 µl of each dNTP (2 mM), 0.04 µl Taq DNA polymerase (5U) (PE-Applied Biosystems) and 3.46 µl distilled water. The amplification profile consisted of initial denaturation of the template DNA at 95^{o}C for 3 min, followed by 35 cycles, each for 30 sec at 95^{o}C (denaturation), 1 min at 56^{o}C (annealing), and 1 min at 72^{o}C (extension), and a final extension at 72^{o}C for 30 mins was included to minimise the +A overhang.

Capillary Electrophoresis

After the PCR, a few samples from each primer pair product were randomly selected and checked for proper amplification by comparing DNA samples with a known molecular weight of λ-DNA after running them on 2% agarose gel. An ABI plate was prepared with a total volume of 10 µl (9.0 µl from a mix of an injection solution mixed by vortexing (1ml) formamide (HIDI) (Perkin Elmer-Applied Biosystems) and 12.0 µl GS500 LIZ (Perkin Elmer-Applied Biosystems) was aliquoted into 96-well plates and 1.0 µl of pooled PCR products from each of the 6-FAM, VIC, NED and PET-labelled PCR products was added. DNA fragments were denatured at 95^{o}C for 3 min, chilled quickly for five minutes and size-fractioned using ABI 3730 Capillary DNA sequencer (PE-Applied Biosystems). In this system, the labeled PCR products were detected using a laser and capillary electrophoresis based on their fluorescent dye and fragment size. The peaks were sized and the alleles called using Gene Mapper software version 3.7 (PE-Applied Biosystems) and presented as alleles scored as estimated fragment sizes in base pairs compared to the internal size standard GS500LIZ-3730.

3.6. Data Analysis

3.6.1. Phenotypic data

Analysis of variance

The analysis of variance (ANOVA) for each character was performed using the statistical analysis system soft ware (SAS) version 9.2 (SAS Institute 2008).

The model for alpha lattice design was as follows

$Yijk = \mu + Ti + Pj + Bjk + eijk$

Where, Yij is the value of the observed trait for i-th treatment in the k-th block with in j-th replicate (super block); μ is the overall mean; Ti is the fixed effect of the i-th treatment (i=1,2,3,…,t); Pj is the effect of the jth replicate (super block) (where, j=1,2…r.); Pjk is the effect of the k-th incomplete block within the jth replicate (k=1,2,…,s); eijk is an experimental error associated with the observation of the ith treatment in the kth incomplete block within the jth complete replicate

Estimation of variability

Estimation of phenotypic and genotypic coefficient of variation

Phenotypic and genotypic variances were estimated using the following formula:

Genotypic variance ($\sigma^2 g$) $\sigma^2 g = \frac{MSg - MSe}{r}$

Where, r = number of replications

MSg = mean square due to genotypes and

MSe = mean square of error (environmental variance)

Environmental variance ($\sigma^2 e$) $\sigma^2 e = MSe$

Phenotypic variance ($\sigma^2 p$) $\sigma^2 p = \sigma^2 g + \sigma^2 e$

Where, $\sigma^2 g$ = genotypic variance and

$\sigma^2 e$ = mean square of error (environmental variance).

Phenotypic and genotypic coefficient of variation

Both genotypic and phenotypic coefficients of variability were computed as per the method suggested by Burton and Devane (1953).

Phenotypic coefficient of variation (PCV)

$$PCV=\frac{\sqrt{\sigma 2p}}{\bar{X}} * 100$$

Where, σ^2P = phenotypic variance and

$\bar{X}$ = mean of the character being evaluated

Genotypic coefficient of variation (GCV)

$$GCV=\frac{\sqrt{\sigma 2g}}{\bar{X}} * 100$$

Where, σ^2g = genotypic variance and

$\bar{X}$ = mean of the character

GCV and PCV values were categorized as low, moderate and high values as indicated by Deshmukh *et al*. (1986) as follows:

0-10% - Low

11-20% - Moderate

21% and above- High

Heritability (in the broad sense)

Heritability in the broad sense for quantitative characters was computed using the formula suggested by Allard (1999) as:

$$H^2=\frac{\sigma 2g}{\sigma 2P} \times 100$$

Where, H^2= heritability in the broad sense,

σ^2g = genotypic variance and σ^2P = phenotypic variance

Heritability percentage categorized as low, moderate and high as given by

Robinson *et al.* (1949).

0-30- low

31-60% -moderate

61% and above high

Genetic advance expected (GA)

The expected genetic advance (GA) was estimated in accordance with the formula described in Allard (1999) as follows:

$GA = K^* \sigma_p * H^2$

Where, GA = expected genetic advance,

H^2 = heritability in broad sense,

K = the selection differential at 5% selection intensity

$\sigma_{p=}$ phenotypic standard deviation on mean basis

Genetic advance (as percent of mean) was computed using the formula given below to compare the extent of predicted genetic advance of different traits under selection:

$$GAM = \frac{GA}{\overline{X}} \times 100$$

Where, GAM = Genetic Advance as percent of mean

$\overline{X}$ = grand mean of the trait

GA = genetic advance under selection

Genetic advance as percent of the mean was categorized as low, moderate and high following Johnson *et al.* (1995) as follows:

0-10% - Low

11-20%- Moderate

Above 20%- High

3.6.2. Molecular data analysis

Population Structure (Q-matrix) and Kinship (K-matrix) analysis

Population structure (Q-matrix) among the 160 sorghum accessions (152 landraces and 8 released varieties) was conducted using the model based Bayesian clustering algorism-software package "*STRUCTURE"* version 2.3.3 (Pritchard *et al.*2000a). By setting the number of k levels from 1 to 9 with five times repetition for each k, nine independent structure runs were performed with 100,000 burn-in time and 100,000 iterations of Markov chain convergence for each run. The plot of the average *log likelihood* values

over five runs for each *K* ranging the k-values from 1 to 9 showed that the *log likelihood* estimates increase progressively as *K* increases. The ad hoc criterion described by Evanno *et al.* (2005) was used to reliably detect the most probable number of subpopulations and the number of sub populations were found to be at k=4.

On this basis, four sub-populations were assumed in the association mapping. Thus, the population structure matrix (Q) for K = 4 was selected to be assigned as the proportion of membership for each accession and used by MLM in the TASSEL software for the structure-based association analysis. All STRUCTURE runs were performed using the admixture model with the option of correlated allele frequency between populations. Also, the measure of the degree of admixture, alpha, was allowed to be inferred from the data (Pritchard *et al.*, 2000a), and Lambda, the parameter of the distribution of allelic frequencies, was set to one.

The matrix of kinship coefficient comparing all pairs of the 160 lines using 39 SSR markers was calculated by the software package SPAGeDi as described by Loiselle *et al.* (1995). Negative kinship values between two individuals, indicating that there was less relationship than that expected between two random individuals, were changed to 0 and the diagonal was set to 2 (Pritchard *et al*., 2000a).

Linkage disequilibrium analysis

LD (R^2) between SSR markers /loci/ was evaluated using Tassel soft ware version 2.0.1. The LD was calculated using the statistical coefficient of determination (R^2) which is a measurement of correlation between a pair of variables (Shi *et al*, 2010). Alleles with frequencies less than 0.05 were not included for LD calculation.

Analysis of Marker- Trait Association

By fitting the population structure and kinship matrix into the model to avoid spurious associations, the trait marker association was evaluated by employing a mixed linear model (MLM) using "TASSEL" Software version 3.0.1. To achieve linear independence,

the structure matrix (Q-matrix) with one column less than the number of sub populations was used (Prichard *et al*., 2000a). The statistical model used for identifying SSR markers associated with traits was as follows:

$Y_{klmnf} = \mu + Qk + ML + Am(ML)Kn + \varepsilon_{kmn}$

where **μ** is the general mean, Qk is the fixed effect of *kth* subgroup of the population structure (Q-matrix), **M**L is the fixed effect of *Lth* marker, **A**m**(ML)K**n is the random effect of *mth* accession nested in the *Lth* marker associated with *nth* kinship coefficient , $\boldsymbol{\varepsilon}_{kmn}$ is the error. In this model Qk and ML represent fixed effects, and Am (ML) Kn and ε_{kmn} represent random effects.

Only markers with an allele frequency ≥ 5% were included in the association analysis. The significance of associations between loci and traits was based on an F-test with P values calculated by TASSEL at 5% significant level (Wang *et al*., 2011).

4. RESULTS AND DISCUSSIONS

4.1. Analysis of Variance

The analysis of variance (ANOVA) for the studied characters is presented in appendix Table 1. The result showed that there was a highly significant ($P < 0.0001$) difference among genotypes for all characters indicating wide variability in performance among the genotypes for these characters.

4.2. Mean Performance of the Landraces

The individual entry mean values and the average mean performance of the 160 sorghum accessions studied is presented in Table (4).

Table 4. The mean performance of the studied materials.

Entry	Accession name	50%FL	PHT	PWT	GWPP	PHI	THGT	NGPP
1	69046	96.33	1.84	55.86	43.5	0.78	28.67	1556.48
2	69057	94.33	2	46.5	47.81	0.79	38.33	1206.77
3	69094	90	1.34	58.1	48.13	0.83	29.5	1616.55
4	69096	95.33	1.33	49.48	37.69	0.8	35.67	1098.09
5	69105	100	2.68	60.58	50.89	0.84	25	2027.82
6	69183	114	2.61	105.15	83.95	0.8	23	3650
7	69192	101	2.41	60.34	48.21	0.8	34.67	1320.4
8	69210	94.5	1.94	64.34	48.15	0.74	22.5	2183.04
9	69227	119	2.15	86.2	76.33	0.88	30.5	2690
10	69231	101.5	2	94.6	81.89	0.87	42.5	1926.73
11	69241	114	2.11	101.15	80.8	0.8	35	2308.57
12	69286	87.67	2.11	65.52	52.91	0.81	32.33	1641.86
13	69371	86	1.69	54.88	46.54	0.85	36.67	1279.59
14	69391	81.67	1.67	47.29	40.21	0.85	37.33	1078.71
15	69392	100	2.17	95.81	70.93	0.84	30	2343.77
16	69442	80	2.09	42.59	33.61	0.79	28.33	1202.13
17	69492	104.33	1.76	62.84	49.8	0.79	28.33	1764.88
18	69452	94	2.48	38.72	30.12	0.78	28.33	1095.57
19	69494	98.5	2.08	37.74	28.42	0.74	26	1145
20	239210	98	2.4	82.1	69.98	0.85	45.5	1536.68
21	69516	100	1.74	44.43	27.88	0.71	22.67	1223.57
22	70068	110	1.66	43.57	35.65	0.82	18	1961.56
23	70069	112	1.82	54.05	44.57	0.82	20.67	2172.52
24	70075	88	2.77	60.11	49.06	0.82	28	1757.69
25	70189	91	1.86	49.39	38.83	0.79	30	1379.07
26	70216	105	2.8	45.99	31.2	0.74	25.33	1232.08
27	70301	112	2.3	55.93	36.28	0.66	19.67	1840.69
28	70306	98.33	1.96	53.64	39.33	0.75	29.67	1324.88
29	70537	101.5	2.24	86.81	77.63	0.89	34.5	2236.13
30	71021	85.33	1.77	44.52	37.19	0.84	22.33	1667.17
31	71370	114	1.8	46.62	38.65	0.82	21	1777.02
32	71418	86.33	2.54	54.88	42.48	0.77	24	1775.3
33	71421	91.67	2.13	53.14	51.95	0.84	24	1997.88
34	71422	75	1.89	51.84	40.14	0.77	27.33	1473.24
35	71425	128	1.81	42.95	32.73	0.76	16.5	1965
36	71522	89	2.24	65.25	54.1	0.83	30.67	1775.64
37	71536	90	2.16	34.41	26.66	0.76	22	1438.95
38	69468	98.5	2.1	58.83	44.75	0.76	24	1873.48

Table 4 (continued)

39	71562	88	2.11	36.81	29.42	0.8	28	1048.79
40	71565	97	1.87	62.59	46.42	0.75	31.33	1495.68
41	71570	86.67	1.75	41.87	32.38	0.77	28.67	1131.87
42	71590	94.67	2.03	46.73	38.89	0.84	35.5	1069.72
43	71625	85.67	1.78	54.85	41.44	0.75	26.33	1569.63
44	71657	91.67	2.03	63.84	51.53	0.8	27.67	1886.3
45	71739	90	2.3	45.25	36.78	0.81	28.33	1294.58
46	71748	99.67	1.89	61.84	44.21	0.72	21.5	2051.07
47	71788	93	2.45	44.18	35.94	0.83	24.5	1447.71
48	71795	98.5	2.05	35.38	25.32	0.71	29	797.59
49	71810	86.67	1.72	57.27	44.81	0.79	26.33	1619.14
50	72437	107	2.34	65.56	53.48	0.82	41.5	1463.41
51	72438	103	1.91	41.58	31.65	0.76	30.67	1034.23
52	72470	101	2.45	44.04	34.69	0.79	28.33	1258.2
53	72506	106	2.39	43.75	35.65	0.81	27.67	1229.31
54	72588	110	2.1	72.71	61.84	0.85	27.33	2262.84
55	72611	99.67	2.45	29.08	20.83	0.7	22.5	907.24
56	72997	93.33	2.16	50.62	40.97	0.81	34.5	1190.66
57	73067	106.67	2.82	96.33	84.4	0.87	36.5	2300.53
58	73068	100	2.66	83.2	67.2	0.81	28	2400
59	73339	88	2.12	54.28	39.89	0.73	24	1661.72
60	73341	90.5	2.44	37.8	28.34	0.74	22.67	1249.6
61	73637	100	2.37	66.75	51.97	0.78	32.5	1615.5
62	73641	96.33	2.34	62.81	49.93	0.8	42.33	1192.53
63	73646	76	2.26	41.5	31.16	0.74	24.67	1339.82
64	73762	87	1.92	43.17	35.61	0.82	38.67	914.97
65	73797	88.67	1.82	50.84	43.71	0.86	40.67	1069.07
66	73825	78.33	2.11	39.93	31.88	0.8	31	1042.71
67	73940	104	1.89	50.33	41.84	0.83	37	1145.85
68	73995	87.67	2.33	31.12	21.29	0.69	29.33	756.28
69	74097	90.5	1.83	53.72	43.91	0.79	45.33	983.55
70	74108	91	2.29	39.58	30.96	0.78	35.5	959.43
71	74115	94.67	2.5	49.8	38.89	0.79	30.67	1260.85
72	74262	96.33	2.7	48.5	39.79	0.8	29.33	1341.68
73	74268	90.33	2.03	45.15	36.6	0.81	30.67	1169.45
74	74698	110	2.55	69.73	56.88	0.82	32.67	1765.3
75	74761	99	2.04	45.49	35.9	0.78	34	1053.27
76	74766	122	3.44	65.36	47.58	0.73	23	2063.57
77	74935	89	1.96	58.86	49.75	0.85	39.67	1261.02
78	74999	119	2.59	28	21.04	0.75	17	1237.65
79	75000	111.33	2.47	63.84	47.85	0.75	35	1370.82
80	75033	103	1.99	52.15	40.28	0.77	31.33	1278.19

Table 4 (continued)

