ADDIS ABABA UNIVERSITY

INSTITUTE OF BIOTECHNOLOGY

Genetic Diversity Analysis of Sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia based on Quantitative Traits and Microsatellite Markers

MSc. Thesis

Wubshet Mamo Umie

July 2021

Addis Ababa, Ethiopia

Genetic Diversity Analysis of Sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia based on Quantitative Traits and Microsatellite Markers

A Thesis Submitted to the Institute of Biotechnology, Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biotechnology

By

Wubshet Mamo Umie

July 2021

Addis Ababa, Ethiopia

Thesis Declaration

We certify that Wubshet Mamo's MSc Thesis Research entitled "**Genetic Diversity Analysis of Sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia based on Quantitative Traits and Microsatellite Markers**" has been conducted under our direct supervision. We recommended that the thesis has been accepted as fulfilling the requirement for the Degree of Master of Science in Biotechnology.

Supervisors:	**Signature**	**Date**
Tileye Feyissa (PhD, Associ. Prof.), IoB, AAU		Aug. 27, 2021
Tilahun Mekonnen (PhD), IoB, AAU		Aug. 27, 2021
Muluken Birara (MSc., PhD candidate), IoB, AAU		Aug. 27, 2021

We, the undersigned, members of the Board of Examiners of the final open defense by Wubshet Mamo have read and evaluated his thesis entitled "**Genetic Diversity Analysis of Sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia based on Quantitative Traits and Microsatellite Markers**" and examined the candidate. This is therefore to certify that the thesis has been accepted in partial fulfilment of the requirement for the Degree of Master of Science in Biotechnology.

Examining Committee:	**Signature**	**Date**
Teklehaimanot Hailesilassie (PhD, Associ. Prof.), IoB, AAU Name of Examiner		Aug. 27, 2021
Demsachew Guadie (PhD), IoB, AAU Name of Examiner		Aug. 27, 2021
Tesfaye Sisay (PhD, Professor) Director, Institute of Biotechnology		Aug. 27, 2021

አዲስ አበባ ዩኒቨርሲቲ
የባዮቴክኖሎጂ ኢንስቲትዩት
Institute of Biotechnology
Addis Ababa University

Dedication

This thesis is dedicated to my Mother, **Konjit Betemariyam**, for her remarkable and endless love, encouragement, sacrifices, prayers, supports and advices in all my life.

Acknowledgements

Above all, I thank the almighty GOD for keeping me strong from the start to the end of this work. I thank Dr. Tileye Feyissa, my major supervisor, for giving me the opportunity to work with him and his team. His dedication and overwhelming attitude in all circumstances made the success of this study. His fatherly supervision, timely advice, meticulous scrutiny, scholarly advice, patience, motivation, guiding and commenting on thesis writing and scientific approach during my work at the field and laboratory have helped me at large to accomplish the study.

I am very glad to express my deepest sense of gratitude to my co-supervisors, Dr. Tilahun Mokennen and Mr. Muluken Birara; they have been always there to listen and give me advice. I am deeply grateful to them, for the long discussions that greatly helped me for successful completion of the thesis work.

My deepest sense of gratitude once again goes to Dr. Tilahun Mekonnen for his close supervisions during the laboratory work, utilization of softwares, data analysis and interpretation of the results. His prompt inspiration, timely suggestion with kindness, enthusiasm and dynamisms have given me strength to complete my thesis.

I am very grateful for Mr. Muluken Birara for his inestimable helps in facilitating all the ways in collecting seeds for the study, preparing working materials, securing the land for field experiment, and all rounded assistances in data collection and analysis.

I would like to extend my special thanks to Dr. Kassahun Tesfaye, Director General of Ethiopian Biotechnology Institute (EBTi) for allowing me to work in his Plant Genetics Research Laboratory at College of Natural and Computational Sciences, Addis Ababa University.

My special thanks also goes to Professor Tesfaye Sisay and Dr. Addis Simachew, Director and Deputy Director to the Institute of Biotechnology, respectively for their help to use their laboratory facilities.

This work would not be possible without the financial support of the Swedish International Development Agency (SIDA) under the project (sub project 1.2 entitled “Sorghum: Screening and association mapping of drought tolerant sorghum genotypes in Ethiopia”). Hence, I am very thankful for the SIDA program.

I would like to thank the Institute of Biotechnology, Addis Ababa University, for all academic supports I got during the study period. The National Agricultural Biotechnology Research Center (NABRC) is also thanked for all the support I received during my study.

I would like to thank Mr. Girma Gashu for all his support and encouragement during the field data collection.

Special thanks also go to Mr. Kiros Tekle, Mr. Temesgen Dagnaw and Mrs. Genet Atsbeha for their continued support and encouragement during my stay in the laboratory. I appreciate and value all your kind support, help and hospitality.

I am thankful to my-home Institute; the Ethiopian Biodiversity Institute (EBI) for financial support and giving me a study leave. I am thankful for the core directorate members, administration and all the supportive staff members of the Institute.

Finally, I would like to present limitless thanks to my parents for all their encouragements and supports throughout my career. Thank you all the family members!

Table of Contents

Contents Page

List of Figures

Contents **Page**

List of Tables

Contents **Page**

List of Acronyms and Abbreviations

AMOVA	Analysis of molecular variance
ANOVA	Analysis of variance
CV	Coefficient of variation
CSA	Central Statistics Agency
DAP	Di-ammonium phosphate
DArT	Diversity Arrays Technology
EBI	Ethiopian Biodiversity Institute
EIAR	Ethiopian Institute of Agricultural Research
FAO	Food and Agriculture Organization of the United Nation
He	Expected heterozygosity
FAOSTAT	Food and Agriculture Organization Corporate Statistical Database
Ho	Observed heterozygosity
IBPGR	International Board for Plant Genetic Resources
ICISAT	International Crops Research Institute for the Semi-Arid Tropics
MARC	Melkassa Agricultural Research Center
MCMC	Monte Carlo Markov Chain
METAR	Multi Environment Trial Analysis for R
MHARC	Mehoni Agricultural Research Center
Na	Number of alleles
NJ	Neighbor Joining
PCA	Principal component analysis

PCoA	Principal coordinate analysis
PCs	Principal components
PIC	Polymorphic information content
SSR	Simple sequence Repeat
UPGMA	Unweighted Pair-group Method using Arithmetic Averages
USDA	United States Department of Agriculture

Genetic Diversity Analysis of Sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia based on Quantitative Traits and Microsatellite Markers

Wubshet Mamo

Institute of Biotechnology, Addis Ababa University, P.O.Box: 1176, Addis Ababa, Ethiopia.

Email: wubma53@gmail.com

ABSTRACT

Sorghum [*Sorghum bicolor* (L.) Moench] is a very important food security crop in Ethiopia and in the semi-arid and tropical parts of Africa. Genetic analysis based on DNA markers coupled with morphological traits could provide important information on the genetic structure of the crop. Therefore, the present study was targeted to investigate the extent and patterns of genetic variation in 324 sorghum accessions using six quantitative traits measured at three locations in two replications. Among the studied traits, plant height showed the highest genetic advance as percent of mean (38.99%) followed by biomass yield (18.68%). Moreover, the genetic structure of 100 sorghum accessions were further assessed using 15 microsatellite markers. The results revealed the presence of significant differences among the genotypes for all the studied quantitative traits, suggesting the possibility to improve through selection. The polymorphic information content of the loci ranged from 0.68 to 0.89 with overall mean of 0.80, indicating their higher resolution power for genetic analysis. Structure analysis confirmed the three sub-groups with greater degree of genetic admixture. We would like to confirm that the combined application of morphological and marker technology had successfully elucidated the genetic structure of sorghum accessions in Ethiopia. The information will be useful for sorghum improvement through selection, conservation, and wise and sustainable use.

Keywords/Phrases: Gene flow, Genetic structure, Polymorphic information content, Quantitative traits, Sorghum.

1. INTRODUCTION

Sorghum bicolor (L.) Moench, commonly known as sorghum (2n = 2x = 20), is the most important grain crop in the world's semi-arid tropics because of its adaptation to a wide range of climate and ecological conditions (Ali *et al.*, 2011). It is predominantly a self-fertilizing species, with a 7-30% outcrossing rate (Ellstrand and Foster, 1983; Djè *et al.*, 2004). Based on inflorescence and grain morphology, sorghum has been traditionally classified into five basic races: Bicolor, Caudatum, Durra, Guinea and Kafir, and ten intermediate races in all pairwise combinations of the five basic races (Harlan and De Wet, 1972).

Globally, sorghum ranked fifth in production after maize, wheat, rice and barley. With respect to cultivation area, sorghum ranks second after maize, accounting for about 22% of total area under cereal production which in turn followed by cereal grass (USDA, 2019). It is widely cultivated for its multiple purposes including food, feed, fodder, and fuel. It also has various industrial applications like health, pharmaceutical diagnosis, synthesis of organic molecules, and utility items (Hariprasanna and Rakshit, 2016).