81	75065	98	2.25	51.41	42.59	0.82	29	988.28
82	75066	90.5	2.78	65.96	55.74	0.84	33.33	1672.35
83	75129	110	1.79	85	65	0.76	31	2096.77
84	75216	110	1.97	50.99	39.41	0.77	40.33	979.18
85	75220	128	2.15	49.55	31.7	0.64	13	2438.46
86	75353	102	2.28	36.54	27.74	0.76	19	1460
87	75455	91.33	1.69	51.41	39.98	0.78	29.67	1683.39
88	200546	84.67	1.72	47.52	37.46	0.8	26	1499.7
89	200547	93.33	2.45	59.4	46.57	0.76	33	1550.14
90	201299	100	1.56	45.95	36.23	0.79	36.33	991.96
91	201302	90	1.55	48.06	40.13	0.84	33.5	1087.67
92	201318	82.67	2.42	54.51	34.97	0.65	34.33	1017.34
93	201349	78.33	1.94	64.71	52.31	0.8	36	1457.84
94	201491	96.33	2.61	46.85	39.4	0.84	35.67	1069.04
95	201499	100	2.45	61.08	48.8	0.8	27.33	1782.07
96	201502	98.33	2.36	44.41	37.03	0.83	35.5	1038.6
97	204776	103	2.4	61.67	50.36	0.82	24.67	2064.41
98	206910	110	2.05	41.44	30.44	0.71	18.67	1566.87
99	206921	84	1.77	40.92	33.9	0.82	30.33	1113.17
100	206923	82.67	2	47.1	36.22	0.77	32	1139.88
101	206936	85	2.01	33.23	24.97	0.74	28.33	862.99
102	206943	80.67	1.73	34.17	27.2	0.79	26.67	1014.4
103	206952	91.5	1.94	67.27	54.15	0.8	28	1945.76
104	206962	110	1.93	35.18	26.23	0.73	17.33	1477.11
105	206963	98	2.27	44.33	34.09	0.77	30.5	1100.27
106	207573	87.67	1.85	46.98	36.39	0.75	38	987.65
107	213008	110	2.2	77.38	61.27	0.79	29	2188.13
108	213386	101	2.53	57.3	44.2	0.77	31	1425.81
109	213393	104.33	2.07	54	46.9	0.87	42	1116.67
110	214046	86	1.66	62.57	51.38	0.82	34.67	1464.79
111	214067	83.5	1.93	39.82	31.83	0.8	31	999.84
112	215053	88	2.38	55.85	43.78	0.78	31	1414.02
113	215054	104	1.74	67.05	53.8	0.79	30	1807.9
114	215330	92.67	1.59	41.02	34.21	0.83	31.67	1082.36
115	216742	92	1.61	60.52	51.39	0.85	34.67	1491.59
116	216824	110	2.31	28.43	22.97	0.81	40	574.17
117	220251	84	1.96	44.65	34.14	0.77	38.67	888.61
118	220252	84.33	1.76	37.49	30.8	0.82	33	945.85
119	220267	75	1.81	48.62	40.02	0.82	30.33	1211.24
120	222882	86	1.58	56.99	47.1	0.82	33.67	1407.01
121	222884	80.33	1.38	52.56	44.93	0.85	32.67	1394.96
122	226080	104.33	1.78	64.09	51.83	0.8	34.67	1484.26

Table 4 (continued)

123	229232	98	1.93	41.78	32.89	0.78	37.67	870.14
124	229238	110.33	2.76	55.15	46.95	0.85	28	1637.6
125	229843	93.33	2	40.32	31.49	0.78	30.33	1044.72
126	229844	91	2	41.43	31.24	0.75	33	945.55
127	231200	105	2.18	56.69	42.52	0.75	36.67	1164.64
128	231230	106.67	1.98	74.55	60.45	0.81	30.67	1985.47
129	235459	89.33	1.8	64.42	52.26	0.81	38	1290.28
130	235469	96.33	1.69	49.83	38.52	0.77	33.67	1143.26
131	235913	86	2.04	39.69	33.13	0.8	26.67	1210.4
132	235921	112	1.99	36.85	27.35	0.74	29	943.1
133	235922	110	2.04	51.49	39.87	0.77	27.33	1446.05
134	69238	103.33	2.27	55.93	46.81	0.84	37.33	1215.75
135	69306	100.67	2.36	41.15	32.8	0.83	41	800.12
136	210902	110	2.82	19.08	8.68	0.45	23	377.39
137	210922	101	2.16	128.76	101.09	0.79	39	2470.64
138	216739	86.5	1.5	87.4	73.2	0.84	28	2319.8
139	217685	93.33	1.84	63.66	47.88	0.75	27.67	1798.95
140	217705	92	1.88	56.35	46.43	0.81	33.67	1363.14
141	229230	100	2.23	32.2	25.07	0.77	40.33	615.77
142	231191	97	2.06	56.97	44.96	0.79	39.33	1136.28
143	237274	94.33	2.07	38.86	31.65	0.82	42.67	746.14
144	237287	89.33	1.71	42.3	34.56	0.82	35.33	978.95
145	237289	90	1.38	55.72	44.6	0.79	35	1261.62
146	235820	89.33	1.71	46.95	37.29	0.79	24.67	1520.32
147	239232	93.67	2.26	78.59	64.72	0.82	30	2143.3
148	241183	104.33	2.28	91	66	0.72	28	2405.5
149	241227	128	2.2	79.73	64.5	0.81	29	2235.4
150	241705	85.67	1.4	51.82	41.79	0.81	33.67	1243.17
151	241728	94	2.17	52.19	48.46	0.54	27.67	1585.78
152	243681	90.5	2.2	42.64	38.44	0.52	43	890.96
153	76T1-23	83.67	1.07	38.52	29.07	0.75	18.33	1563.79
154	B35	76.33	0.92	43.47	37.61	0.78	26.67	1477.31
155	Baji	103	1.22	62.59	48.71	0.78	23	2117.83
156	Birmash	111.33	1.49	68.89	53.03	0.77	22	2414.24
157	E36-1	81	1.46	56.73	47.65	0.84	35.33	1351.33
158	Gambela	85.67	1.58	62.05	61.13	0.81	28	1942.58
159	Meco-1	85.5	1.49	56.73	47.47	0.83	33.33	1417.59
160	Teshale	93.67	1.34	52.18	41.67	0.8	32.5	1280.25
	Average	96.31	2.05	54.44	43.52	0.79	30.52	1462.96
	$LSD_{0.05}$	7.32	0.38	15.18	11.65	0.07	5.053	428.56
	CV (%)	4.726	11.41	17.35	16.65	5.474	10.3	18.22

LSD = least significant difference, **CV**=coefficient of variation, **50%FL**= days to 50% flowering, **PHT**=plant height, **PWT**= panicle weight, **GWPP**=grain weight per panicle, **PHI**=Panicle harvest index, **THGT**=1000 seed weight, **NGPP**=number of grains per panicle

Days to 50% Flowering (days)

Days to 50% flowering for entry means ranged from 75 to 128 days with mean value of 96.31 days across entries. Two genotypes (71422 and 220267) showed the minimum value of days to 50% flowering (75 days) where as three genotypes (71425, 75220, and 241227) showed the maximum value (128 days) of days to 50% flowering. The average value for released varieties was 90.02 days. A total of 82 accessions including all released varieties except Baji (103 days) and Birmash (111 days) showed mean values less than the grand mean (96.31).

Plant height (m)

The mean plant height of the genotypes ranged from 0.92 (B-35) to 3.44 (74766) with average value of 2.05 for all entries. The average for released varieties was 1.32. Eighty five accessions had plant stature below the grand mean (2.05).

Panicle weight (g)

The panicle weight ranged from 19.08 to 128.76 across entries. Among the genotypes, 210902 and 210922 showed the minimum (19.08) and maximum (128.76) values, respectively. The average panicle weight for all entries was 54.44. The average panicle weight for released varieties was 55.15. In this study, the panicle weight of 70 accessions was above the grand mean (54.44) including Birmsh (68.89), Gambella (62.05), Meco-1 (56.73), and E36-1 (56.73).

Grain weight per panicle (g)

Grain weight per panicle ranged from 8.68 (210902) to 101.09 (210922) with an average value of 43.52. The first five best performing sorghum lines were 210922, 73067, 69183, 69231, and 69241 in that order. The average grain weight per panicle among the released varieties was 45.79. The mean values of grain weight per panicle for 72 accessions were greater than the grand mean (43.52). The remaining 88 accessions (including 76T1-23, B-35, and Teshale) had mean values below the grand mean. Gambella (61.43), Birmash (53.03), Baji (48.71), E36-1 (47.65), and Meco-1 (47.47) were among the accessions

whose means were greater than the grand mean (43.52) indicating the potential of these released varieties to be continued in production. However, the first 17 best performers among the entries were the landraces. The superior performance of the land races under water stress condition over that of the released lines is an indication for the presence of wide genetic variability among the studied land races from which elite sorghum lines that can produce better yield than the existing released varieties could be developed for moisture stressed environments. Similar results of superior performance of the land races over that of the released lines for grain yield were reported in sorghum by Ghebru *et al*. (2002) and Tesfaye *et al*. (2011) who evaluated 28 and 200 sorghum land races, respectively and declared the existence of wide genetic variability among sorghum land races for future use in breeding programs.

In the present study, the majority (55.6 % or 40 accessions) of the land race accessions which produced grain yield per panicle above the grand mean had medium to relatively longer days to 50% flowering and taller plant height (greater than the grand mean for both traits). The superior performance of these genotypes as compared to the remaining early flowering ones might be due to the fact that medium to late maturing genotypes have the genetic potential to produce more assimilating organs (eg. Leaf area) and accumulate higher biomass which leads to increased grain filling duration and more grain yield. On the contrary, early flowering cultivars usually produce fewer assimilating organs which results in less production of assimilates which in turn leads to less grain yield (Habyarrimal *et al*., 2004). Studies showed that early maturing sorghum genotypes yielded less biomass than medium to late maturing sorghum genotypes (Habyarrimal *et al*., 2004).

As crops rely on remobilization of stored carbohydrates to combat the influence of drought that occurs especially at reproductive stage, genotypes with better biomass accumulation efficiency before the stress occurs will produce relatively better grain yield than less efficient ones in biomass accumulation (Beheshti and Fard, 2010). The better biomass accumulation capacity of the land races might also be, in part, due to their better potential to extract water from the soil under stressed conditions. According to Reynolds *et al*. (2007), land races have generally better water uptake characteristics than the released lines in water stressed conditions. Due to their adaptation to harsh environmental conditions, the local landraces (LLRs) might be more tolerant to various abiotic stresses including water

stress as compared to modern varieties. Yong'an *et al*. (2010) reported that varieties with higher ability of soil moisture capture were the varieties with higher grain yield and drought tolerance.

Thus, the superior performance of the land races for grain yield over that of the released lines might be in part due to their superior genetic potential to accumulate higher biomass than the released varieties (even in the absence of water stress for both) and in part due to their better efficiency of extracting moisture from the soil and assimilation of more photosynthates than the released lines in the face of stress. The question of "which factor contributes more for higher grain yield performance by the land races" requires further detailed study using two water regimes. Similar results of higher grain yield performance by the land races were also reported by Maposa *et al.* (2010) using sixteen sorghum lines (two released lines and 14 land races) and Reynolds *et al*. (2007) using four wheat lines.

Thousand seed weight (g)

For thousand seed weight, accession no.75220 showed the minimum value (13.00) where as the maximum value (45.50) was observed by 239210 with an average value of 30.52 for all entries. The average value of thousand seed weight for released varieties was 27.40 g. A total of 77 accessions, including E36-1(35.33), Meco-1(33.33), and Teshale (32.5), had mean values greater than the grand mean (30.52). The differences among the genotypes for thousand seed weight might be due to the differences in grain filling duration. Reductions of sorghum grain yield due to drought stress before anthesis are related to decreases in grain number, while a decrease in thousand seed weight and a smaller grain size is responsible for yield losses when water deficits occurs after anthesis (Manjarrezsandoval *et al*., 1989).

Panicle harvest index

The minimum and maximum values for panicle harvest index were 0.45 and 0.89 from accessions 210902 and 70537, respectively with an average value of 0.79 for all entries studied. The average value for released varieties was 0.80. A total of 80 accessions including E36-1(0.84), Meco-1(0.83), Gambella (0.81), and Teshale (0.8) had mean value of panicle harvest index greater than the grand mean (0.79)

Number of grains per panicle (number)

Of all accessions studied, the minimum value for the number of grains per panicle was found to be 377.39 and the maximum value was 3650.00 and the accessions showed these values were 210902 and 69183, respectively. The average value for all entries was 1462.96. The average value for the released varieties was 1695.62. A total of 69 accessions including Birmash (2414.24), Baji (2117.83), Gambella (1942.58), 76T1-23 (1563.79), and B-35 (1477.31) had mean number of grains per panicle greater than the grand mean (1462.96).

4.3. Estimates of Variability

Phenotypic and genotypic coefficient of variation

The extent of Variability with respect to range, phenotypic and genotypic variance and coefficient of variations (phenotypic and genotypic coefficient of variations) for the seven characters studied is given in (Table 5).

The phenotypic coefficient of variation (PCV) was relatively higher than genotypic coefficient of variation (GCV) for all characters (Table 5) indicating the influence of environment in expression of the traits. Similar results of higher phenotypic coefficient of variation over that of genotypic coefficient of variation were reported in sorghum by Khan *et al*. (2005), Sundaresha (2006), Tariq *et al*. (2007), and Shinde *et al.* (2010).

All characters showed moderate to high phenotypic coefficient of variations (PCV) that ranged from 11.04 (50%Fl) to 35.23 (NGPP). Moderate PCV (%) values were recorded for plant height (20.13), panicle harvest index (12.66) and, days to 50 % flowering (11.04), while high values were recorded for NGPP (35.23),GWPP (33.59), PWT (32.12), and THGWT (21.43). Similar results of high PCV were reported in sorghum by Reddy *et al*. (1996) and Can and Yoshida (1996).

The genotypic coefficient of variations (GCP) ranged from 5.21 (PHI) to 30.15 (NGPP). High values of GCV were observed for NGPP (30.15), GWPP (29.17), and PWT (27.03).

Moderate GCV values were observed for THGT (18.78) and PHT (16.32) whereas low GCV values were recorded for PHI (5.21) and 50%Fl (9.97). The low GCV values for PHI and 50%Fl indicate the presence of limited improvement through selection for these characters. The high values of GCV for NGPP, GWPP, and PWT are evident for the presence of high genetic variability among the entries that in turn offers good scope for genetic gain by selection. Similar results of high GCV for NGPP (30.15), GWPP (29.17), and PWT (27.03) were reported in sorghum by Can and Yoshida (1996), Sundaresha (2006), and Shinde *et al.* (2010).

In the present study, high values of both GCV and PCV were simultaneously observed for number of grains per panicle, grain weight per panicle and panicle weight indicating the presence of high variability among the genotypes in terms of phenotype and genotype which suggests positive response to selection. Similar results of both high GCV and PCV were reported by William *et al.* (1987) while studying the effect of environment on yield components of sorghum using three sorghum lines.

Table 5. Variability measures viz. range, phenotypic and genotypic variance and coefficient of variations (phenotypic and genotypic) for the seven characters studied

Characters	Range	σ^2g	σ^2p	σ^2e	GCV (%)	PCV (%)	H^2 (%)	GA	GAM (%)
Days to 50%flowering	75-128	92.21	112.93	20.72	9.97	11.04	81.65	17.88	18.57
Plant height	0.92-3.44	0.11	0.17	0.05	16.32	20.13	67.17	0.57	27.86
Panicle weight	19.08-128.76	216.61	305.85	89.22	27.03	32.12	70.83	25.52	46.87
Grain weight per panicle	8.68-101.09	161.13	213.65	52.52	29.17	33.59	75.42	22.71	52.19
Panicle harvest index	0.45-0.89	0.0016	0.01	0.0001	5.21	12.66	71.25	0.15	18.58
1000 grain weight	13-45.5	32.87	42.75	9.88	18.78	21.43	76.89	10.36	33.94
Number of grains per panicle	377.4-3650	194566	265.649	71083.33	30.15	35.23	73.24	777.64	53.16

σ^2g= Genotypic variance, σ^2p =phenotypic variance, σ^2e = environmental variance, GCV= genotypic coefficient of variation, PCV=phenotypic coefficient of variation, H2 = heritability in broad sense, GA = expected genetic advance, and GAM (%) = genetic advance as percent of the mean.