In Ethiopia, sorghum ranks third next to maize and teff in area coverage, and fourth in production after maize, teff and wheat (CSA, 2020). It is cultivated in most parts of the Ethiopia, and in 2019 it was cultivated on 1.8 million hectare with total production of 5.27 million tonne (FAOSTAT, 2019).

The present average sorghum productivity in Ethiopia is 2.88 tonne ha^{-1}, which is substantially higher than the global average productivity (1.44 tonne ha^{-1}) (FAOSTAT, 2019). Irrespective of

this comparative higher productivity, there is huge gap between sorghum production and demand in Ethiopia, likely due to the increased population size of the country. Consequently, many studies have been focused on its improvement in terms of grain yield. This breeding target usually limited in addressing other agronomically desirable traits such as biomass and related traits, which are also the needs and preferences of farmers and consumer. In this regard, landraces are highly preferred by the farming community for their biomass that contributes to biofuel generation in addition to its other uses (Derese Solomon *et al.*, 2018).

Ethiopia is the center of origin and diversity of sorghum (Doggett, 1965). It was thought to be domesticated in North-eastern Africa, particularly in Ethiopia, from where it had been introduced to other regions of the world with diverse agro-climatic conditions (Li *et al.*, 2010). Ethiopia also ranks first among countries that have contributed many germplasm collections to the world collections of sorghum at both International Crop Research Institute for Semi-Arid Tropics (ICRISAT) (Roa *et al.*, 2002) and Germplasm Resource Information Network (GRIN, 2012), indicating that the country is a rich source of sorghum landraces. Landraces are cultivated forms of a crop species, which have evolved over generations of selections by farmers (Frankel *et al.*, 1995). They are characterized by high genetic heterogeneity, good adaptation to local environmental conditions and low productivity. They are also considered as key sources of useful genes required for further increment and maintenance of the productivity of modern varieties (Amsalu Ayana, 2001).

The knowledge of genetic diversity of a crop usually helps the breeder in choosing desirable parents for breeding program to improve desirable agronomic traits such as yield, yield related

traits, biomass traits, biotic stress resistances and abiotic stress tolerances (Allard, 1999). Genetic diversity analysis is also very useful for wise utilization of the available genetic resource and also for economical germplasm conservations. The genetic variability in sorghum could be dissected using morphological assays that generally require neither sophisticated equipment nor preparatory procedures (Shehzad, 2009; Asfaw Adugna, 2014) which made the method less reliable due to the environmental influences (Van Beuningen and Busch, 1997). Recently, the less environmentally influenced genetic molecular markers become the method of choice for genetic diversity analysis. Molecular marker-based germplasm characterization does not require previous pedigree information (Bohn *et al.*, 1999).

So far , various marker systems have been developed and widely used in sorghum genetic diversity analysis like Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), microsatellite, Diversity Arrays Technology (DArT) and single nucleotide polymorphism (SNP) (Ruiz-Chután *et al.*, 2019; Zhu *et al.*, 2020).

Microsatellites also known as simple sequence repeats (SSRs) markers have shown their great potential in detecting the genetic diversity and relationships of organisms due to its allele specific and co-dominant nature. Several genetic diversity studies based on morphological traits and molecular markers have been reported for Ethiopian sorghum (Asfaw Adugna *et al.*, 2013; Asfaw Adugna, 2014). Although these studies revealed the significance of an integrated use of agro-morphological traits with molecular markers in sorghum genetic analysis, such studies are very few as relative to the huge sorghum germplasm collections the country is endowed with.

Although Ethiopia is rich in sorghum landraces as sources of desirable genes to screen genotypes for better agronomic performances (Amsalu Ayana *et al.* , 2000), genetic diversity analysis based on quantitative morphological traits coupled with molecular marker technology is greatly missing. Therefore, the present study was designed with the following general and specific objectives.

1.1. General Objective

To investigate the genetic diversity of sorghum landraces collected from various regions of Ethiopia through integrated application of microsatellite markers and quantitative morphological traits.

1.1.1. Specific Objectives

- To characterize sorghum landraces collected from different regions of Ethiopia using quantitative morphological traits.

- To determine the genetic diversity and population structure of sorghum landraces collected from different regions of Ethiopia using microsatellite markers.

2. LITERATURE REVIEW

2.1. Origin, Domestication and Distribution of Sorghum

The North-eastern part of Africa is thought to be the center of origin and domestication for cultivated sorghum, most likely in the modern Ethiopia and Sudan areas, where cultivation began about 4000-3000 BC (Dillon *et al.*, 2007). Archaeological evidence also suggests that cereal domestication was introduced to Egypt about 3000 BC from Ethiopia (Doggett, 1965). The high genetic diversity of wild sorghum crops in Ethiopia (Amsalu Ayana, 2001; Tesfaye Tesso *et al.*, 2008; Asfaw Adugna *et al.*, 2013) suggests that Ethiopia is most likely the center of origin and diversity base. *S. bicolor* is now distributed all over the world, but Africa is the continent with the greatest diversity of cultivated sorghums (Kimber *et al.*, 2013).

2.2. Taxonomy

Sorghum [*Sorghum bicolor* (L.) Moench, (2n = 2x= 20) belongs to the grass family *Poaceae* (*Gramineae*), subfamily *Panicoideae*, tribe andropogoneae, subtribe Sorghinae (Clayton and Renvoize, 1986). It is of tropical origin C4 crop (Dillon *et al.*, 2007). Since the sixteenth century, a number of writers have attempted to classify sorghum into various categories; however, (Snowden, 1936) published the first systematic classification and concise definition of sorghum. The genus *Sorghum* has 31 cultivated species and 17 related wild species (Snowden, 1955). His research paved the way for further sorghum taxonomic research in the future. Cultivated, wild, and weedy sorghums are known as *Sorghum bicolor*, which is further divided into three subspecies: *S. bicolor* subsp. *bicolor*, *S. bicolor* subsp. *drummondii*, and S. *bicolor* subsp. *Verticilliflorum* (De Wet, 1978).

Sorghum bicolor subsp. *bicolor* is the name given to cultivated sorghums, which are divided into five basic races and ten intermediate races (De Wet, 1978). Bicolor, guinea, kafir, caudatum, and durra are the five basic races. The race bicolor is considered to be of a primitive origin. Hybridization of these races results in intermediate races that have traits from both parents. These hybrid races are guinea-bicolor, caudatum-bicolor, kafir-bicolor, dura-bicolor, guinea-caudatum, guinea-kafir, guinea-durra, kafir-caudutum, durra-caudutum and kafir-durra.

Wild sorghum is grouped under *S. bicolor* subsp. *drummondii* while *S. bicolor* subsp. *verticilliflorum* constitutes hybrid derivatives of cultivated and wild sorghums (De Wet, 1978). According to Venkateswaran *et al.* (2019), the subsp. *verticilliflorum* (former *arundinaceum*) consists of four races of wild progenitors: *aethiopicum, arundinaceum, verticilliflorum*, and *virgatum*. The following is a summary of the five sorghum races' geographic distribution:

S. bicolor* subsp. *bicolor* race *durra:

The race Durra is grown in India, Burma, the Nile Valley, and Ethiopia, and this race is widely cultivated in Arabia and Asia Minor. The Ethiopian-Sudan region, the Near East, and India tend to be hotspots of three morphological diversity. Doggett (1988) argued that durra is Ethiopian in origin since the entire range of wild type, bicolor-durra, is grown extensively in the highlands of the country (Harlan and De Wet, 1972).

S. bicolor* subsp. *bicolor* race *bicolor:

The race bicolor has the highest diversity in Asia, but it is also widely distributed in Africa. Some cultivars are only found in Africa, while others are only found in Asia, Southeast Asia, and

along the coast of South China. The bicolor race is believed to have been originated in Eastern Africa from the variety aethiopicum though the high genetic diversity is found in Asia, which was arose after its introduction to the region (Kimber *et al.,* 2013).

S. bicolor* subsp. *bicolor* race *caudatum:

The race caudatum classification is based on the subseries Caffra species S. caudatum and S. nigricans (De Wet, 1978). In parts of Sudan, Chad, Nigeria, and Uganda, these are the most common. Agronomically, this race tends to be very significant when combined with other races. Caudatum is a sorghum race that has been around for a long time (De Wet, 1978).

S. bicolor* subsp. *bicolor* race *guinea:

The race is widely spread in the West Africa region and eastern Africa's mountains, which receive a high amount of rain annually. It is well suited to wet and humid environments, which other domesticated races find difficult to adapt. The race guinea was most likely derived from selection among wild members of the variety *arundinaceum*, based on morphological affinities and distribution (Kimber *et al.*, 2013).

S. bicolor* subsp. *bicolor* race *kafir:

In most parts of the sorghum-growing region of the world and West Africa including the northern Nigeria to northern Ghana, this race is widely grown, where there is a gene flow between the races guinea and kafir (Kimber *et al.,* 2013). This race appears to be mostly African in nature and its distribution and morphological affinities indicate that it is descended from the variety *verticilliflorum*.