The present study also revealed that the magnitude of the difference between the phenotypic and genotypic coefficient of variation was found less in the studied characters except for PHI (7. 45) with the least difference for 50%FL (1.07), followed by THGT (2.64), PHT (3.81), GWPP (4.43), NGPP (5.08), and PWT (5.09). Panicle harvest index showed relatively wider gap between PCV and GCV (7.45) indicating the relatively higher influence of the environment on its phenotypic expression. The more or less equivalent magnitude of PCV and GCV for the rest of the characters indicated that these characters were less influenced by environmental fluctuations and the genotype had significant effect on their phenotypic expression indicating the possibility of selecting the superior genotypes based on the expression of phenotypes for further improvement. In agreement with the present study, Manonmani *et al.* (2002) using red grain sorghum genotypes and Adissu (2011) using 226 RILs of sorghum genotypes reported the presence of narrow gap between PCV and GCV for characters viz., THGT, PHT, NGPP, PWT, and GWPP.

Heritability

Broad sense heritability (H^2), expected genetic advance (GA) and genetic advance as percent of the mean (GAM) for the studied characters is given in Table 5 and outlined below.

Effective selection can be achieved when heritability estimates are high (Shinde *et al*., 2010). In the present study, the estimates of heritability (in broad sense) were high for all characters ranging from 67.17 for plant height to 81.65 for days to 50% flowering (Table 5). High heritability values indicate lesser influence of the environment on the phenotypic expression of the characters and suggested good scope of genetic improvement through selection. In agreement with the present study, high heritability was previously reported in sorghum by Manonmani *et al.* (2002) for grain yield, by Chaudhary and Arora (2001) for panicle weight and grain yield per plant, by Tariq *et al*. (2007) for harvest index, and by Shinde *et al* (2010) for days to 50% flowering, plant height, number of grains per panicle, and grain yield per plant.

Genetic advance and genetic advance as percent of mean

The estimates of genetic advance as percent of mean were moderate to high ranging from18.57 (50%Fl) to 53.16 (NGPP) for all characters. High values of genetic advance as percent of mean were recorded for NGPP (53.16), GWPP (52.19), PWT (46.87), THGT (33.94), and PHT (27.86). Moderate values were observed for 50% FL (18.57) and PHI (18.58). However, all of the studied characters had high heritability (Table 5). High heritability along with high genetic advance as percentage of mean for NGPP, GWPP, PWT, THGT, and PHT in the present study indicates the role of additive genes in the expression of these characters suggesting high possibility of improvement upon selection.

Although effective selection can, in fact, be achieved when heritability estimates are high, Johnson *et al* (1955) noted that heritability in conjunction with genetic advance is more effective and reliable in predicting the resultant effect of selection than heritability alone. High heritability could not always be associated with high genetic advance (Shinde *et al*., 2010). if heritability is mainly due to non additive effects (dominance and epistasis), the genetic advance will be low, where as if heritability is due to additive effects it would be associated with high genetic advance (Panse, 1957). Thus, heritability values alone may not provide clear predictability of selections made. Hence, heritability values along with estimates of genetic advance would be more reliable than heritability alone (Johnson *et al.,* 1955). Also, heritability values indicate only the magnitude of inheritance of the quantitative character, while genetic advance is helpful in formulating the selection procedure to be adopted (Adisu, 2011).

Thus, in the present study, the presence of high heritability coupled with high genetic advance as percent of mean for NGPP, GWPP, PWT, THGT, and PHT indicated that these characters are most probably controlled by additive gene action and hence these traits can be fixed by selection. Similar results of high heritability coupled with high genetic advance as percent of mean were reported in sorghum for plant height, one thousand seed weight, and seed yield per panicle by Deepalakshmi *et al*. (2007) and Shinde *et al.* (2010) using sixteen white grain sorghum genotypes and 120 F_6 sorghum lines, respectively.

In this study, plant height and days to 50% flowering showed high heritability coupled with moderate genetic advance. The decreased in genetic advance (moderate) indicates the

influence of environment and hence these traits are less amenable for selection as they are most likely to be controlled by both additive and non-additive gene actions. Similar result of high heritability coupled with moderate genetic advance for plant height and days to 50% flowering was reported by Kamatar *et al*. (2011) using 60 sorghum lines.

4.4. Population Structure Analysis

The present study showed that the 160 sorghum accessions contained four distinct sub groups as shown in Fig. 2 and 3.

In the present study, the plot of the average log likelihood values over five runs for each K (ranging the k-values from 1 to 9) showed that the log likelihood estimates increase progressively as K increases (Fig. 1) and did not show a clear peak to determine the true K (number of sub groups). The ad hoc criterion described by Evanno *et al.* (2005) was used to reliably detect the most probable number of sub-populations and the number of sub-populations were found to be 4 (Fig.2).

The first group, G-1 (Fig.3) consisted of 52 accessions of which 20 were from Amhara, 16 from Oromia, 9 from Tigray, 1 from Afar and other 5 accessions which their geographical origin was not available. The second group G-2 (Green) which is the smallest of the four sub populations consisted of 13 accessions: 6 were from Amhara, 4 from Oromia, 2 from SNNP, and 1 from Tigray. The third group, G-3(Blue) consisted of 24 accessions of which, 9 were from Gambella, 6 from SNNP, 3 from Amhara, 2 from Oromia and 4 released varieties. The fourth group G-4 (Yellow) which is the largest of the four identified sub populations consisted of 71 accessions of which 19 were from Oromia, 18 from Amhara, 11 from Tigray, 10 from SNNP, 4 from Oromia, 2 from Eritrea, 1 from Tigray, 2 accessions which their origin was not available and 7 are released varieties.

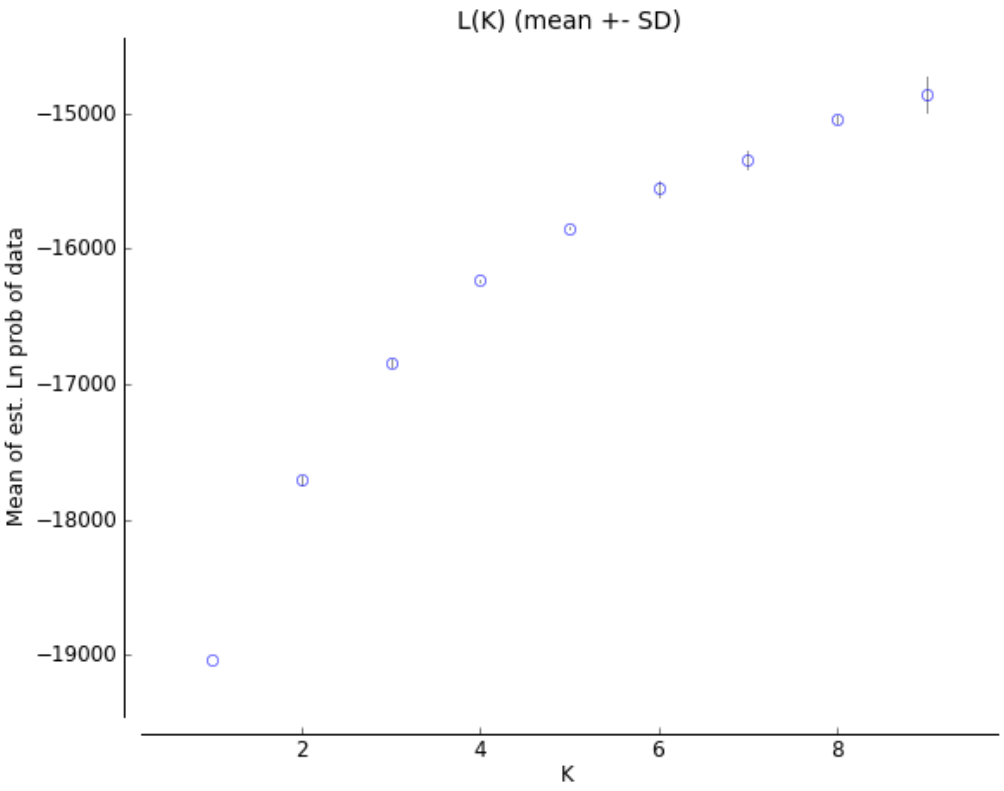


Figure 1. Posterior probability, ln $P(D)$, of the data as a function of the number of subpopulations (k), where k was allowed to range from 1 to 9. Circles represent the average of the five independent runs for each value of k.

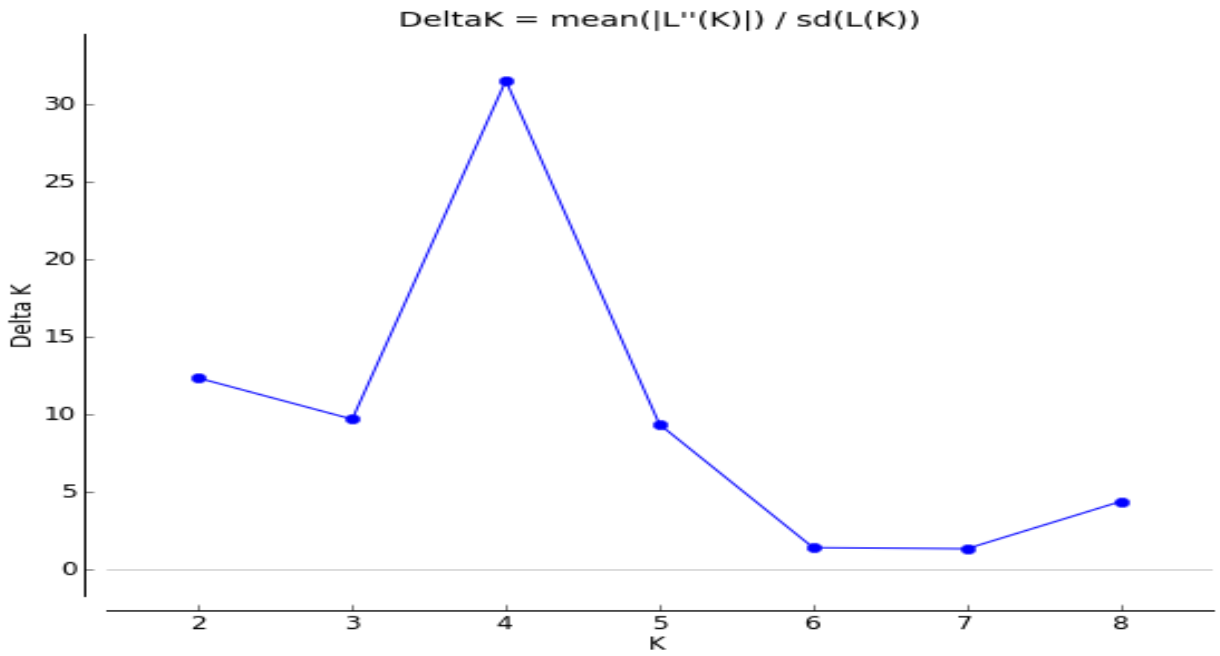


Figure 2. Values of K (x-axis), with its modal value used to detect the true K (y-axis) of four groups (K =4).

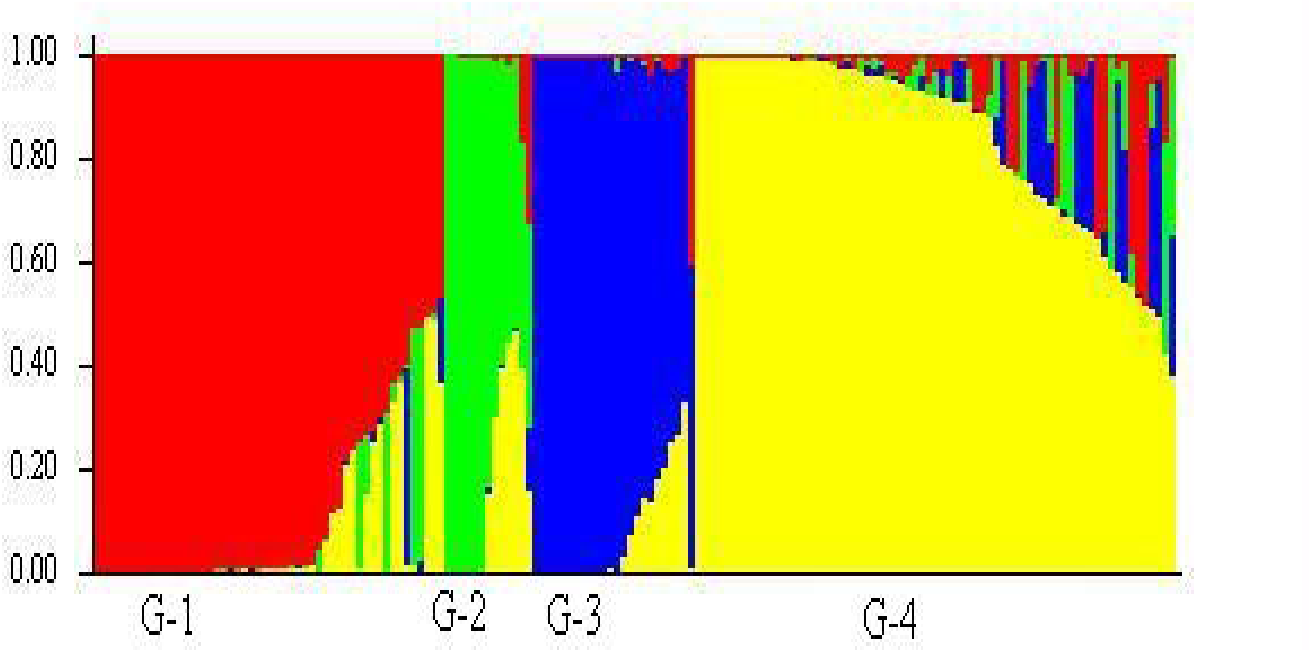


Figure 3. Population structure in the studied entries. The subpopulations obtained with K= 4 are represented by different colors as indicated at the bottom (G-1=red, G-2=green, G-3=Blue, and G-4 = yellow).

Generally, groups were composed of accessions from different areas of collection. The distribution of accessions into the four groups without reflecting their region of origin might be an indication of the presence of wide variations among accessions within the regions as well as lack of strong regional differentiation which might be due to gene flow between the regions. This result is in agreement with Atnafu (2010) who conducted a diversity study of 205 sorghum accessions using 39 SSR markers including the 160 sorghum accessions used in this study. In that study it was reported that sorghum accessions were not clustered according to their area of collection; rather clusters were from different areas of collection. Similar results that showed lack of clustering based on the site of collection in sorghum accessions were also reported by Nemera *et al.,* (2006), Perumal *et al.,* (2007), and Alemu (2009).

4.5. Level of Linkage Disequilibrium

The present study did not show a clear trend on linkage disequilibrium decay in the studied germplasms as shown below in fig.4.

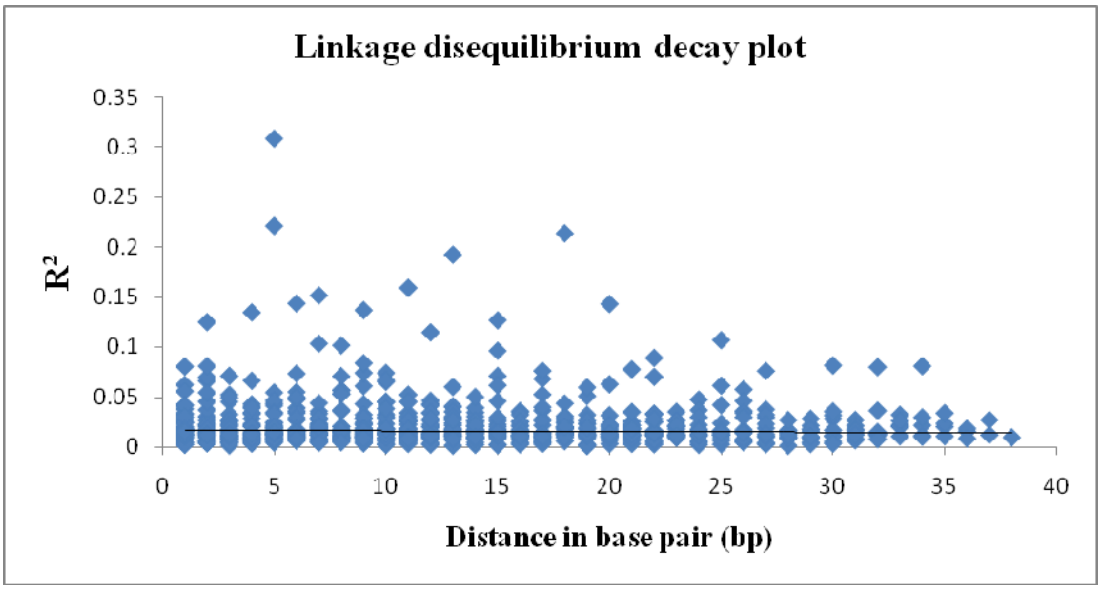


Figure 4. Linkage disequilibrium decay plot generated by 39 SSR markers

In this study, all 39 SSR markers were used to estimate the presence of LD in all accessions. After filtration of the data to exclude markers with less than 5% allele frequencies from the analysis, there were 325 pair wise locus comparisons for all

accessions and the majority of loci pairs (67.077%) were independent loci (non-significant). In all accessions, 107 loci pairs (32.92%) had a significant ($p < 0.05$) mean LD of 0.19, with an $R^2 > 0.2$ for 33 evaluated loci pairs.

The present study did not show a clear trend on linkage disequilibrium decay (Fig.4) and no clear conclusions can be made regarding the decay of LD. This result might be explained by low number of markers used in this study. Similar results were reported by Shehzad *et al*. (2009) using 107 sorghum accessions and 98 SSR markers and Li *et al* (2010) using 26 sorghum inbred lines.

4.6. Association Mapping

In this study, a total of 10 significant marker-trait associations ($P \leq 0.05$) were detected (Table 6) and the phenotypic effect of SSR marker alleles on the associated characters and the number of accessions carrying each significantly associated marker allele were identified (Table 7).

Table 6. Associations between SSR markers and six agronomical traits.

Trait	Marker	Chr.	Position(Mb)	F-value	p-value	Marker R2
GWPP	gpsb123	C8	52.282	5.34565	0.0236	0.07626
PWT	gpsb123	C8	52.282	4.46473	0.03802	0.06574
PHI	gpsb123	C8	52.282	4.4633	0.03805	0.04638
PHT	mSbCIR223	C2	4.657	4.23342	0.04232	0.04077
PHT	mSbCIR248	C5	4.746	5.79222	0.01803	0.05389
THGT	Xcup53	C1	72.905	4.22234	0.04333	0.03798
THGT	Xtxp114	C3	60.794	4.72867	0.0313	0.02642
PHT	Xtxp145	C6	49.285	10.4089	0.0021	0.17757
50%FL	Xtxp145	C6	49.285	4.12649	0.04697	0.08347
50%FL	Xtxp278	C7	51.120	4.14561	0.04426	0.04183

GWPP=grain weight per panicle, PWT=panicle weight, PHI=panicle harvest index, PHT=plant height, THGT=thousands grain weight, 50FL=Days to 50% flowering. Only SSR markers with a significant marker-trait associations are reported ($P < 0.05$). The *P*-value determines whether a QTL is associated with a marker, and the marker R^2 evaluates the magnitude of the QTL effects (percentage of total variation explained by the marker).

The 10 significant marker-trait associations were identified using 7 different SSR markers for six agronomical characters (50Fl, PHT, PWT, GWPP, THGT, and PHI), with R^2 ranging from 2.6 % (Xtxp114 with THGT) to 17.76 % (Xtxp145 with PHT) as briefly out lined below.

Days to 50% flowering

Two SSR markers (Xtxp145 & Xtxp278) having a significant association (P≤0.05) with days to 50% flowering were detected on chromosome 6 and 7, respectively. Xtxp145 had an effect of explaining 8.35% of the total phenotypic variation, where as Xtxp278 had an effect of 4.18% of the total phenotypic variation. SSR markers linked to QTLs that control flowering time in sorghum were previously reported on chromosome six by Mannai *et al*. (2011) using association mapping and on chromosome 7 by Sirinivas *et al.*(2009) and Shiringani *et al*. (2010) using conventional QTL mapping.

Plant height

Three loci (mSbCIR223, mSbCIR248, Xtxp145) with a significant association (P≤0.05) with plant height were detected on chromosome 2, 5 and 6, respectively. Marker mSbCIR223 had an effect of 4.1% of the total phenotypic variation; mSbCIR248 had an effect of 5.39 % of the total phenotypic variation, whereas Xtxp145 had an effect of 17.76 % of the total phenotypic variation. Wang *et al*. (2011), using pool based genome wide association mapping, reported four SSR markers that were closely associated with plant height on chromosomes 2, and 6. Similarly, Sirinivas *et al*. (2009), using conventional method confirmed the presence of QTLs for plant height on chromosome 6 and 7 in sorghum.

Panicle weight, grain weight per panicle, and panicle harvest index

Locus gpsb123 showed simultaneous significant associations (P≤0.05) with three characters, namely panicle weight, grain weight per panicle, and panicle harvest index on chromosome 8. This locus had an effect of explaining 6.6 %, 7.6 %, and 4.64% of the total phenotypic variation for panicle weight, grain weight per panicle, and panicle harvest index, respectively.

Thousand seed weight

Two loci (Xcup53 and Xtxp114) on chromosome 1 and chromosome 3, respectively, showed significant association ($P \leq 0.05$) with thousand seed weight. Xcup53 had an effect of 3.8 % of the total phenotypic variation, whereas Xtxp114 had an effect of explaining 2.64 % of the total phenotypic variation. A QTL controlling seed weight was previously reported on chromosome one by Sirinivas *et al.* (2009) using conventional QTL mapping.

The present study also showed that in each of the identified SSR marker locus there were two genotypes having variant alleles in the studied accessions (Table 7). Most of the two genotypes at each locus had different magnitudes on the expression of the phenotype. As shown in table 7, for example, for Xcup53 on chromosome 1, there were two genotypes (182:182 and 182:186) which were significantly associated to thousand seed weight (THGT). At this marker locus, there were 33 lines with genotype 182:182 and 48 lines with genotype 182:186. For the trait THGT, the difference between the two genotypes (182:182 and 182:186) was 3.5836. In other words, the presence of allele 182 in its homozygous state (182:182) increased the weight of thousand seeds by 3.5836 in 33 accessions compared to its heterozygous state (182:186) in 48 accessions. Similarly, for Xtxp114 on chromosome 3 at position 60.794 Mb which also linked to THGT, there were two genotypes (231:231 and 233:233). For this marker there were 37 accessions with genotype 231:231 and 112 accessions with genotype 233:233. For this trait (THGT), the difference between the two genotypes (231:231 and 233:233) was 2.37984. In the same way, Xtxp145 on chromosome 6 which associated simultaneously with two traits: days to 50% flowering and plant height, had two homozygous genotypes (214:214 and 212:212). There were 46 accessions with genotype 214:214 and 15 accessions with genotype 212:212. The difference between the two genotypes (214:214 and 212:212) was 7.19 for days to 50% flowering and 0.39742 for plant height. In other words, the presence of marker allele 214 in its homozygous form at this position decreased the days to 50 % flowering in 46 lines by 7.19 and the plant height by 0.39742 compared to its variant allele in homozygous form (212:212) in 15 accessions.

Table 7. The phenotypic effect of marker alleles at loci associated with traits and the number of accessions carrying each marker allele in the studied sorghum accessions.

Characters	Marker	Chr.	Pos(Mb)	Genotype(bp)	Effect	Observation
50%FL	Xtxp145	C6	49.285	214:214	-7.1919	46
50%FL	Xtxp145	C6	49.285	212:212	0	15
50%FL	Xtxp278	C7	51.12	248:248	4.92614	62
50%FL	Xtxp278	C7	51.12	242:248	0	48
GWPP	gpsb123	C8	52.282	290:290	-15.866	44
GWPP	gpsb123	C8	52.282	292:292	0	34
PHI	gpsb123	C8	52.282	290:290	-0.063	44
PHI	gpsb123	C8	52.282	292:292	0	34
PHT	mSbCIR223	C2	4.657	105:111	0.17516	40
PHT	mSbCIR223	C2	4.657	105:105	0	62
PHT	mSbCIR248	C5	4.746	91:91	-0.2505	54
PHT	mSbCIR248	C5	4.746	101:101	0	46
PHT	Xtxp145	C6	49.285	214:214	-0.3974	46
PHT	Xtxp145	C6	49.285	212:212	0	15
PWT	gpsb123	C8	52.282	290:290	-17.307	44
PWT	gpsb123	C8	52.282	292:292	0	34
THGT	Xcup53	C1	72.905	182:182	3.5836	33
THGT	Xcup53	C1	72.905	182:186	0	48
THGT	Xtxp114	C3	60.794	231:231	-2.3798	37
THGT	Xtxp114	C3	60.794	233:233	0	112

50FL=Days to 50% flowering, PHT=plant height, PWT=panicle weight, GWPP=grain weight per panicle, PHI=panicle harvest index, and THGT=thousands grain weight.

Similarly, Xtxp278 on chromosome 7 which was found to be significantly linked to days to 50% flowering had two genotypes (248:248 and 242:248) for the studied accessions. There were 62 accessions with genotype 248:248 and 48 accessions with genotype 242:248. For days to 50% flowering, the difference between the two genotypes (248:248 and 242:248) was 4.926136 days.

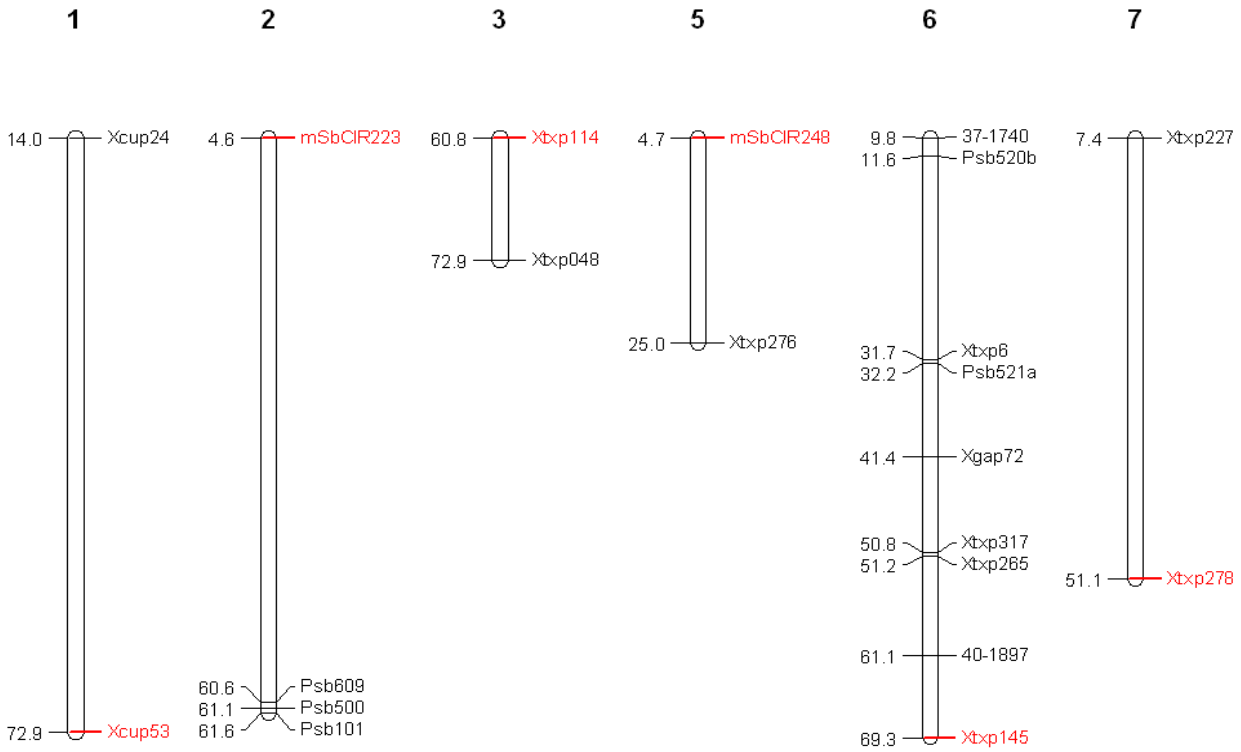
In the same way, gbsp123 on chromosome-8, which was found to be significantly associated with panicle weight and grain weight per panicle, had two genotypes (290:290 and 292:292). There were 44 accessions with genotype 290:290 and 34 accessions with 292:292. The presence of allele 290 in its homozygous form on this locus decreased panicle weight and grain weight per panicle by 17.307 and 15.866, respectively, in 44 accessions compared to its homozygous variant allele (292) in 34 accessions for both traits. The difference of the effect on the phenotype between the two genotypes of mSbCIR223 on chromosome 2, mSbCIR248 on chromosome 5, and Xtxp145 on

chromosome 6 which were found to be linked to plant height and between the two genotypes of gpsb123 on chromosome 8 to panicle harvest index ,was negligible.

However, it should be noted that it is not actually the variant marker allele itself which causes the decrease or increase in the expression of the phenotype in any trait of interest. Rather, it is suspected that there is a causative gene that is tightly linked to the variant marker allele which is responsible to cause a decrease or an increase in the expression of the phenotype in any trait of interest. Thus, by following the variant marker allele that is tightly linked to the causative gene, it is possible to follow the effect of the causative gene on the phenotype of the lines under study for the trait of interest.

Generally, most of the variant alleles on the identified SSR marker loci had differences in magnitude of their effect to the phenotype of the trait under study. Some of the variant alleles had an increasing effect to the expression of the phenotype while others had a reducing effect, as stated above. For example, in Xtxp145, allele 214 in its homozygous form (214:214) had a decreasing effect of days to 50% flowering by 7.19 days in 46 accessions and the plant height by 0.39742 meters in 15 accessions compared to its variant allele in homozygous form (212:212). This phenomenon might have useful application in molecular breeding. For example, if an interest arises to develop a variety of a grain cereal having a relatively short stature and earliness in flowering, a statistically linked marker allele with a reducing effect to plant height and days to 50 % flowering will be the target of the breeder. On the other hand, if the interest is to develop a variety for green forage having a relatively taller plant height and late flowering with high biomass accumulation, the marker allele with an increasing effect to plant height and days to 50 % flowering will be the target allele as plants with taller plant height and late flowering tend to accumulate high biomass for the purpose of green forage than with short stature and earliness in flowering (Habyarimana *et al.,* 2004). The same analogous is true for the application of the effects of the remaining linked marker alleles.