Table 2.1. Cultivated sorghum race characteristics (Doggett, 1988).

Race	Distinct characteristics
Bicolor	– Open inflorescences with pendulous branches
	– Elliptic grain
Kafir	– Moderately compact, cylindrical inflorescences
	– Elliptic spikelets
	– Tightly clasping, long glumes
Caudatum	– Compact to open inflorescences
	– Grains with one side flat, opposite side curved
	– Shorter glumes that expose grains
Durra	– Compact inflorescences
	– Flat, ovate shaped sessile spikelets
	– Middle-creased lower glume
	– Distinct texture on tip of lower glume
Guinea	– Large, open inflorescences with pendulous branches
	– Long, separated glumes that expose grains
	– Obliquely twisted grains

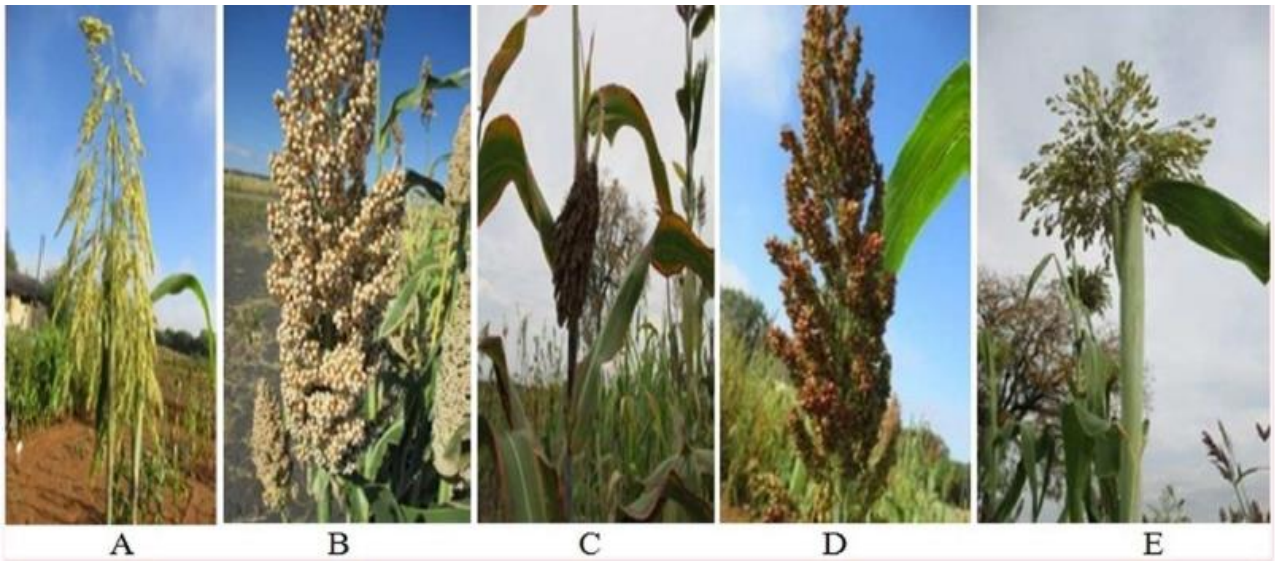

Figure 2.1. Morphological diversity in the five sorghum races: A. Guinea; B. Caudatum; C. Durra; D. Kafir; E. Bicolor. (Photo: Motlhaodi *et al.* (2017)).

2.3. Ecological Requirements

Sorghum is grown in Ethiopia and has a broader agro-ecological adaptation. Sorghum production areas are classified into three major adaptation zones, namely, lowland (below 1600 m.a.s.l), intermediate (1600 to 1900 m.a.s.l) and highland (above 1900 m.a.s.l) (Gebrekidan Brhane, 1981). Sorghum plants need between 10% and 30% clay soil, deep well-drained fertile soil, a medium to strong and reasonably steady rainfall pattern during the growing season, temperate to warm weather, and a frost-free period for optimal growth (Abera Deresa *et al.*, 2005).

Sorghum has a distinct advantage in that it can flourish in harsh conditions that are unsuitable for other crops. It tolerates a wide variety of soil types and can be grown in low soil fertility conditions. It thrives in thick vertisols like those found in the tropics, as well as light sandy soil in dryland areas (Kimber *et al.*, 2013). It is more alkaline salt resistant than other grain crops and thus can be grown successfully on soils with a pH of 5.5 to 8. Due to its ability to adapt to

conditions such as drought, salinity, and high temperatures, sorghum can be considered one of the best crops that tolerates climate change (ICRISAT, 2015).

Sorghum can be adapted in a wide variety of precipitation and temperature conditions. Because of its drought tolerance and adaptability, it has become a common crop in arid and marginal areas. It is a warm weather crop requiring high temperatures to germinate and grow well. It is a short-day plant, which means it requires short days (and long nights) before moving on to the reproductive stage. During flower initiation, sorghum plants are most susceptible to photoperiod (Alemu Tirfessa, 2018).

2.4. Sorghum Production and its Constraints

In 2019, the world-harvested area of sorghum is about 40,074,667 hectares. Sorghum production in the world is estimated to be about 57,893,378 tonne per year (FAOSTAT, 2019). According to (FAOSTAT, 2019), in terms of production share of sorghum by growing region in the year 2019, Africa is the top producer, with about 49.4% (28,619,588) tonne whereas Oceania is the lowest producer which accounts about 2% (1,165,471 tonne). Other countries such as America, Asia and Europe produced about 32.6% (18,875,728 tonne), 13.6 % (7,862,817 tonne), 2.4 % (1,369,774 tonne), respectively.

Ethiopia is the third top global producer of sorghum with 5.265,580 tonne next to Nigeria (6,665,000 tonne) and United States of America (8,673,480 tonne) and Burkina Faso is the lowest sorghum producing country with about 1,871,791 tonne (FAOSTAT, 2021). Sorghum is

grown in almost every region of Ethiopia, covering an approximate total land area of 1.82 million hectares.

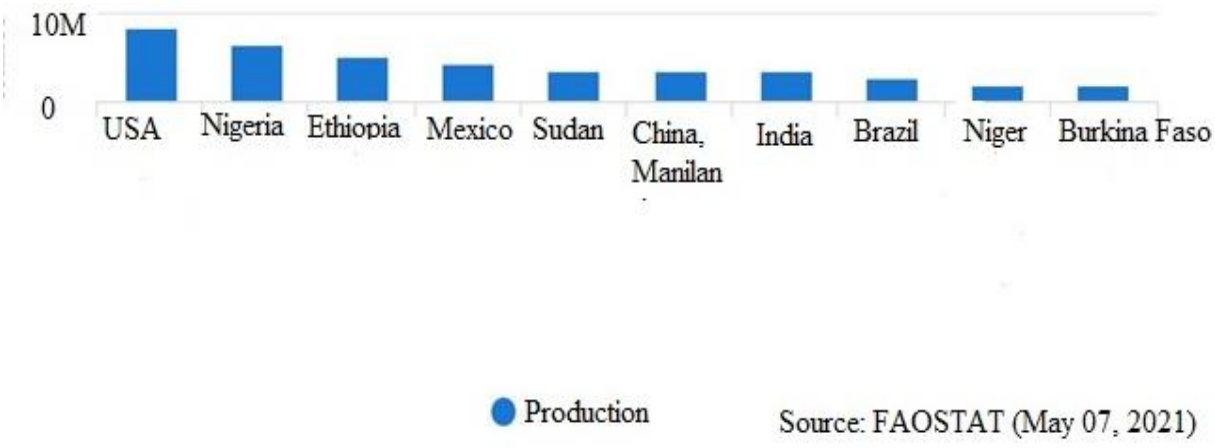

Figure 2.2. Top 10 sorghum producer countries.

According to the Central statistics Agency of Ethiopia (CSA, 2020), out of the total grain crop area, 81.46% (10,478,218.03 hectares) was under cereals. Teff, maize, sorghum and wheat took up 24.11% (about 3,101,177.38 hectares), 17.68% (about 2,274,305.93 hectares), 14.21% (1,828,182.49 hectares) and 13.91% (1,789,372.23 hectares) of the grain crop area, respectively. Cereals contributed 88.52% (296,726,476.94 quintals) of the grain production. Maize, teff, wheat and sorghum made up 28.75% (96,357,345.00 quintals), 17.11% (57,357,101.87 quintals), 15.86% (53,152,703.28 quintals) and 15.71% (52,655,800.59 quintals) of the grain production, in the same order.

Table 2.2. Estimate of area and production of sorghum by region in Ethiopia (CSA, 2020).

Region	Area (hectares)	Production (quintals)
Tigray	232,636.49	7,008,560.77
Afar	2,765.09	54,479.74
Amhara	641,613.53	18,146,464.95
Oromia	714,492.91	21,626,963.10
Somali	56,365.24	942,073.45
Benishsngul Gumuz	52,005.79	1,540,894.15
SNNP	105,255.96	2,849,141.51
Gambella	3,689.05	82,773.67
Harari	8,030.19	176,684.64
Dire Dawa	11,328.24	227,764.61

*SNNP = Southern Nation Nationality and Peoples.