Generally, the identified seven SSR markers were localized on chromosomes 1, 2, 3,5,6,7, and 8 harboring one marker each (Xcup53, bSbCIR223, Xtxp114, mSbCIR248, Xtxp145, Xtxp278, and gbsp123 respectively) as shown below in Figure 5.



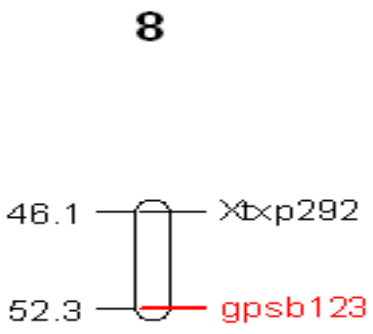


Figure 5. Chromosomal regions of marker trait associations. The linked markers highlighted in red are the present findings where as those in black are previously identified ones. Physical positions in base pairs (Mbp) are indicated on the left of the map and the corresponding marker names are indicated on the right. The sources for marker positions are Mace *et al.* (2009) and Wang *et al.* (2011).

In sorghum, several reports have been published using conventional QTL mapping and some of them are in agreement with the present study. Recently, Srinivas *et al*. (2009) using conventional QTL mapping reported that seed weight in sorghum was influenced by three QTLs residing on chromosomes 1, 4 and 6, with phenotypic variation explained by each QTL ranging from 7 to 14.8% with a major QTL detected on chromosome-one which explained 14.8% of phenotypic variation. Similarly, Shiringani *et al*., (2010) reported that nine significant QTLs associated with plant height were identified on chromosomes 1, 2, 4, 6, 7, 8. 9 and 10, harboring one QTL each except chromosome-four which contained two QTLs and other five QTLs were detected on chromosome 3, 4, 6, 7, and 8 associated with flowering dates.

Some of these previously identified QTLs using conventional QTL mapping in sorghum are similar to the findings of the present study. For example, in the study of Srinivas *et al*. (2009), one major QTL and two other QTLs were detected on chromosome 1 controlling seed weight which is similar to the finding of this study where locus Xcup53 on chromosome 1 is found to be significantly associated with thousand grain weight. In the same way, among the nine significant QTLs associated with plant height in the study of Shiringani *et al*. (2010), one was found to be located on chromosome 2 which corresponds to the present result where locus mSbCIR223 on chromosome 2 found to be significantly associated with plant height. Among the five QTLs detected and associated with flowering dates in the study of Shiringani *et al*. (2010), two of them were located on chromosome 6 and 7. Similar results were obtained from the present study where markers Xtxp145 on chromosome 6 and Xtxp278 on chromosome 7, both found to be significantly associated with days to 50% flowering. Thus, some of the QTLs detected by conventional QTL mapping in previous reports were also detected in the present study by association mapping approach indicating the reliability of association mapping in unraveling the complex marker-phenotype associations in molecular breeding.

However, there are also discrepancies between the findings of the present study and the previous conventional QTL mapping studies. For example, in Srinivas *et al*. (2009), QTLs controlling for panicle weight and grain yield were detected on chromosome 6 which are in discrepancies from the present finding where both panicle weight and grain weight per panicle were found to be significantly associated with locus gpsb123 on chromosome 8.

Moreover, in both Srinivas *et al*. (2009) and Shiringani *et al*. (2010), several QTLs simultaneously localized on more than three chromosomes were detected for each trait. For example, Srinivas *et al*. (2009) detected nine QTLs for days to an thesis, of which two on chromosome 1, two on chromosome 2, one each on Chromosome 3, 5, 6, 7, and 8. Similarly, in the study of Shiringani *et al*. (2010), QTLs for controlling flowering date were detected on chromosome 3, 4, 6, 7, and 8. Whereas in the present study only plant height, days to 50% flowering, and thousand grain weight were found to be associated with more than one chromosome simultaneously. For example, plant height was found to be associated with chromosomes 2, 5 and 6; days to 50% flowering with chromosomes 6, and 7, whereas thousand grain weight with chromosome 1 and 3. Similarly, Srinivas *et al*. (2009) identified three QTLs controlling seed weight on chromosome 1, 4, and 6 whereas only Xcup53 on chromosome 1 was found to be significantly associated with thousand grain weight in this study.

There are several possible causes for the discordance between this study and previous conventional QTL mapping studies. One reason might be that this study may not have detected all the existing major QTL because of the small number of markers used in the study. Another cause might be that a major QTL detected by a bi parental cross QTL mapping may not have large effect in the phenotypic variation of a germplasm collection and may be difficult to be detected with association mapping approach (Shahzad *et al*., 2009).

5. SUMMARY AND CONCLUSION

Sorghum [*Sorghum bicolor* (L.)Moench] serves as a food staple for more than 500 million of the world's most food insecure people, particularly in the semi-arid tropics of Asia and Africa including Ethiopia. Sorghum is known for its unusual tolerance to conditions of limited moisture and be productive during periods of extended drought circumstances which could not be achieved by most other cereals. However, this crop is seriously affected by drought at the reproductive stage especially during and post-flowering stage, implying that the true potential of sorghum can only be realized through genetic improvement (Ejeta and Knoll, 2007). In Ethiopia drought has remained to be the leading cause of disaster and human suffering in terms of frequency, area coverage and number of people affected, and sorghum in Ethiopia is predominantly cultivated in moisture stress areas that cover nearly 66% of the total area of the country (Tadesse *et al*., 2008) necessitating the development of drought tolerant varieties.

Although conventional breeding in the past has played a great role in developing abiotic stress tolerant varieties in different cereal crops, in quantitatively inherited traits such as drought tolerance selection using conventional methods has been neither efficient nor reliable due to the complex nature of the trait and the complicating effects of the environment (Ejeta and Knoll, 2007). Genome assisted selection (GAS) using molecular markers has been thought to be efficient for such traits (Haussmann *et al,* 2004). However, QTL mapping (conventional and/ or association mapping) is pre requisite for molecular markers to be used as selection criteria for crop improvements (Pennisi, 2008). In conventional QTL mapping, one must grow the plants for at least three years which is expensive in terms of time and man power. Also, the precision of QTL location is limited because of the relatively low amount of recombination that can occur in most, often relatively small, bi-parental offspring populations, and only two alleles at any particular locus can be assessed.

In order to overcome these limitations, association mapping has recently been employed by molecular breeders to map QTLs associated with desired traits, with no need of mapping population development but with a higher level of resolution (Hamblin *et al*, 2005). In association mapping studies, not only a collection of cultivars and breeding lines but also landraces can be used as mapping population (Sorkheh *et al*., 2008). Of particular

interest to sorghum breeders is, the possibility of using existing germplasm resources (land races) for gene and allele discovery on the basis of association mapping strategies. Ethiopia is a rich source of sorghum landraces which are valuable sources of desirable genes such as for drought tolerance (Amsalu *et al*., 2000). However, no studies to detect marker-trait association for drought tolerance in sorghum have been previously reported using association mapping in Ethiopia. Therefore, the objectives of this study were to identify drought tolerant sorghum genotypes, to map chromosomal regions (QTLs) associated with agronomically important and/or drought tolerance traits, and to identify SSR markers that were linked to agronomically important traits in sorghum, using association mapping.

In this study, one hundred sixty sorghum genotypes (152 land races and 8 released varieties) were used. These populations were evaluated in field at kobo in the off-season using irrigation in an alpha lattice design replicated three times. The phenotypic data were collected including days to 50% flowering, plant height, panicle weight, grain weight, grain weight per panicle, panicle harvest index, one thousand grain weight and number of grains per panicle.

Analysis of variance showed highly significant ($P<0.0001$) differences among the genotypes for all characters indicating wide variability for these characters among the genotypes. Grain weight per panicle ranged from 8.68 (210902) to 101.09 (210922) with an average value of 43.52. The first 17 best performers among the entries were the landraces indicating the presence of better performing genotypes in the landraces than the existing released varieties and the five best performing genotypes were 210922, 73067, 69183, 69231, and 69241 in that order.

Most of the characters showed moderate to high phenotypic and genotypic coefficient of variation. Heritability was high for all of the studied characters. High heritability coupled with high genetic advance as percent of mean is the most promising clue for possibility of improvement by selection and was observed for plant height, panicle weight, grain weight per panicle, one thousand grain weight, and number of grains per panicle.

Linkage disequilibrium (LD) analysis was performed using the 39 SSR markers. In all accessions, 107 locus pairs (32.92%) had a significant ($p < 0.05$) mean LD of 0.19, with an $R^2 > 0.2$ for 33 evaluated locus pairs. Association mapping requires population structure to be considered to avoid false positive association results. Population structure analysis showed that there were four distinct clusters in the studied materials. A total of 10 marker-trait associations related with agronomically important and/or drought tolerance characters were identified using 7 different SSR markers with R^2 ranging from 2.64 to 17.76. The seven SSR markers were localized on chromosomes 1, 2, 3,5,6,7, and 8 harboring one marker each (xcup53, bSbCIR223, Xtxp114, mSbCIR248, Xtxp145, Xtxp278, and gbsp123 respectively).

Most of the identified markers were linked with those previously identified as linked to agronomically important and/or drought tolerance-related traits using conventional QTL mapping approach supporting the reliability of the present findings. Although for the genome wide association study a large number of molecular markers are necessary, this study can serve as initial effort for the association mapping studies in sorghum particularly in our country as the associated SSR markers are potential candidates for marker assisted selection to improve drought tolerance in sorghum. However, as this study is the first attempt in the identification of QTLs for drought tolerance using association mapping, the identified QTLs need to be validated in independent or related populations and in different environments before their use in marker-assisted selection. Based on the result of this study it is recommended that:

- ✓ Based on further validation, the markers that showed association with traits should be used to select genotypes with desirable features for a trait.
- ✓ Land races which were found to be superior in their performance could be used for developing new varieties.

6. REFERENCES

Abdurakhmonov, I.Y., and A. Abdukarimov (2008). Application of AssociationMapping to Understanding the Genetic Diversity of Plant Germplasm Resources. *International Journal of Plant Genomics*, 1-18.

Addissu Gebre Ayele (2011). Heritability and genetic advance in recombinant inbred lines for drought tolerance and other related traits in sorghum (Sorghum bicolor). *Continental J. Agricultural Science* **5** (**1**): 1 – 9.

Agrarwissenschaften, D. (2008). Association mapping for drought stress related traits in a structured population with wild and cultivated barley. A PhD dissertation presented to school of graduate studies at the University of Bonn, Germany: 117p.

Alemu, T. (2009). Analysis of Genetic diversity among Ethiopian Sorghum [*Sorghum bicolor* (L.) Moench] germ plasm collections. An MSc. Thesis presented to the school of graduate studies of Hawasa University 98p.

Ali, M. A., S. Niaz, A. Abbas, W. Sabir, K. Jabran (2009). Genetic diversity and assessment of drought tolerant sorghum landraces based on morph-physiological traits at different growth stages. *Plant Omics J*., **2**: 214-227.

Allard, R.W. (1999). Principles of plant breeding 2nd (ed.). 95-115pp. John Wiley and Sons. Inc. New York.

Amsalu, A., Endeshaw B. and T, Bryngelsson (2000). Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). *Genetic Resources and Crop Evolution,* **47**: 471-482.

Amsalu, A. (2001). Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. A PhD dissertation presented to school of graduate studies of Addis Ababa University, Addis Ababa, Ethiopia. 175p.

Anithakumari, A. M. (2011). Genetic Dissection of Drought Tolerance in Potato (*Solanum tuberosum*.L). A PhD dissertation presented to school of graduate studies of Wageningen University, Wageningen, The Netherlands. 161p.

Anjum, S.A., X. Xie, L. Wang, M.F. Saleem., C. Man, and W. Lei (2011).Morphological, physiological and biochemical responses of plants to drought stress. *African Journal of Agricultural Research,* **6** (**9**), 2026-2032.

Aranzana, M.J., S. Kim, K. Zhao, E. Bakker, and S. Horton (2005). Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS Genet.,* **1** (**5**): 531-539.

Ardlie, K.G, L. Kruglyak, M. Seielstad (2002). Patterns of linkage disequilibrium in the human genome. *Nat. Rev. Genet*.; **4**:299-309.

Asfaw, A. (2007). Assessment of yield stability in sorghum. African Crop Science Journal, **15** (**2**): 83 – 92.

Atnafu Kebede (2010). Assessment of genetic diversity of Ethiopian sorghum (*Sorghum bicolor* (L.) Moench) germplasm using simple sequence repeat (SSR) markers. An M.Sc. Thesis presented to school of graduate studies of Jimma University college of Agriculture and Veterinary medicine (JUCAVM). 92p.

Avise, J. C. (2004). Molecular markers, natural history, and evolution. Journal of molecular Evolution **17**:214-226.

Badi, S., B. Pedersen, L. Monowar, and B.O. Eggum (1990). The nutritive value of new and traditional sorghum and millet foods for Sudan. *Plant Foods and Human Nutrition* **40**: 5-19.

Bantilan, M.C.S, U.K Deb, C.L.L. Gowda, B.V.S. Reddy, A.B. Obilana and R.E. Evenson (eds.) (2004). Sorghum genetic enhancement: research process, dissemination and impacts. Patancheru 502324, Andhra Pradesh, India: International Crops Research Institute for the Semi -Arid Tropics. 320 pp.

Beer, S. C., W. L. S. Siripoonwiwat, E. O'donoughue, Souza, D. Matthews, and M. E. Sorrels (1997). Associations between molecular markers and quantitative traits in an oat germplasm pool: can we infer linkages?" *Journal of Agricultural Genomics*, **3**: 185-197.

Beheshti, A.R., and B. B. Fard (2010). Dry matter accumulation and remobilization in grain sorghum genotypes (*Sorghum bicolor* L. Moench) under drought stress. *AJCS* **4**(**3**):185-189.

Bhattramakki, D., J. Dong, A. K. Chhabra, and G. E. Hart (2000). An Integrated Ssr and Rflp Linkage Map Of Sorghum (Sorghum Bicolor (L.) Moench). *Genome*, **43**: 988-1002

Bodmer, W.F. (1986). Human genetics: the molecular challenge. Cold Spring Harb. Symp. *Quant. Biol.* ***51***: 1–13.

Boyer, J. S. (1976). Photosynthesis at low water potentials. Philosophical Transactions of the Royal Society. **273**: 501-512.

Boyer, J.S. 1982. Plant production and environment. *Science*, **218**:443–448.

Bruce,W.B., G..O. Edmeades and T.C. Barker (2002). Molecular and physiological approaches to maize improvement for drought tolerance. *J. Exp. Bot.*, **53**(**366**): 13-25.

Bubeck DM, M.M. Goodman, W.D. Beavis, and D. Grant (1993). Quantitative trait loci controlling resistance to gray leaf spot in maize. *Crop Sci.* **33**:838–47.

Burton, G.W and E.H. DeVane. (1953). Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. *Agron. J.,* **45**: 478-81.

Can, N.D. and Yoshida, T.(1999). Genotypic and Phenotypic variances and covariance's in early maturing grain sorghum in a double-cropping. *Plant Prod Sci.* **2**(**1**):67-70.

Casa, A.M., G. Pressoir, P.J. Brown, S.E. Mitchell, W.L. Rooney, M.R. Tuinstra, C.D. Franks, and S. Kresovich. (2008). Community resources and strategies for association mapping in sorghum. *Crop Sci.* **48**:30–40.

Ciacci, C., L. Maiuri, N. Caporaso, C. Bucci, L. D. Giudice, D. R. Massardo, P. Pontieri , N. D. Fonzo, S. R**.** Bean, B. Ioerger, and M. Londei (2007). Celiac disease: in vitro and in vivo safety and palatability of wheat-free sorghum food products. Clinical Nutrition. **26**:799-805.