There are a variety of sorghum production systems that can be distinguished based on the amount of rainfall obtained and the intended use. To combat the effects of drought, multiple cropping systems have been used in low moisture areas, but sole cropping has been identified as the dominant cropping method in Ethiopia's main sorghum growing area (Taye Mindaye *et al.*, 2016).

In Ethiopia, biotic and abiotic factors, primarily drought and a parasitic weed called Striga (*Striga spp.)* in the lowland climate, have hampered sorghum productivity. Drought stress is a common occurrence that poses serious challenges to sorghum productivity; in particular, late-

maturing sorghum landraces suffer as a result of the prolonged dry spell during the early stages of plant growth. In lowland ecosystems, the increasing prevalence of *Striga spp*., combined with extreme moisture stress, complicates varietal growth and sorghum production (EIAR, 2014).

2.5. Importance and Utilization of Sorghum

Sorghum is mainly used as a food grain in developing countries. It is also used for feed stalk, biofuel, and fencing in developing countries, but it is mostly used for animal feed and biofuels in developed countries. It has been grown as a field crop for centuries and is the primary source of food for 500 million people in 30 countries across Sub-Saharan Africa and Asia (Gabisa Ejeta and Grenier, 2005; Taye Mindaye *et al.*, 2016). In this area, sorghum is widely used for food and it is the principal source of energy, protein, vitamin and minerals. In Western countries, there is growing interest in developing sorghum's potential for use in human foods and beverages, especially as a gluten-free food source (Norwood, 2015).

Sorghum grain is also used for alcoholic beverages, fodder, and marketing, while the stalk is used for fencing, fodder, firewood, and the production of brooms. Sorghum is often used as a dual-purpose crop, with cattle grazing on the stubble after the grain is harvested. Animal feed includes both sorghum grain and plant biomass (leaves and stalks). Due to its adaptability to dry conditions, it is a less expensive alternative to maize and needs less water to achieve comparable yields (FAOSTAT, 2015). Sugar and syrup are made from the stalks of sweet sorghum varieties with high sugar content (CGIAR, 2015).

Sorghum ethanol and ethanol by-products are in high demand today. In response to concerns about climate change, biofuels are being produced to replace fossil-based transportation fuels. The sugars stored in the stalks of sweet sorghum varieties and the starch in the seeds of grain sorghum varieties are used to make ethanol in the biofuel industry (O'Hara *et al.*, 2013). Different races or cultivars of *S. bicolor* may be described as grain sorghum, fodder sorghum, or sweet sorghum depending on their morphology or end use (Purseglove, 1972).

For the preparation of the staple leavened bread (injera), sorghum grain is preferred next to teff, a small cereal grain crop. While grain quality varies depending on the end-use product, larger seed size, white, and light red sorghum grains are commonly used to make injera. The grain has also been used for the preparation of locally prepared beverages. In addition, the stories are equally valued as the grain, which can be used as animal feed, fuel wood and for construction purposes (Taye Mindaye *et al.*, 2016).

2.6. Biomass of *S. bicolor*

Biomass has been recognized as one of the most promising renewable energy sources for biofuel production, since it is an abundant resource, has low CO_2 emissions, and low cost (Berndes *et al.*, 2003; Antonopoulou *et al.*, 2008). Biomass is a broad term that includes phytomass (plant biomass) which contains chemical energy converted from solar energy through the photosynthesis process in plants (Saidur *et al.*, 2011). Sorghum is a very important high biomass producer C4 plant (Arya *et al.*, 2014; Kidanemaryam Wagaw, 2019). Number of effective tillers, stalk diameter, number of leaves, internode length and plant height are among the important

traits of sorghum biomass components. The sorghum stem can be utilized as a carbon precursor to prepare highly porous carbon materials (Kong *et al.*, 2020).

Unlike many other sources, all parts of sorghum including its grain, its sugary juice from the stem, and its biomass can be converted to ethanol (Barbanti *et al.*, 2006; Wang *et al.*, 2008). Furthermore, plant's biomass can contribute to a stabilization of farmers' incomes, and can maintain and improve ecological and social sustainability (Parikka, 2004; Xiong *et al.*, 2008). Sweet sorghum is one of the most important energy crops. It was found to contain 43.6-58.2% soluble sucrose, glucose and fructose in the stalk and is recognized as one of the most promising ethanol crops (Amaducci *et al.*, 2004).

2.7. Genetic Diversity and Population Structure of *S. bicolor*

Genetic diversity is the variation among alleles of genes in different individuals of populations of a species (IBPGR and ICRISAT, 1993). Knowledge of genetic diversity of a crop usually helps the breeder in choosing desirable parents for the breeding program and gene introgression from distantly related germplasm to maintain sustainable crop productivity concerning disease and insect resistance as well as tolerance to abiotic stresses, such as drought, cold, and salinity (Allard, 1999). The more diverse genotypes or accessions are crossed to supply superior hybrids with resistance to abiotic and biotic stresses (Jayaramachandran *et al.*, 2011). Morphological assays generally require neither sophisticated equipment nor preparatory procedures (Shehzad *et al.* , 2009; Asfaw Adugna, 2014).

However, morphological traits based characterization alone has its own drawbacks because of limited number of markers, which are influenced by the environment and show continuous variation, resulting in low heritability and high genotype by environment interactions (Van Beuningen and Busch, 1997). On the other hand, molecular markers reflect genetic similarity and differences without being influenced by the environment. Furthermore, molecular marker-based characterization does not require previous pedigree information (Bohn *et al.,* 1999).

To overcome these problems many molecular techniques have been widely used in genetic diversity assessment because at the DNA sequence level, DNA markers reveal abundant and neutral variation (Collard *et al.,* 2005). Molecular markers have played a crucial role in conservation, and use of plant genetic resource improvement programs of various crops (Aldrich and Doebley, 1992; Rami *et al.*, 1998; Deu *et al.*, 2006; Wang *et al.*, 2006; Morris *et al.*, 2013; Motlhaodi *et al.,* 2017; Alemu Tirfessa *et al.,* 2020). Various marker systems like Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), microsatellite, Diversity Arrays Technology (DArT) and single nucleotide polymorphism (SNP) were employed in sorghum research (Menz *et al.,* 2002; Mace *et al.*, 2009; Ramu *et al.*, 2009; Ruiz-Chután *et al.*, 2019; Zhu *et al.*, 2020).

Microsatellites also known as simple sequence repeats (SSRs) markers have shown their great potential in detecting the genetic diversity and relationships of organisms due to its allele specific and co-dominant nature. Additionally, microsatellite markers are the popular marker system for several sorghum genomics and molecular breeding applications due to its technical simplicity, high throughput level and potential for automation (Caniato *et al.*, 2007; Ali *et al.*, 2008; Deu *et*

al., 2008). Several morphological traits and molecular markers based studies have been reported for sorghum in Ethiopia (Gebrekidan Brhane, 1981; Amsalu Ayana and Endashaw Bekele, 1998; Assar *et al.*, 2005; Tesfaye Tesso *et al.*, 2008; Asfaw Adugna *et al.*, 2013; Asfaw Adugna, 2014; Tesfaye Disasa *et al.*, 2017).

The non-random distribution of genotypes among individuals within a population is referred to as population structure; in structured populations, mating does not occur at random. It is caused by the unequal distribution of alleles among different ancestry subpopulations. In a selection of genotypes, population structure correction is important, particularly in a breeding program where genetic relationships between breeding lines are highly variable (Sukumaran and Yu, 2014). The study of population structure in *S. bicolor* has been investigated using SSR markers (Wang *et al.*, 2009; Billot *et al.*, 2013; Tesfamichael Abraha *et al.*, 2014; Weerasooriya *et al.*, 2016; Alemu Tirfessa *et al.*, 2020; Zhu *et al.*, 2020).

3. MATERIALS AND METHODS

3.1. Characterization of Sorghum Accessions for Morphological Quantitative Traits

3.1.1. Description of the study areas

The field experiment was conducted in Ethiopia at three study locations, Melkassa (Melkassa Agricultural Research Center, MARC), Mehoni (Mehoni Agricultural research Center, MHARC) and Meisso (sub center of MARC at Miesso site). The details of the study locations are presented in Table 3.1.

Table 3.1. Summary information about the study locations (EIAR, 2021).

Location	Distance from Addis Ababa (km)	Latitude	Longitude	Altitude (m.a.s.l)	Average annual Rain fall (mm)	Annual Temperature (°C)	
						Min	Max
Melkassa	117	8°24’99" N	39°19’53" E	1550	763	14	28.4
Meisso	360	9° 03' 49" N	40° 52' 26" E	1855	980	21	29
Mehoni	678	12° 07' 31" N	39°17' 40" E	1574	525	18	25

3.1.2. Plant materials

The entries for this study consisted of 324 (320 landraces and 4 improved genotypes, B35, Melkam, Argiti and ESH-1). Of these 265 accessions were obtained from Melkassa Agricultural Research Center (MARC) which were originally collected by the Ethiopian Biodiversity Institute (EBI) and 59 accessions were newly collected (code = SB) from different parts of the country.

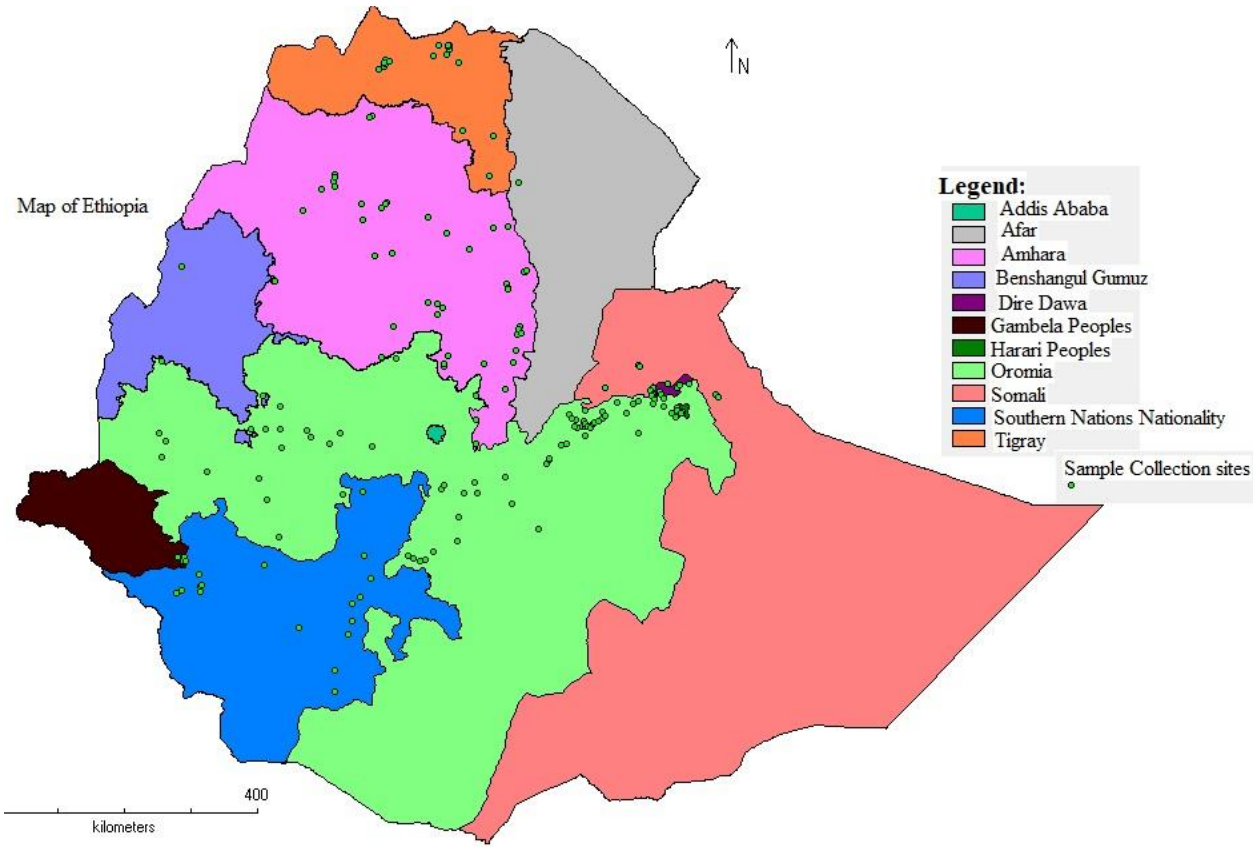

Figure 3.1. Sorghum accessions collection sites and regions depicted on the map of Ethiopia based on geographic coordinate information.

3.1.3. Experimental design

The field experiment was conducted at three locations during the main cropping season of 2018/19. The experiment was laid in α- lattice design with two replications having 12 incomplete blocks per replication and 27 entries per block. The spacing was maintained at 20 cm between plants, 75 cm between accessions, 1m between blocks and 2m between replications. After two weeks of sowing, thinning out was done. After thinning, each plot consisted of an average of 15 sorghum plants in a row of 3m long with an area of 2.25 square meters (3m x 0.75m). Five plants from each plot were randomly selected for substituent collections of data on morphological quantitative traits. Fertilization followed the recommended amount (100 kg/ha)

of di-ammonium phosphate (DAP) during planting. Similar amount of urea was supplemented as a split application following the procedure used by (Amsalu Ayana, 2001).

All the necessary agronomic practices were followed according to the recommended practices of each site. A minimum of five plants were selected immediately after flowering and the heads were covered with suitable and durable paper bags in order to protect from neighboring pollen contamination. The panicles were further covered with selected cloth bags designed for the protection of seeds against birds.

3.1.4. Quantitative trait data collection

For every accession, five individuals were used for recording data for quantitative traits on plot basis in each replicate. Data on six quantitative traits (number of effective tillers, stalk diameter, number of leaves, internode length, plant height and dry biomass) were collected based on sorghum descriptor list of International Plant Genetic Resources Institute and International Board for Plant Genetic Resources (IBPGR and ICRISAT, 1993) as shown in Table 3.2.

Table 3.2. Summary of quantitative traits used to assess the diversity of 324 sorghum accessions (IBPGR and ICRISAT, 1993).

No	Parameter	Description of the character	Time of scoring
1	Number of effective tillers	Count of all effective basal branches of a plant	Physiological Maturity
2	Stalk diameter (mm)	Measured at 20 cm above ground using digital caliper (World Precision Instruments, Shanghai Trading Co., Ltd)	
3	Number of leaves	Count of all leaves on a main stalk	
4	Internode length (cm)	Length measured at the center of plant stalk	
5	Plant height (cm)	Length of plant from ground to the tip of the panicle	
6	Dry biomass (Kg)	Weight of sun-dried biomass	

3.1.5. Quantitative trait data analysis

The normality of the data for the six traits for all locations and their combination was determined using the R software version 4.0.5. The same software program was used to analyise mean performance, descriptive statistics, analysis of variance (ANOVA), heritability in broad-sense, genetic advance, clustering, principal component analysis and eigenvalue across the entire accessions (R: The R Project for Statistical Computing (r-project.org)). Before clustering and principal component analysis, records for all traits were pre-standardized to means of zero and variance of unit to avoid bias due to differences in measurement scales according to Sneath and Sokal (1973) using the same software.

All factors were treated as random variables. The estimates of phenotypic and genotypic correlation coefficient were analyzed using META-R (Multi Environment Trail Analysis with R for Windows) Version 6.04.

3.1.5.1. Analysis of variance

Table 3.3. The skeleton of alpha lattice design structural analysis of variance (ANOVA) (Patterson and Williams, 1976).

Sources of variation	Degrees of Freedom	Sum of squares	Mean squares	F
Replications	r-1	SSr	MSr	MSr/MSe
Blocks (within replications)	rs-r	SSb	MSb	MSb/MSe
Treatments (adjusted for blocks)	t-1	SSt	MSt	MSt/MSe
Error	rt-rs-t+1	SSe	MSe	
Total	tr-1	SST		

The linear model of observations in alpha lattice design at each location was:

$$y_{ijk} = \mu + ti + rj + bjk + e_{ijk},$$

where; y_{ijk} represents the value of the observed trait for i-th treatment received in the k-th block within j-th replicate (superblock), t_i is the fixed effect of the i-th treatment (i = 1,2,…,t); r_j is the effect of the j-th replicate (super block) (j = 1,2,…,r), b_{jk} is the effect of the k-th incomplete block within the j-th replicate (k = 1,2,…s), and eijk is an experimental error associated with the observation of the i-th treatment in the k-th incomplete block within the j-th complete replicate.

The linear model of observations in alpha lattice design for combined locations was:

$y_{ijksl} = \mu + ti + rj + bjk + Ls + tilk + e_{ijksl}$,

where; y_{ijksl} represents the value of the observed trait for i-th treatment received in the k-th block within j-th replicate (superblock), t_i is the fixed effect of the i-th treatment (i = 1,2,…,t); r_j is the effect of the j-th replicate (super block) (j = 1,2,…,r), b_{jk} is the effect of the k-th incomplete block within the j-th replicate (k = 1,2,…s), Ls = locations' effect, tilk = location by treatment effect and eijksl is an experimental error associated with the observation of the i-th treatment in the k-th incomplete block within the j-th complete replicate (Table 3.3).