Chantereau J., G. Trouche, J.F. Rami, M. Deu, C. Barro, L. Grivet (2001): RFLP mapping of QTLs for photoperiod response in tropical sorghum. *Euphytica* , **120**:183-194.

Chaudhary, L. and S. Arora (2001). *Agricultural science. Digest*, **21**: 219-222.

Craufurd, P.Q., D.J. Flower, and J.M. Peacock. (1993). Effect of heat and drought stress on sorghum. I. Panicle development and leaf appearance. *Exp. Agric*. **29**:61–76.

Decker, W.L. (1983). Probability of drought for humid and sub-humid regions, p. 11-19. *In* C.D. Raper and P.J. Kramer (ed.) Crop reaction to water and temperature stresses in humid temperate climates. West View Press, Boulder, Colorado.

Deepalakshmi, A. J., and K. Ganesamurthy (2007). Studies on genetic variability and Character association in kharif sorghum {Sorghum bicolor (L.) Moench}. *Indian J. Agricultural Research*. **41** (**3**): 177-182.

Deshmukh, S.N.N., M.S. Basu and P.S. Reddy (1986) . Genetic variability, character association and path coefficient analysis of quantitative traits in *Virginia bunch* varieties of ground nut. *Indian J. Agric. Sci*., **56**: 515-518.

Dicko, M.H., H. Gruppen, A.S. Traore, G.J. Alphons, A.G.J. Voragen and W.J.H. Van Berkel (2006). Sorghum grain as human food in Africa: Relevance of content of starch and amylase activities. *African Journal of Biotechnology* **5**: 384-395.

Doggett, H. (1976). Sorghum bicolor (Gramineae-Andropogoneae). In, N.W. Simmonds (Ed.). Evolution of crop plants. Longman and New York.

Ejeta, G. and J. E., Knoll. (2007). Marker-assisted selection in sorghum pp. In: R. Vashney and R. Tuberosa (eds) Genomic Assisted Crop Improvement: Vol 2, Genomic Applications in Crops. pp187-205.

Evanno, G., S. Regnaut, and J. Goudet. (2005). Detecting the number of clusters of individuals using the software *Structure*: A simulation study. *Mol. Ecol*. **14**:2611–2620.

FAOSTAT Food and Agriculture Organization *(United Nations)* Statistics Division (2010) accessed 07 October 2010.

Farnir, F., Coppieters W., Arranz. J.J., P. Berzi., N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, M. Mni, C. Nezer, P. Simon, P. Vanmanshoven, and D. Wagenaar (2000). Extensive genome-wide linkage disequilibrium in cattle. *Genome Res*. **10:** 220–227.

Farooq, S. and F. Azam (2002). Molecular Markers in Plant Breeding-II. Some Pre-requisites for Use. *Pakistan Journal of Biological Sciences* **5** (**10**): 1141-1147.

Farooq, M., A. Wahid, N. Kobayashi, D. Fujita, S.M.A. Basra (2009). Plant drought stress: effects, mechanisms and management. *Agron. Sustain. Dev*., **29**: 185-212.

Fenster, C. (2003). White food sorghum in the American diet. In US Grains Council 43rd Board of Delegates' Meeting July 2003 Minneapolis MN http://www.grains.org/news/latest_news/43_delegates_meeting_upd ates.html

Flexas, J., J. Bota, F. Loreto, G. Cornic, T.D. Sharkey (2004). Diffusive and metabolic limitations to photosynthesis under drought and salinity in C_3 plants. *Plant Biol*., **6**: 1-11

Flint-Garcia, S.A., J.M. Thornsberry and E.S. IV. Buckler (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol*. **54**: 357–374.

Flint-Garcia, S.A., A.C.Thuillet, J. Yu, G. Pressoir, S.M. Romero, S.E. Mitchell, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. (2005). Maize association population: A high-resolution platform for quantitative trait locus dissection. *Plant J*., **44**:1054–1064.

Gaaris, A.J, S.R. McCouch, S. Kressovch (2003) Population structure and its effect on haplotype diversity and Linkage disequilibrium surrounding the Xa5 locus of Rice (*Oryza sativa L*). *Genetics***, 165** : 759-769.

Gebhardt, C. & J.P.T., Valkonen (2001).Organization of genes controlling disease resistance in the potato genome. *Annu Rev Phytopathol,* **39**: 79–102.

Georges, M. (2000). Extensive genome-wide linkage disequilibrium in cattle. *Genome Res.,* **10** (**2**)**:**220-227.

Geremew, G., Asfaw A., Taye T., Tesfaye T., Ketema B. and Hilemichael H. (2004). Development of Sroghum varieties and hybrieds for dry land areas of Ethiopia. *Uganda journal of agricultural science*, **9**:594-605.

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

Gupta, P. K., S. Rustgi, S. Sharma, R. Singh, N. Kumar and H. S. Balyan (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Molecular Genetics and Genomics*, **270:** 315-323.

Gupta, P.K., S. Rustgi and P. L. Kulwal (2005). Linkage disequilibrium and association studies in higher plants:Present status and future prospects. Plant Molecular Biology. **57**:461–485.

Habyarimana, E., D. Laureti, M. Deninno, and C. Lorenzoni (2004). Performance of biomass sorghum [*sorghum bicolor* (l.) Moench] under different water regimes in mediterranean region. Industrial crops and products **20:** 23-28.

Hamblin, M. T., S. E. Mitchell, G..M. White, J. Gallego, R. Kukatla, R. A.,Wing, A. H. Paterson, and S. Kresovich (2004). Comparative Population Genetics Of Panicoid Grasses: Sequence.

Hamblin, T. Martha, G. Maria, S. Fernandez, M. Alexandra, S. E. Casa, Mitchell, H. Andrew, Paterson, and Stephen Kresovich (2005). Equilibrium processes cannot explain high levels of short- and medium-range linkage disequilibrium in the domesticated grass *Sorghum bicolor. Genetics*, **171**: 1247–1256

Hao, Z. X., C. Li, J. Xie , M. Weng, D. Li , X. Zhang, L. Liang, S. Liu and S. Zhang. (2011). Identification of Functional Genetic Variations Underlying Drought Tolerance in Maize Using SNP Markers. *Journal of Integrative Plant Biology*, **53** (**8**): 641–652.

Harlan, J.R., and De Wet, J.M.J. (1972). A simplified classification of cultivated plants. Taxon **20**: 509-517.

Harlan, J.R., J.M.J. De Wet & A.B.L. Stemler (1976). Origin of African plant domestication. Mouton, The Hague. PP 409-452.

Harris, K., PK Subudhi, A. Borrell, D.R. Jordan, D.T. Rosenow, H.T. Nguyen, P.E. Klein, R.R. Klein, and J. Mullet (2007). Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence. *J Exp Botany*, **58**:327-338.

Haussmann, B.I.G., D.E. Hess, N. Seetharama, H.G. Welz And H.H. Geiger (2002). Construction Of A Combined Sorghum Linkage Map From Two Recombinant Inbred Populations Using Aflp, Ssr, Rflp, And Rapd Markers, And Comparison With Other Sorghum Maps. *Theoretical and Applied Genetics*, **105**: 629-637.

Homma K., T Horiea, T. Shiraiwaa , S. Sripodokb, and N. Supapoj (2004). Delay of heading date as an index of water stress in rainfed rice in mini-watersheds in Northeast Thailand field. *Crops Res.,* **88**: 11–19.

House, L. R**. (**1985). A guide to sorghum breeding. Second edition. Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics. 216 pp.

Hsiao, T. C. (1973). Plant responses to water stress. *Ann. Rev. Plant Physiol*. **24**: 519-570.

Hulse, J.H., E.M. Laing and O.E. Pearson (1980). Sorghum and the millets: Their composition and nutritive value, Academic Press, London, pp. 1-32.

Hussain, S. S. (2006). Molecular breeding for a biotic stress tolerance: drought Perspective *Pakistan Acad. Sc.,* **43** (**3**) 289-210.

International Plant Genetic Resources Institute (IPGRI) (1993). Descriptors for *Sorghum [Sorghum bicolor* (L.) Moench*]*. IBPGR/ ICRISAT. Rome, Italy.

Inuyama, S. (1978). Varietal differences in leaf water potential, leaf diffusive resistance and grain yield ofgrain sorghum affected by drought stress. *Jpn. J. Crop Sci*., **47**: 255-261.

Jannink, J. L. and Walsh, B. (2002). Association Mapping in Plant Populations. In: Quantitative Genetics, Genomics and Plant Breeding, Kang, M.S. (Ed.). CAB International, Oxford, UK., ISBN-13: 978-0851996011, pp: 59-68.

Jiménez, A., R. Holderegger, D. Csencsics and L. G. Quintanilla (2010). Microsatellites reveal substantial among-population genetic differentiation and strong inbreeding in the relict fern *Dryopteris aemula*. *Annals of Botany,* **106**: 149-155.

Johnson, H. W., H.F. Robinson and H. F. Comstock (1955). Estimates of genetic and environmental variability in soybean. *Agronomy Journal*, **47**: 314-318.

Johnson, H.F. Robinson and R.G. Cosmtock (1995). Genotypic and phenotypic correlation in Soy bean and their application in selection. *Agron.J.,* **47**:477-483.

Jordan, D. R., R. E. Casu, P. Besse, B.C. Carroll, N. Berding and C. L. Mcintyre (2004). Markers Associated With Stalk Number and Suckering In Sugarcane Colocate With Tillering And Rhizomatousness Qtls in Sorghum. *Genome*, **47**(**5**): 988-993.

Jun H., K. Van, Y.M. Kim, H. Lee and R.D. Walker (2008). Association analysis using SSR markers to find QTL for seed protein content in soybean. *Euphytica*, **162**:179–191.

Kamhara, A.Y, A. Menkir, B. Badu-apraku, O. Ibikunle (2003). The influence of drought stress on growth, yield and yield components of selected maize genotypes. *J. Agric. Sci*., **141**: 43-50.

Kamatar, M., Y., Kotragouda M., Shinde, D., G., and Salimath, P., M. (2011). Studies on Variability, Heritability and Genetic Advance in F3 Progenies of *Kharif* × *Rabi* And *Rabi* × *Rabi* Crosses Of Sorghum (*Sorghum Bicolour* L. Moench). *Plant Archives.* **11** (**2**): 899-901.

Kang, H.M., N.A. Zaitlen, C.M. Wade, A. Kirby, D. Heckerman, M.J. Daly, and E. Eskin, (2008). Efficient control of population structure in model organism association mapping. Genetics, **178**: 1709-1723.

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

Khan, M.B, Hussain, N., Iqbal, M. (2001). Effect of water stress on growth and yield components of maize variety. YHS 202. *J. Res*. (Science), **12**: 15-18.

Khan, M.Q., S.I. Awan and M.M. Mughal (2005). Estimation of genetic parameters in spring wheat genotypes under rain-fed conditions. *Indus J. Biol. Sci*., **2**: 367–70.

Klein, R.R, P.E. Klein, J. Mullet, P. Minx, W.L. Rooney, K.F. Schertz (2005). Fertility restorer locus *Rf* 1 of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. *Theor Appl Genet*, **111**:994-1012.

Kole, C. (2006). Genome Mapping and Molecular Breeding in Plants: Cereals and Millets. Volume 1. Springer-Verlag Berlin Heidelberg. 366 pp.

Kraakman, A.T.W., R.E. Niks, P.M. Van der Berg, P. Stam and F.A. Van Eeuwijk (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics,* **168**:435-46.

Kumar, A.A., B.V. Reddy, H.C. Sharma, C.T. Hash, P.S. Rao, B. Ramaiah and P.S. Reddy (2007). Recent Advances in Sorghum Genetic Enhancement Research at ICRISAT. *American Journal of Plant Sciences*, **2**: 589-600.

Li, M., N. Yuyama, L. Luo, M. Hirata and H. Cai (2009) *In silico* mapping of 1758 new SSR markers developed from public genomic sequences for sorghum. *Mol Breed*, **24**:41–47.

Li, Y., S. Bhosale, B.I. Haussmann, B. Stich, A.E. Melchinger and H.K. Parzies (2010). Genetic diversity and linkage disequilibrium of two homologous genes to maize *D8*: Sorghum *SbD8* and pearl millet *PgD8. Journal of Plant Breeding and Crop Science*. **2** (**5**):117-128.

Lipinsky, E.S., and S. Kresovich (1980). Sorghums as energy sources. p. 91-93. Proc. Bio-Energy ’80 World Congress and Exposition, Atlanta, GA. 21-24 Apr. 1980. The Bio-Energy Council, Washington, D.C.

Litt, M. And J.A. Luty (1989). A Hyper Variable Microsatellite Revealed By In Vitro Amplification of A inucleotide Repeat Within The Cardiac Muscle Actin Gene. *The American Journal of Human Genetics*, **44**: 397-401.

Liu L., L. Wang, Y. Zheng, and C. Zhao (2010). Association mapping of six agronomic traits on chromosome 4A of wheat (*Triticum aestivum* L.). *Molecular Plant Breeding*, **1** :(**5**) 1 -10.

Loiselle BA, V.L. Sork, J. Nason, and C. Graham (1995) Spatial genetic structure of a tropical understory shrub, Psychotria officinalis (Rubiaceae). *American Journal of Botany*, **82**:1420-1425.

Ludlow, M.M. and R.C. Muchow (1990). A critical evaluation of traits for improving crop yields in water-limited environments. *Adv. Agron*., **43**:107-152.

Mace E., J. Rami, S. Bouchet, P. Klein, R. Klein, A. Kilian, P. Wenzl, Xia L., K. Halloran and R. Jordan (2009). A consensus genetic map of sorghum that integrates multiple

component maps and high-throughput Diversity Array Technology (DArT) markers. *BMC Plant Biology,* **9**:13, 1-14.

Mace ES, H.K. Buhariwalla, and J.H. Crouch (2003) A high-throughput DNA extraction protocol for tropical molecular breeding programs. *Plant Mol Biol Rep.*, **21**: 459a-459h.

Mahajan R C, P.B. Wadikar, S.P. Pole and M.V. Dhuppe (2011). Variability, Correlation and Path Analysis Studies in Sorghum. *Research Journal of Agricultural Sciences.* **2**(**1**): 101-103.

Magill, C. W., K. Boora, R. Sunithakumari, O. Jairo, B.J. Oh, B. Gowda, Y. Cui, And R. Frederikson (1997). Tagging of Sorghum Genes For Disease Resistance. Expectations And Reality. In: Proceedings of the International Conference on Genetic Improvement of Sorghum and Pearl Millet, September 23-27, 1996, Lubock, Texas, Pp. 316-325.

Malosetti, M., C.G. van der Linden, B. Vosman and F.A. van Eeuwijk (2007). A mixed-model approach to association mapping using pedigree information with an illustration of resistance to Phytophthora infestans in potato. *Genetics,* **175**: 879–889.

Mann, J.A., C.T. Kimber, and F.R. Miller (1983). The origin and early cultivation of sorghums in Africa. Bull. No. 1454. Texas Agric. Exp. Stn., College Station, TX.

Mannai, Y. E., T. Shehzad and K. Okuno (2011). Variation in flowering time in sorghum core collection and mapping of QTLs controlling flowering time by association analysis. *Genet Resour Crop Evol*. **58**:983–989.

Manjarrez-sandoval, P., V.A. Gonzai-Ez-Hernandez, L.E. Mnnnoza-Onlornl,and E.M. Ano Enci-Euen (1989). Drought Stress Effects On The Grain Yield And Panicle Development Of Sorghum. Can. J. Plant Sci. **69**: 631-641.