3.1.5.2. Estimation of heritability in broad sense and components of variance

The phenotypic and genotypic variance were estimated according to the methods suggested by (Burton and Devane, 1953) and these components of variance (σ^2p, σ^2e and σ^2g) were used for the estimation of coefficients of variation (PCV, GCV) as described by Singh Chaudhary (1977). Genetic advance (GA) and genetic advance as percent of the mean (GAM) for each traits and heritability in broad sense (H^2) (K=2.06 at 5% selection intensity) were calculated for each character based on the formula developed by Allard and Wiley (1960):

Genetic variance (σ^2g) = (MSG-MSgl)/rl

Phenotypic variance (σ^2p) = $\sigma^2g + (\sigma^2gl)/l + \sigma^2 e /rl$

Where, MSG = mean square due to genotype, σ^2gl = mean square due to genotype by environment interaction, l = location, δ^2e = combined error mean square

The phenotypic coefficients of variation (PCV) and genotypic coefficients of variation (GCV) were estimated as percentage of the corresponding phenotypic standard deviation (σp), genotypic standard deviation (σg) of the trait and dividing this value to μ (the grand populations mean for the trait).

Phenotypic coefficient of variation (PCV) = (σpx100)/μ

Genotypic coefficient of variation (GCV) = (σgx100)/μ

Heritability (H^2) = (σ^2gx100)/ σ^2_p

Expected genetic advance (GA) = H^2 x k x σ_p

Expected genetic advance as percentage of mean (GAM) = (GAx100)/μ

Where, k is a constant value at selection intensity of 5% (k = 2.06), σ_p is the phenotypic standard deviation; H^2 is broad sense heritability; and μ is the grand populations mean for the trait under considerations.

3.1.5.3. Genetic diversity and cluster analysis

The genetic diversity among sorghum accessions was done based on multivariate analysis of hierarchical clustering of genotypes with the ward D^2 method and the cluster analysis was performed using a measure of similarity levels and Euclidean distance (Eisen *et al.*, 1998) using R software. The numbers of clusters were decided by using the 'within sum of square' technique where the location of a bend (knee) in the scree plot is generally considered as an indicator of the appropriate number of clusters.

3.1.5.4. Principal component analysis

According to Iezzoni and Pritts (1991), principal component analysis (PCA) is probably the most popular multivariate statistical technique, which is used by almost all scientific disciplines and a principal component has inherently more information than would any single variable alone. A principal component based on the correlation matrix was computed using R software. The principal components having an Eigenvalue greater than one were considered as significant (Kaiser, 1961).

3.2. Genetic Diversity and Population Structure of Sorghum Accessions using Microsatellite Markers

3.2.1. Plant material

From the 324 sorghum accessions collected from major sorghum growing regions of Ethiopia described in Appendix 1 and Figure 3.1, the first 100 accessions were used for genetic diversity analysis. These accessions were grown on field at Melkassa Agricultural Research Center (MARC) and fresh leaves of five plants from two-week-old seedlings were harvested in bulk, dried with silica gel, and then used for DNA extraction.

3.2.2. DNA extraction

Six milligram of leaf sample from each individual was used to form a bulk sample making a pool of 30 mg leaf sample. Genomic DNA was extracted according to DArT plant DNA extraction protocol (DArT, 2019) (Appendix 2). Quality were checked with Nano-drop spectrophotometer and on 1% agrose gel electrophoresis. High molecular weight extracted DNA was stored in -20 °C until use for genotyping.

3.2.3. Primer selection and optimization

The SSR primers were referred from previously published sorghum scientific literatures (Ng'uni *et al.* , 2012; Tesfaye Disasa, 2016; Alemu Tirfessa *et al.*, 2020). The primers' polymorphism and reproducibility were checked on eight randomly selected sorghum genotypes. Before use in PCR reaction, the extracted DNA concentration was normalized to 100 ng/µl using nuclease free sterile (double distilled) water. For genotyping, a total of 15 polymorphic primers were used

(Table 3.4). The primers annealing temperatures were optimized using gradient PCR (HIMEDIA – Prima-96™ Thermal Cycler 96 wells block) machine. The PCR program for the first four primer pairs involved an initial denaturation at 95 ºC for 5 minutes followed by 40 cycles of denaturation at 95 ºC for 30 seconds, optimized annealing temperatures of 44, 50 or 52 ºC for 2 min (Table 3.4), and primer extension at 72 ºC for 2 minutes followed by a final extension at 72 ºC for 10 minutes and holding at a temperature of 4 ºC. The PCR conditions for the next seven primer pairs (5-11) in Table 3.4 involved an initial denaturation at 94 °C for 5 min followed by 35 cycles with 1 min of denaturation at 95 °C, 1 min annealing at 54 or 56 °C. The primer extension was at 72 °C for 2 min followed by a final extension step of 10 min at 72 °C and holding at 4 °C (Table 3.1; Figure 4.3).

The PCR program for the last four primer pairs (12-15) involved an initial denaturation at 94 °C for 5 min followed by 44 cycles with 1 min of denaturation at 95 °C, 2 min annealing at 58 or 60 °C. The extension was at 72 °C for 2 min followed by a final extension step of 10 min at 72 °C before holding at 4 °C. The PCR products were separated on 3% agarose gel electrophoresis using 1 X TBE buffer at 100 V for 3 hr. The gel was stained in gel red (2 µl), visualized under UV light and subsequently photographed using a BIO-RAD Gel Doc TM EZ Imaging System. A 100 bp DNA ladder was used to estimate the size of the amplified fragments (Table 3.4; Figure 4.3).

Table 3.4. Primers sequence, repeat motif, chromosome number, annealing temperatures and their range of molecular weight detected using 15 microsatellite loci in 100 accessions representing five populations of sorghum in Ethiopia.

Marker	Forward and Reverse primer (5′ to 3′)	Repeat Motif	Chromosome Number	[a]Annealing temperature (°C)	[b]Fragment Size (base pair)	
					Min	Max
SbAGB02	F: CTCTGATATGTCGTTGTGCT R: ATAGAGAGGATAGCTTATAG-CTCA	(AG)35	7	44	95	126
Xcup74	F: GTCGCCATTGTGATGAAGAG R: CAGTAGTCCAGCAAAACGGC	(TG)9	2	56	174	185
Xtxp136	F: GCGAATAGCATCTTACAACA R: ACTGATCATTGGCAGGAC	(GCA)5	5	56	238	242
Xcup50	F: TGATTGATTGAGGCAGGCAC R: TTCCGGTCTCTGTCCATTTC	(ACAGG)5	10	58	150	159
Xcup67	F: GGTCAGTGCTTACACAGATTCC R: GGGGATTGCAGGTGTCATAG	(TA)6	10	56	264	294
Xcup14	F: TACATCACAGCAGGGACAGG R: CTGGAAAGCCGAGCAGTATG	(AG)10	3	60	209	237
Xgap084	F: CGCTCTCGGGATGAATGA R: TAACGGACCACTAACAAATGATT	(AG)14	2	50	176	209
Xtxp010	F: ATACTATCAAGAGGGGAGC R: AGTACTAGCCACACGTCAC	(CT)14	9	60	139	148
Xtxp012	F: AGATCTGGCGGCAACG R: AGTCACCCATCGATCATC	(CT)22	4	56	175	200
Xtxp021	F: GAGCTGCCATAGATTTGGTCG R: ACCTCGTCCCACCTTTGTTG	(AG)18	4	52	175	190
Xtxp141	F: TGTATGGCCTAGCTTATCT R: CAACAAGCCAACCTAAA	(GA)23	10	50	141	182
Xtxp145	F: GTTCCTCCTGCCATTACT R: CTTCCGCACATCCAC	(AG)22	6	54	120	143
Xtxp265	F: GTCTACAGGCGTGCAAATAAAA R: TTACCATGCTACCCCTAAAAGTGG	(GAA)19	6	56	183	220
Xtxp273	F: GTACCCATTTAAATTGTTTGCAGTAG R: CAGAGGAGGAGGAAGAGAAGG	(TTG)20	8	56	198	231
Xtxp320	F: TAAACTAGACCATATACTGCCATGATAA R: GTGCAAATAAGGGCTAGAGTGTT	(AAG)20	1	58	260	288

Max = minimum, Max = maximum, [a], [b] = resulted in this study. Sources: (Ng'uni *et al.*, 2012; Tesfaye Disasa, 2016; Alemu Tirfessa, 2020).

3.3. Microsatellite Markers Data Scoring and Analysis

The sizes of all PCR-amplified SSR regions were estimated using the PyElph software package (Pavel and Vasile, 2012), in relation to the size marker. The scored marker data were used to analyze genetic diversity and population structure. The standard indices of genetic diversity were calculated using different statistical software packages. PowerMarker v3.25 software (Liu and Muse, 2005) was used to compute locus-based diversity indices across the entire populations, including major allele frequency (MAF) and polymorphic information content (PIC).

POPGENE version 1.3.1 (Yeh, 1999) was used to assess population differentiation (Gst) and gene flow (Nm = 1/Fst) −1]/4 across populations. Over the entire population, allelic frequency, the number of observed alleles (Na), gene diversity (h), effective number of alleles (Ne) and Shannon's Information index (I) were computed using GenAlEx ver. 6.501 software (Peakall and Smouse, 2006). This software was also used to calculate population diversity indices over all loci including Na, Ne, I, the number of private alleles (NPA), Nei's gene diversity (h) and Percentage of polymorphic loci (PPL). The same software was used to compute pairwise population genetic distances and gene flow, and to perform the genetic differentiation test (Gst and p- values) over 999 bootstrap replications.

Arlequin ver. 3.5.2.2 (Excoffier and Lischer, 2006) was used to compute the analysis of molecular variance (AMOVA) and the estimation of the variance components. Gene flow (Nm) among populations was estimated using the formula, Nm (diploid) = [(1 / Fst) – 1] / 4, where Fst = the variance among populations/total genetic variations. Nei's standard genetic distance (Nei and Takezaki, 1983) based Unweighted Pair Group Method with Arithmetic Mean (UPGMA)

trees were generated using PowerMarker v3.25 (Liu and Muse, 2005). The resulting trees were visualized using MEGA software (Kumar *et al.*, 2018) and Tree View (built in PowerMarker v3.25). Significance was tested based on 1000 bootstrap replications (Felsenstein, 1985). This software was also used to compute PIC, heterozygosity (Ho) and gene diversity (expected heterozygosity, He) for each marker as well as the average across markers. Rare (<5%), common (5-50%), abundant (>50%), and private alleles (alleles only present in one group not in other) were manually counted in MS Excel (Microsoft Inc., USA).

Population structure and admixture patterns were determined using STRUCTURE software ver. 2.3.4 with a Bayesian model-based clustering algorithm (Pritchard *et al.*, 2000). To estimate the true number of population cluster (K), a burn-in period of 100,000 was used in each run, and data were collected over 250,000 Markov Chain Monte Carlo (MCMC) replications for K = 1 to K = 10 using 20 iterations for each K. The optimum K value was predicted following the simulation method of (Evanno *et al.*, 2005) using the web-based STRUCTURE HARVESTER ver. 0.6.92 (Earl, 2012). A bar plot for the optimum K was determined using Clumpak beta version (Kopelman *et al.*, 2015).

4. RESULTS

4.1. Genetic Diversity Analysis of Sorghum Accessions based on Quantitative Traits

4.1.1. Visual classification and quantitative traits performance analysis of sorghum accessions

The 324 sorghum accessions were found to have a wide diversity in terms of numbers of effective tillers, stalk diameter, leaf numbers, inter-node length, plant height and dry biomass across the study areas. The mean of the studied traits ranged from 3.90 – 3.99 for number of effective tillers, 19.13 – 20.20 cm for stalk diameter, 17.86 – 18.33 for leaf numbers, 18.23 – 20.17 cm for internodes length, 265.60 – 277.20 cm for plant height and 11.51 – 12.36 tonne per hectare for dry biomass (Table 4.1).

4.1.1.1. Mean performance and range of sorghum accessions over locations

Plant height and stalk diameter

The tested genotypes showed significant difference for mean plant height length and stalk dimeter. The tallest plant was observed in accession 69544 (382.83 cm) followed by accession 69284 (372.08 cm). Conversely, the shortest plant height was recorded from B35 (90.83 cm). Across all accessions, the average plant height was 272.12 mm (Table 4.2). Accessions planted at Miesso had the highest mean plant height (277.20 cm), followed by Melkassa (273.50 cm). Mehoni, on the other hand, had the lowest mean height value of 265.60 cm (Table 4.1).

The largest stalk diameter measure was observed in accession 243219 (22.60 mm) followed by 27921 (22.27 mm), whereas the lowest was recorded from accession 19625 (16.43 mm). In terms of stalk diameter, the thickest was observed in accessions from Melkassa (22.6 mm) followed by Mehoni (19.28 mm) (Table 4.1). While, the smallest stalk diameter was scored at Miesso (19.13 mm).

Leaf number and internodes length

The large number of leaves was observed in accession 241208 (20.53) followed by 69529 (19.90), whereas the least was recorded from SB120 (14.93). The average number of leaf number is 18.04 (Table 4.2). The highest leaf number was obtained from Melkassa (18.33) followed by Mehoni (17.86) while the lowest value was recorded from Miesso (17.86) (Table 4.1). The longest internodes length was observed in accession 223513 (21.57 cm) followed by 241239 (21.53 cm), whereas the shortest was recorded from SB16 (15.67 cm). The highest internodes length was obtained from Melkassa (20.17 cm) followed by Mehoni (18.68 cm), and the smallest was recorded from Miesso (18.23 cm) (Table 4.1).

Number of effective tillers

The largest number of effective tillers was observed in accession 223517 (4.54) followed by 236767 (4.51), whereas the smallest was recorded from accession 237038 (3.40). The average value of the number of effective tillers for all study areas was 3.96 (Table 4.2). The largest number of effective tillers were recorded from Melkassa (3.99) followed by Mehoni (3.96) and the lowest at Miesso (3.90) (Table 4.1).

Dry biomass

The highest biomass yield was exhibited by the accession 69232 (20.90 tonne ha^{-1}) followed by accession 241187 (19.333 tonne ha^{-1}), whereas the lowest biomass yield was recorded from accession 210957 (8.22 tonne ha^{-1}). The highest dry biomass was obtained at Miesso (12.36 tonne ha^{-1}) followed by Melkassa (11.73 tonne ha^{-1}) and the smallest value was recorded at Mehoni (11.51 tonne ha^{-1}). The average value for the entire locations was 11.85 tonne ha^{-1} (Table 4.1; Appendix 1).

Table 4.1. A summary of mean, range and standard deviation of the six quantitative traits of 324 sorghum accessions over three environments.

Trait	Environment	Mean	Range	Standard Deviation	CV (%)	Pr > F
NET (number)	Melkassa	3.99	1.43	0.29	8.96	***
	Meisso	3.90	1.51	0.24	8.07	**
	Mehoni	3.96	1.95	0.26	8.12	***
SD (mm)	Melkassa	20.20	9.20	1.32	8.61	**
	Meisso	19.13	10.40	1.57	10.87	**
	Mehoni	19.28	9.20	1.28	9.69	ns
LN (Number)	Melkassa	18.33	7.55	1.33	9.13	***
	Meisso	17.86	7.80	1.35	9.59	***
	Mehoni	17.94	6.50	1.25	8.89	***
IL (cm)	Melkassa	20.17	7.80	1.32	8.77	*
	Meisso	18.23	9.10	1.51	10.42	***
	Mehoni	18.68	8.50	1.35	9.87	*
PH (number)	Melkassa	273.50	310.00	55.02	21.74	***
	Meisso	277.20	310.00	52.85	20.81	***
	Mehoni	265.60	305.00	52.16	22.58	***
DB(tonne ha^{-1})	Melkassa	11.73	14.18	0.40	17.90	***
	Meisso	12.36	12.62	0.42	19.35	***
	Mehoni	11.51	10.84	0.32	14.70	***

NET = number of effective tillers, SD = stalk diameter, LN = number of leaf, IL = internode length, DB = dry biomass, mm = millimetres, cm = centimeters, ha = hectare, CV (%) = coefficient of variation in (per cent), Pr > F = Probability levels: *** = highly significant at $p < 0.001$, ** = highly significant at $p < 0.01$, * = Significant at $p < 0.05$ and ns = non-significant.

Table 4.2. Variance and descriptive statistics for six traits of 324 sorghum accessions

Variable	Maximum	Mean	Minimum	Standard Error	Standard deviation	Variance	CV (%)
NET	4.54	3.95	3.40	0.01	0.20	0.04	5.05
SD	22.60	19.54	16.13	0.045	0.85	0.72	4.33
LN	20.53	18.04	14.93	0.05	0.92	0.85	5.12
IL	21.57	19.02	15.67	0.05	0.90	0.81	4.73
PH	382.83	272.12	90.83	2.58	46.43	2155.42	17.06
DB	23.67	11.85	5.47	0.02	0.30	0.09	11.26

NET = number of effective tillers (No.), SD = stalk diameter (mm), LN = number of leaf (No.), IL = internode length (cm), PH = plant height (cm), DB = dry biomass (tonne ha^{-1}), mm = millimeters, cm = centimeters, CV (%) = Coefficient of variation (per cent).

4.1.1.2. Estimation of analysis of variance

Genotype by environment interaction

The accessions and location result was highly significant for all of the studied variables (number of effective tillers, stalk diameter, number of leaves, internodes length, and dry biomass), according to the results of a combined analysis of variance (Table 4.3). The replication effect was significant for the internodes length and plant height. The effect of block was highly significant for plant height. The effect of accession by environment interaction was highly significant for leaf number, internodes length, plant height and dry biomass and was significant for number of effective tillers (Table 4.3).

Table 4.3. Combined analysis of variance for six quantitative traits of sorghum genotypes tested at Melkassa, Miesso and Mehoni during the main cropping season of 2018/19.