Manonmani S, M. Suresh, A.K.F. Khan (2002). Genetic variability and correlation studies in red grain sorghum hybrids under rainfed condition. *Madras Agric. J*., **89** (**3**): 85-88.

Maposa, D., E. Mudimu, and O. Ngwenya (2010). A multivariate analysis of variance (MANOVA) of the performance of sorghum lines in different agro ecological regions of Zimbabwe. African Journal of Agricultural Research, **5**(**3**):196-203,

Matsuoka, M., R.T. Furbank, H. Fukayama, and M. Miyao (2001). Molecular engineering of C4 photosynthesis. Annual Review of Plant Physiology and Plant Molecular Biology, **52**: 297–314.

Mcintyre, C. L., R.E. Casu, J. Drenth, D. Knight, V.A. Whan, B.J. Croft, D.R. Jordan, And J.M. Manners (2005). Resistance Gene Analogues In Sugarcane And Sorghum And Their Association With Quantitative Trait Loci For Rust Resistance. Genome, **48**(**3**): 391-400.

Mekbib, F. (2009). Farmers' Breeding of Sorghum in the Centre of Diversity, Ethiopia: ISocio-Ecotype Differentiation, Varietal Mixture And Selection Efficiency. *Maydica*, **54**: 25-37.

Menz, M.A., R.R. Klein, J.E. Mullet, J.A. Obert, N.C. Unruh And P.E. Klein (2002). A High-Density Genetic Map Of Sorghum Bicolor (L.) Moench Based On 2926 Aflp, Rflp And Ssr Markers. *Plant Molecular Biology*, **48**: 483-499.

Mohan M, S. Nair, A. Bhagwat, T.G. Krishna, M. Yano, C.R. Bhatia, T. Sasaki (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding,* **3**:87-103.

Moose S.P and R.H. Mumm (2008) Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology,* **147**: 969-977.

Myles, J. P. J. Brown, E.S. Ersoz,a Zhiwu Zhang, E. Costich and E.S. Buckler (2009). Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. *The Plant Cell*, **21**: 2194–2202.

Nair, S. K., B.M. Prasanna, A. Garg, R.S. Rathore, S.A. Setty, And N.N. Singh (2005). Identification And Validation of QTLs Conferring Resistance To Sorghum Downy Mildew (Peranosclerospora Sorghi) And Rajasthan Downy Mildew (P. Heteropogony) In Maize. *Theoretical And Applied Genetics*, **110**(**8**):1384-1392.

Nagaraj, N. J., J.C. Reese, M.R. Tuinstra, C.M. Smith, P.S. Amand, M.B. Kirkham, K.D. Kofoid, L.R. Campbell And G.E. Wilde (2005). Molecular Mapping of Sorghum Genes Expressing Tolerance To Damage By Green Bug (Homoptera: Aphididae). *Journal of Economic Entomology*, **98**:595-602.

Nayyar, H., D. Gupta (2006). Differential sensitivity of C_3 and C_4 plants to water deficit stress: association with oxidative stress and antioxidants. *Environ. Exp. Bot.*, **58**: 106-113.

Nemera, G., Maryke T. Labuschagne and C.D. Viljoen. (2006). Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morphoagronomical markers. *Biodiversity and Conservation*, **15**: 3251-3265.

Nguyen, T.T., N. Klueva, V. Chamareck, A. Aarti, G. Magpantay, A. C. Millena, M.S. Pathan and H.T. Nguyen (2004). Saturation mapping of QTL regions and identification of putative candidate genes for drought tolerance in rice. *Mol. Genet. Genomics,* **272**:35-46.

Nordborg (2005). Genome-wide association mapping in *Arabidopsis* identifi es previously known fl owering time and pathogen resistance genes. *PLoS Genet*, **1**(**5**):531-539.

Panse, V. G. (1957) Genetics of quantitative characters in relation to plant breeding. *Indian Journal of Genetics and Plant Breeding*, **17**: 318-327.

Parh DK, D.R. Jordan, E.A.B. Aitken, B.J. Gogel, C.L. McIntyre, I.D. Godwin (2006). Genetic Components of Variance and the Role of Pollen Traits in Sorghum Ergot Resistance. *Crop Sci*, **46**:2387-2395.

Parmesan, C. and G. Yohe (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**: 37-42.

Paterson A, K. Schertz, Y. Lin, Z. Li (1998). Case history in plant domestication: sorghum, an example of cereal evolution. In Molecular dissection of complex traits, ed. Paterson A., pp. 187-195 Boca Raton, Florida: CRC Press.

Paterson, A.H. (2008). Genomics of sorghum (A review). *International Journal of Plant Genomics*. 1-6. *An open access journal.*.

Paterson, AH, JE Bowers, R. Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G Haberer, U Hellsten, T Mitros, A Poliakov, J Schmutz, M Spannagl, H Tang1, X Wang, T Wicker, AK Bharti, J. Chapman, F.A. Feltus, U Gowik, IV Grigoriev, E Lyons, CA Maher, M Martis, A. Narechania, RP Otillar, BW Penning, AA Salamov, Y Wang, L Zhang, NC Carpita, M Freeling, AR Gingle, CT Hash, B Keller, P Klein, S Kresovich, MC McCann, R M, D.G. Peterson, M. Rahman, D Ware, P Westhoff, KFX Mayer, J Messing and D.S. Rokhsar. (2009). The Sorghum bicolor genome and the diversification of grasses. *Nature.* **457**: 551-556.

Pennisi E. (2008). Plant genetics: the blue revolution, drop by drop, gene by gene. *Science,* **320** : 171–173.

Pereira, M. G. and M. Lee (1995). Identification of genomic regions affecting plant height in sorghum and maize. *Theoretical and Applied Genetics*, **90:**380-388.

Perumal, R., R. Krishnaramanujam, M.A. Menz, S. Katile, J. Dahlberg, Majil, and W.L. Rooney (2007). Genetic diversity among sorghum races and working groups based on AFLPs and SSRs. *CropSci.,* **47**:1375-1383.

Prasad, P.V.V, and S. A. Staggenborg (2008). Impacts of drought and/or heat stress on physiological, developmental, growth, and yield processes of crop plants. *Plant Science and Entomology*, 1-56.

Price, A.L., N.J. Patterson, R.M. Plenge,M.E. Weinblatt, N.A. Shadick and D. Reich (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet*. **38**: 904–90.

Prichard, J.K., M. Stephens, N.A. Rosenberg and P. Donnelly (2000a) Inference of population Structure using multilocus genotype data. *Genetics,* **155**: 945-959.

Prichard, J.K., M stephens, N.A. Rosenberg and P. Donnelly (2000b). Association mapping in structured populations. *Am. J. Hum. Genet.,* **67**: 170-181.

Rafalski, J. & S. Tingey (1993). Genetic diagnostics in Plant Breed: RAPDs, microsatellites and machines. *Trends Genet,* **9**: 275–280.

Rao, S.P, S.S. Rao, N. Seetharama, A.V. Umakanth, P.S. Reddy, B.V.S. Reddy and C.L.L. Gowda (2009). Sweet sorghum for biofuel and strategies for its improvement. Information Bulletin No. 77. Patancheru 502 324, Andhra Pradesh, India: ICRISAT. 80 pp. ISBN: 978-92-9066-518-2.

Rashidi, M. and K. Seyfi (2007). Effect of water stress on crop yield and yield components of cantaloupe. *Int. J. Agri. Biol.,* **9:** 271-273.

Reddy, P. R. R., N. D. Das, G. R. M. Sankar and A. Girija (1996). Genetic parameters in winter sorghum (*Sorghum bicolor*) genotypes associated with yield and maturity under moisture stress and normal conditions. *Indian Journal of Agricultural Sciences*, **66(11):** 661-664.

Reddy, B. V. S., B. Ramaiah, A. Kumar, and P.S.Reddy (2007). Evaluation of sorghum genotypes for the stay-green trait and grain yield. An open access journal published by ICRISAT, 1-4.

Reynolds, M., F. Dreccer and R. Trethowan (2007). Drought-adaptive traits derived from wheat wild relatives and landraces. *Journal of Experimental Botany*, **58** (**2**): 177–186.

Ribaut, J.M. & D. Hoisington, 1998. Marker-assisted selection: New tools and strategies. *Trends Plant Sci*., **3**: 236–239.

RibautJ-M.&Poland D.(2004). http://www.cimmyt.org/english/docs/proceedings/molecApproaches/pdfs/MolecularApproach.

Robinson, H.F., R.E. Cosmtock and P.H. Harvey (1949). Estimates of heritability and the degree of dominance in Corn. *Agron. J*., **41** : 353-359.

Samarah NH, A.M. Alqudah, J.A. Amayreh, G.M. McAndrews (2009). The effect of late-terminal drought stress on yield components of fourbarley cultivars. *J. Agron. Crop Sci*., **195**: 427-441.

Samson, F. and F. Knopf (1994). Prairie conservation in North America. *Bioscience*, **44**:418-421.

Sanchez, A.C., P.K. Subudhi, D.T. Rosenow And H.T. Nguyen (2002). Mapping Qtls Associated With Drought Resistance In Sorghum *(Sorghum Bicolor* (L) Moench). *Plant Molecular Biology*, **48**:713-726.

SAS Institute. Inc. (2008). Statistical analysis soft ware version 9.2, Cary NC: SAS institute inc. USA.

Sax, K. (1923). The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris. *Genetics*, **8**:552–60.

Selkoe, K.A, R.J. Toonen (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**: 615–629.

Senay, G.B. and J. Verdin (2002). Evaluating the performance of a crop water balance model in estimating regional crop production pecora 15/land satellite information IV/ISPRS commission I/FIEOS 2002 confrerence proceedings.

Sharma, K.K. and M. Lavanya (2002). Recent developments in transgenics for abiotic stress in legumes of the semi arid tropics In: Ivanaga M (ed) Genetic engineering for crop plants for abiotic stress, JIRCAS Working Report No. 23: 61-73, Tsukuba, Japan.

Shehzad, T., H. Iwata, and K. Okuno (2009). Genome wide association mapping of quantitative traits in sorghum (*Sorghum bicolor* (L.) Moench) by using multiple models. Breeding Science. **59**:217-227.

SHI A., P. CHEN, B. ZHANG and A. HOU (2010). Genetic diversity and association analysis of protein and oil content in food-grade soybeans from Asia and the United States. *Plant breeding*, 1-7.

Shinde, D. G., B.D. Biradar, P.M. Salimath, M.Y. Kamatar, A.R. Hundekar and S.K. Deshpande (2010). Studies on genetic variability among the derived lines of B x B, B x R and Rx R crosses for yield attributing traits in rabi sorghum (*Sorghum bicolor*(L.) Moench). *Electronic Journal of Plant Breeding*, **1**(**4**): 695-705

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

Simko, I, K.G. Haynes, E.E. Ewing, S. Costanzo, B.J. Christ and R.W. Jones (2004) Mapping genes for resistance to *Verticillium albo-atrum* in tetraploid and diploid potato populations using haplotype association tests and genetic analysis. *Mol Genet Gen.,* **271**:522-531.

Duncan C., Thomas, W. Robert, Haile and D. Duggan (2005) Recent developments in genomewide association scans: a workshop summary and review. *Am. J. Hum. Genet*., **77**: 337–345. *An open access journal*

Sinaki, J. M., E.M. Haravan, A.H.S. Rad, G. Noormohammadi and G. Zarei (2007). The effects of water deficit during growth stages of canola (*Brassica napus* L.). *American-Eurasian J. Agric. and Environ. Sci.* **2:** 417-422.

Singh, Faujdar, K.N. Rai, Reddy, V.S. Belum and D. Diwakar (eds.) (1997). Development of cultivars and seed production techniques in sorghum and pearl millet. Training manual. Training and Fellowships Program and Genetic Enhancement Division, ICRISAT, India. 118 pp.

Smith, C.W. and R.A. Frederiksen (2000). Sorghum: Origin, history, technology and production, John Wiley and Sons Inc., New York. NY 824, p. 668.

Snowdon, R. & W. Friedt (2004). Molecular markers in *Brassica* oilseeds breeding: Current status and future possibilities. *Plant Breed.,* **123**: 1–8.

Sorkheh, K., L.V. Malysheva-Otto, M.G. Wirthensohn, S. Tarkesh-Esfahani, and P. Martínez-Gómez (2008). Linkage disequilibrium, genetic association mapping and gene localization in crop plants. *GenetMolBio.,* **l31**:805-814.

Srinivas, G., Satish, K., Madhusudhana K., Seetharama, N. (2009). Exploration and mapping of microsatellite markers from subtracted drought stress ESTs in Sorghum bicolor (L.) Moench. *Theor Appl Genet.,* **118**:703–717.

Moose, S. P. and R.H. Mumm (2008). Molecular Plant Breeding as the Foundation for 21st Century Crop Improvement. *Plant Physiology*, **147**: 969–977.

Stewarti, M., J., and K. S. Kantartzi (2008). Association Mapping of Important Phenotypic Traits in *Gossypium arboreum* Accessions. Summaries of Arkansas Cotton Research. 24-28.

Sundaresha (2006). Identification of genomic regions harboring QTLs for stay-green and related traits insorghum [*Sorghum bicolor* (L.) Moench].*M Sc* (*Agri.*) An M.Sc.Thesis presented to the school of post graduate studies at University of Agricultural sciences. Dharwad. PP72.

Szalma, S.J, I.V.E.S. Buckler, M.E. Snook, M.D. MucMullen (2005). Association analysis of candidate genes for mysin and chlorogenic acid accumulation in maize silks. *Theor Appl Gennet* **110** : 1324-1333.

Tadesse, T., Tesso, T. and Ejeta, G. (2008). Combining ability of introduced sorghum parental lines for major morpho-agronomic traits. *Journal of SAT Agricultural Research,* **6**:1-7.

Tanksley, S. D. (1993). Mapping polygenes. *Annual Review of Genetics*, **27:** 205-233.

Tao, YZ, A.Hardy, J. Drenth, R.G. Henzell, B.A. Franzmann, D.R. Jordan, D.G. Butler, C.L. McIntyre (2003). Identifications of two different mechanisms for sorghum midge resistance through QTL mapping. *Theor Appl Genet*, **107**:116-122.

Tariq, M., S. Iqbalawan and M. Irshad ul haq (2007). Genetic Variability and Character Association for Harvest Index in Sorghum under Rainfed Conditions. *International journal of agriculture & biology*, 470–472.

Tesfaye, T., Alemu T. and Hussein M. (2011). Association between morphological traits and yield components in the Dura sorghums of Ethiopia. *Hereditas*, **148** (**3**): 98-109.

Teshome, A., D. Patterson, Asfaw., J.K. Torrance & J.T. Arnason (2007). Changes of *sorghum bicolor* landrace diversity and farmers' selection criteria over space and time, Ethiopia. *Genet Resour Crop Evol*., **54**:1219-1233.

Thoday J.M. (1961). Location of polygenes. *Nature,* **191**:368–70.

Thornsberry, J.M., M.M. Goodman, J. Doebley, S. Kresovich, D. Nielsen and E.S.I.V. Buckler (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet,* **28**:286-289.

Tuinstra, M.R., E.M. Grote, P.B. Goldsbrough and Ejeta**,** G. (1996). Identification of quantitative trait loci associated with preflowering drought tolerance in sorghum. *Crop Sci*. **36**:1337–1344.

Tuinstra, M. R., E.M. Grote, P. B. Goldsbrough and Ejeta**,** G. (1997). Genetic Analysis Of Post-Flowering Drought Tolerance And Components Of Grain Development In Sorghum Bicolor (L.) Moench. Molecular Breeding, **3**:439-448.

Varshney, R.K and R. Tuberosa (2007). Genomic-assisted crop improvement. *Genomics applications in Crops*, **2**:187-205.

Virk, P. S., B.V. Ford-Lloyd,M.T. Jackson, H.S. Pooni, T.P. Clemeno and H.J. Newbury (1996). Predicting quantitativevariation within rice germplasm using molecular markers," *Heredity*, vol. **76** (**3**) 296–304.

Voorrips, R.E. (2002). Map Chart: Software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity*, **93** (**1**): 77-78.

Wang, Y. and B. Rannala (2005). In silico analysis of disease-association mapping strategies using the coalescent process and incorporating ascertainment and selection. *Am J Hum Genet,* **76**:1066-1073.

Wang, J., P.L.R. McClean, R.J. Goos, and T. Helms (2008). Association mapping of iron deficiency Chlorosis loci in soyabean (Glycine max L. merr.). Advanced breeding lines. *Theor Appl. Genet.,* **116:** 777-787.

Wang,Y., R. Loganantharaj, P. Bible, H.D., Upadhyaya (2011). Identification of SSR markers associated with height using pool-based genome-wide association mapping in sorghum. 281-292. *An open access journal*.

Weir, B.S. (1996). Genetic Data Analysis II. Sinauer, Sunderland, MA, USA.

Wilhelm, W., W., J., M., F., Johnson, J.L. Hatfield, W.B. Voorhees, and D.,R., Linden. (2004). Crop and soil productivity response to corn residue removal: A literature review. *Agron. J*. **96**:1-17.

Williams, K.J. (2003). The molecular genetics of disease resistance in barley. *Aust J Agric Res* **54**: 1065–1079.

William, W.T., C.A.P. Boundy and A.J. Millington (1987). The effect of sowing date on the growth and yield of three sorg-hum cultivars. The component of growth and yield. Aust. J. *Agric. Res*., **28**: 381-387

Wilson, J.R, M.M. Ludlow, M.J. Frsher, and E.D. Schulze (1980). Adaptation to water stress of the leaf relations of four tropical forage speeies. *Aust J Plant Physio*., **l 7**: 207-220.

Woldeamlak , B. (2009). Rainfall variability and crop production in Ethiopia; Case study in the Amhara region; proceeding of the 16th International conference of Ethiopian Studies, Addis Ababa, Ethiopia.

Xu, W., P.K. Subudhi, O.R. Crasta, D.T. Rosenow, J.E. Mullet and H.T. Nguyen (2000). Molecular Mapping of QTLs Conferring Stay-Green In Grain Sorghum (Sorghum Bicolor L. Moench). *Genome*, **43**:461-469.

Xu, J.L, H.R. Lafitte, Y.M. Gao, B.Y. Fu, R. Torres, and Z.K. Li (2005). QTLs for drought escape and tolerance identified in a set of random introgression lines of rice. *Theor Appl Genet*., **111**: 1642–1650.

Yonemaru, J., T. Ando, T. Mizybayashi, S. Kasuga, T. Matsumoto and M. Yano (2009). Development of genome-wide simple sequence repeat markers using whole-genome shotgun sequences of sorghum (Sorghum bicolor (L.) Moench). *DNA Research* **16**:187–193.

Yong'an, L., D. Quanwen, C. Zhiguo and Z. Deyong (2010). Effect of drought on water use efficiency, agronomic traits and yield of spring wheat landraces and modern varieties in Northwest China. *African Journal of Agricultural Research,* **5** (**13**): 1598-1608.

Young, N.D., D. Danesh, D. Menancio-Hautea, and L. Kumar (1993). Mapping oligogenic resistance to powdery mildew in mungbean with RFLPs. Theor. *Appl. Genet*. **87**:243–49.

Yu, X., G. Bai, N. Luo, Z. Cheng, S. Liu, J. Liu, S.E. Warnke, and Y. Jiang (2011). Association of simple sequence repeat (SSR) markers with submergence tolerance in diverse population of perennial ryegrass. *Plant Science*, **180**:391-398.

Yu, J., G. Pressoir, W.H. Briggs, I. Vroh Bi, M. Yamasaki,J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet*., **38**: 203–208.

Zhao, H.D, D. Nettleton, M. Soller, and J.C.M. Dekkers (2005). Evaluation of linkage disequilibrium measures between multi-alleleic markers as predictors of linkage disequilibrium between marker and QTL. *Genet. Res*., **86**: 77-87.

Zhao, K., M.J. Aranzana, S. Kim, C. Lister, C. Shindo, C. Tang, C. Toomajian, H. Zheng, C. Dean, P. Marjoram and M. Nordborg (2007). An Arabidopsis example of association mapping in structured samples. *PLoS Genet*., **3**: 71-82.

Zhao, W., E. Park, J. Chung, Y. Park, I. Chung, J. Ahn, and J. Kim (2009). Association Analysis of the Amino Acid Contents in Rice. *Journal of Integrative Plant Biology*, **51** (**12**): 1126–1137.

Zhou, J., A. You, Z. Ma, L. Zhu, and G. He (2012). Association analysis of important agronomic traits in japonica rice germ plasm. *African Journal of Biotechnology,* **11**(**12**): 2957-2970.

7. APPENDIX

Appendix Table 1 . Analysis of variance for seven agronomical traits in 160 sorghum genotypes evaluated for drought tolerance at Kobo

Source of variation	Mean squares						
	DF	PHT	PWT	GWPP	PHI	TGWT	NGPP
Replication	9.21	0.351	197.323	86.434	0.0008	36.97	127660.4
P-value	ns	0.0019	ns	ns	ns	ns	ns
Block(rep)	16.464	0.114	121.711	88.56	0.0031	10.75	89564.1
P-value	ns	0.0002	ns	ns	0.0083	ns	ns
Genotypes	297.35	0.389	739.05	535.91	0.01	108.48	654782
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Error	20.71866	0.054506	89.22	52.52	0.00	9.88	71083.33
CV(%)	4.726	11.40806	18.828	20.772	5.474	10.743	19.852
SE (±)	0.52	0.02	0.82	0.69	0.003	0.31	24.26
R^2	0.909	0.835372	0.8223	0.812403	0.7296	0.8694	0.8358
Efficiency overRCBD	115	108.67	105	117	95.6	123.12	132.45

The degree of freedoms for replication, block, and genotypes were 2, 45, and 159 respectively.
ns = none significant probability value
DF = Days to 50% flowering, PHT = Plant height, PWT = panicle weight, GWPP = grain weight per panicle, PHI = panicle harvest index, TGWT = thousands grain weight, and NGPP = number of grains per panicle

Appendix Table 2. Mean monthly temperature and rain fall data at the experimental site during the cropping period (January-June 2010/2011).

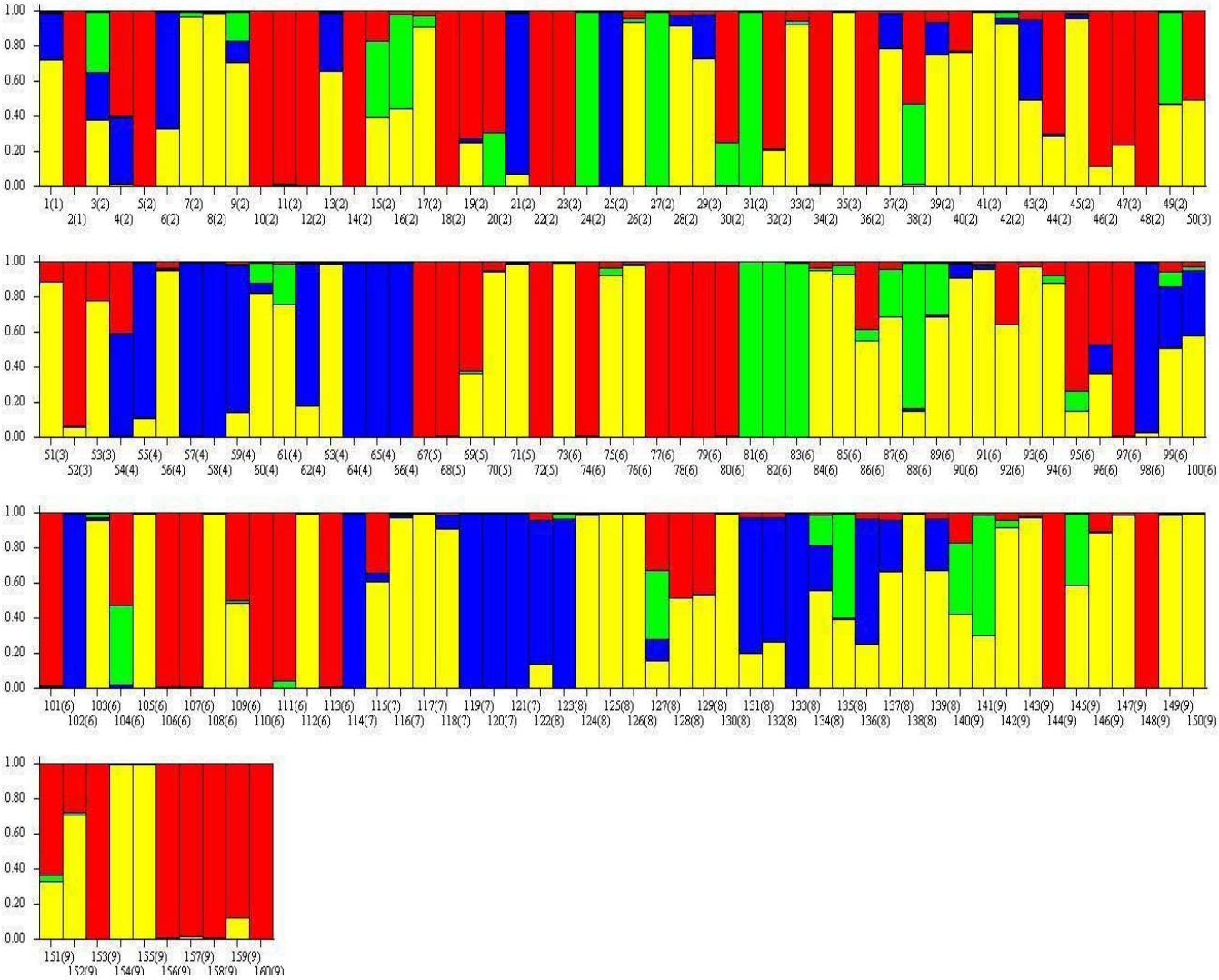
Month	Temperature (oc)		Rain fall (mm)	
	Minimum	Maximum	Minimum	Maximum
December	5.5	20.55	-	-
January	6.75	21.35	-	-
February	6.0	22.75	0.85	2.3
March	8.75	23.85	1.15	3.15
April	9.5	24.5	1.5	31.05
May	10.5	25.15	2.5	5.5
June	12.5	27.05	1.15	2.5
Mean	**8.5**	**23.6**	**1.43**	**8.9**

Appendix Table 3. Regions from which the accessions were collected

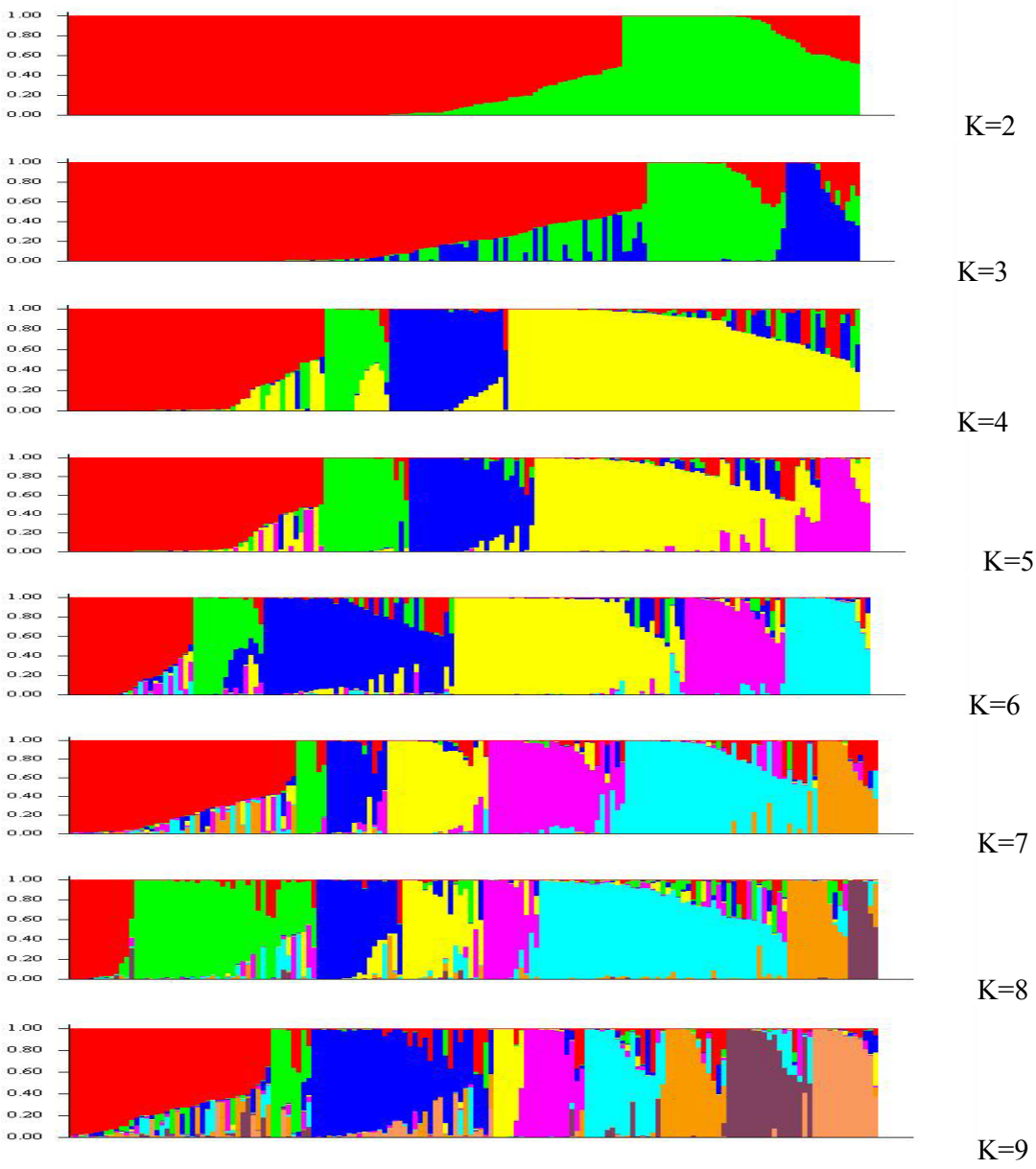
Region	Afar (1)	Amhara (2)	Eritrea (3)	Gambella (4)	NA (5)	Oromia(6)	RV (7)	SNNS (8)	Tigray (9)	Total
N0.of acc.	2	47	4	13	6	41	8	18	21	160

Numbers in brackets are code numbers used to represent the names of regions. RV= Released varieties, NA= Information not available, and N0. of acc. = number of accessions.

8.2. List of figures in the appendix



Appendix fig.1. The Proportional membership of the studied sorghum accessions. Each individual sample is represented by a single vertical line broken into one to four-colored segments (Blue, Green, Yellow and Red) with lengths proportional to each of the four inferred population subgroups. Each individual represents a sorghum variety; the numbers outside the brackets are entry numbers for individuals examined, whereas the numbers in brackets are the code numbers for geographical sampling locations from which the samples were taken. Code numbers are defined in appendix Table 3.



Appendix fig.2. Structure out puts for number of sub populations from 2 up to 9.