Traits	Gen (df=323)	Rep x Loc) (df=1)	Blk (Rep) (df=22)	Loc (df=2)	Gen x Loc (df=646)	Error (df=949)
NET	0.25^{***}	0.18^{ns}	0.09^{ns}	1.48^{***}	0.09^{*}	0.08
SD	4.29^{**}	0.00^{ns}	3.08^{ns}	217.69^{***}	3.68^{ns}	3.36
LN	5.12^{***}	0.06^{ns}	1.98^{ns}	41.77^{***}	2.57^{**}	2.11
IL	4.9^{***}	12.4^{*}	3.3^{ns}	667.10^{***}	3.4^{**}	2.9
PH	12933^{***}	4608^{*}	1993^{***}	22604^{***}	2074^{***}	1267
DB	0.54^{***}	0.24^{ns}	0.171^{ns}	6.10^{***}	0.17^{**}	0.14

NET = number of effective tillers (number), SD = stalk diameter (mm), LN = number of Leaf (number), IL = internode length (cm), PH = plant height (cm), DB = dry biomass (tonne ha^{-1}), df = degree of freedom, Loc = location, CV (%) = coefficient of variation (per cent). Probability levels: *** = highly significant at probability level of $p < 0.001$, ** = highly significant at probability level of $p < 0.01$, * = Significant at probability level of $p < 0.05$ and ns = non-significant.

4.1.2. Estimation of genetic parameters

4.1.2.1. Genotypic and phenotypic coefficient of variation

The GCV values ranged from 2.66% for stalk diameter to 22.58 % for plant height and PCV values ranged from 5.39% for leaf numbers to 26.95 % for plant height (Table 4.4).

4.1.2.2. Trait heritability

In the present study, the highest broad-sense heritability obtained for plant height was 70.22% followed by a value of 45.46% for dry biomass. Number of effective tillers, leaf number and internodes length exhibited a broad sense heritability of 37.78%, 29.20% and 16.07%, respectively. Stalk diameter showed the lowest broad-sense heritability (7.20%) (Table 4.4).

4.1.2.3. Genetic advance as percent of mean

The values of expected genetic advance and genetic advance as percent of the mean (GAM) are presented in Table 4.4. In the present study, the expected genetic advance as percent of mean (GAM) ranged from 1.47% for Stalk diameter to 38.99% for plant height (Table 4.4).

Table 4.4. Mean, phenotypic and genotypic variances and coefficient of variations; heritability in broad sense and genetic advance for six traits of 324 sorghum accessions studied.

Traits	GM	σ^2e	σ^2g	σ^2p	ECV (%)	GCV (%)	PCV (%)	H^2 (%)	GA	GAM (%)
NET	3.95	0.08	0.05	0.14	7.36	5.73	9.33	37.78	0.29	7.26
SD	19.54	3.48	0.27	3.75	9.55	2.66	9.91	7.20	0.29	1.47
LN	18.04	2.29	0.94	3.23	8.39	5.39	5.39	29.20	1.08	5.99
IL	19.02	3.08	0.59	3.68	9.23	4.04	10.08	16.07	0.64	3.34
PH	272.12	1601.50	3776.10	5378.51	14.71	22.58	26.95	70.22	106.09	38.99
DB	11.85	3.05	2.54	5.59	14.74	13.45	19.95	45.46	2.22	18.68

NET = number of effective tillers (number); SD = stalk diameter (mm); LN = number of Leaf (number); IL = internode length (cm); PH = plant height (cm), DB = dry biomass (tonne ha^{-1}). GM= Grand mean, σ^2e= environmental variance, σ^2g= genotypic variance, σ^2p= phenotypic variance, ECV % = environmental coefficient of variance, GCV %= genotypic coefficient of variance, PCV %= phenotypic coefficient of variance, H^2 %= heritability (broad sense), GA= genetic advance, GAM % = genetic advance as percentage of mean.

4.1.3. Correlation analysis

Phenotypic (rp) and genotypic (rg) correlation estimates between all pairs of the six traits considered for the 324 sorghum accessions studied are presented in Table 4.5.

4.1.3.1. Phenotypic and genotypic correlation of dry biomass with other traits

The phenotypic and genotypic correlation coefficients among the six traits are shown in Table 4.5. Genotypic and phenotypic correlation analysis indicated that dry biomass showed positive and highly significant (< 0.001) association with number of effective tillers (rp = 0.53, rg = 0.21), stalk diameter (rp = 0.58, rg = 0.19), leaf number (rp = 0.48, rg = 0.28), internodes length (rp = 0.50, rg = 0.20) and plant height (rp = 0.24, rg = 0.20). This indicates that the increase in dry biomass is because of increase in one or more of the listed characters (Table 4.5).

4.1.3.2. Genotypic and phenotypic correlation among other traits

Number of effective tillers had positive and highly significant ($p<0.001$) phenotypic and genotypic correlation with leaf number (rp = 0.53, rg = 0.27), dry biomass (rp = 0.53, rg = 0.21) and plant height (rp = 0.53). This trait did not show any correlation with stalk diameter for both phenotypic and genotypic relations. It showed positive and highly significant ($p<0.001$) phenotypic correlation (rp = 0.53) but with positive and significant genotypic correlation (rg = 0.11) with internode length (Table 4.5). Stalk diameter showed positive and highly significant ($p<0.001$) phenotypic and genotypic correlation with leaf number (rp = 0.57, rg = 0.14), internode length (rp = 0.98, rg = 0.24), plant height (rp = 0.68, rg = 0.26) and dry biomass (rp = 0.58, rg = 0.19). It had a non-significant relation with number of effective tillers (Table 4.5).

Number of leaves had positive and highly significant ($p<0.001$) phenotypic and genotypic correlation with number of effective tillers (rp = 0.53, rg = 0.27), stalk diameter (rp = 0.57, rg = 0.14), internode length (rp = 0.85, rg = 0.31), plant height (rp = 0.24, rg = 0.18) and dry biomass (rp = 0.48, rg = 0.28) (Table 4.5). Likewise, internode length had positive and highly significant ($p<0.001$) phenotypic and genotypic correlation with stalk diameter (rp = 0.98, rg = 0.24), leaf numbers (rp = 0.85, rg = 0.31), plant height (rp = 0.82, rg = 0.44), dry biomass (rp = 0.50, rg = 0.20) and number of effective tillers (highly significant for rp = 0.53 and significant for rg = 0.11) (Table 4.5). Similarly, plant height had positive and highly significant ($p<0.001$) phenotypic and genotypic correlation with number of effective tillers (rp = 0.53, rg = 0.06ns), stalk diameter (rp = 0.68, rg = 0.26), internode length (rp = 0.82, rg = 0.44), leaf number (rp = 0.24, rg = 0.18) and dry biomass (rp = 0.14, rg = 0.20) (Table 4.5).

Table 4.5. Estimates of phenotypic (Above diagonal) and genotypic (Below diagonal) correlation coefficient of combined means of six traits studied for 324 sorghum accessions tested at Melkassa, Miesso and Mehoni during 2018/19 main cropping season.

Variables	NET	SD	LN	IL	PH	DB
NET	1.00	-0.07^{ns}	0.53***	0.53***	0.53***	0.53***
SD	-0.03^{ns}	1.00	0.57***	0.98***	0.68***	0.58***
LN	0.27***	0.14***	1.00	0.85***	0.24***	0.48***
IL	0.11*	0.24***	0.31***	1.00	0.82***	0.50***
PH	0.06^{ns}	0.26***	0.18***	0.44***	1.00	0.24***
DB	0.21***	0.19***	0.28***	0.20***	0.20***	1.00

NET = number of effective tillers (number); SD = stalk diameter (mm); LN = number of Leaf (number); IL = internode length (cm); PH = plant height (cm), DB = dry biomass (Kg). Probability levels: *** = highly significant at probability level of $p<0.001$, ** = highly significant at probability level of $p<0.01$, * = Significant at probability level of $p<0.05$, and ns = Non-significant.

4.1.4. Cluster and principal component analysis

Cluster analysis based on Euclidean distances as similarity measures and the unweighted pair group method with arithmetic averages (UPGMA) based on the six morphological quantitative traits revealed the presence of six clusters (Table 4.6, Figure. 4.1). The clustering detected four major clusters that further sub-grouped into six clusters. Cluster one (C1) comprised of 103 (31.79%) accessions that are sampled from EBI and newly collected accessions (Table 4.6, Figure. 4.1). The second largest cluster (C2) had 99 (30.56%) accessions that were derived from both EBI and newly collected accessions. Cluster three (C3) and C4 comprised of 60 (18.52%) and 60 (18.52%) sorghum accessions, respectively (Table 4.6, Figure. 4.1). Accessions of C1 and

C4 predominantly came from EBI. Cluster five (C5) and C6 contained 27 (8.33%) and 14 (4.32%) accessions. C5 and C6 also accommodated accessions from both EBI and newly collected accessions.

Table 4.6. Clustering of 324 sorghum accessions in to six clusters using mean of six morphological quantitative characters

Cluster	No. of Accessions
C-1	103 (31.79%)
C-2	99 (30.56%)
C-3	21 (6.48%)
C-4	60 (18.52%)
C-5	27 (8.33%)
C-6	14 (4.32%)

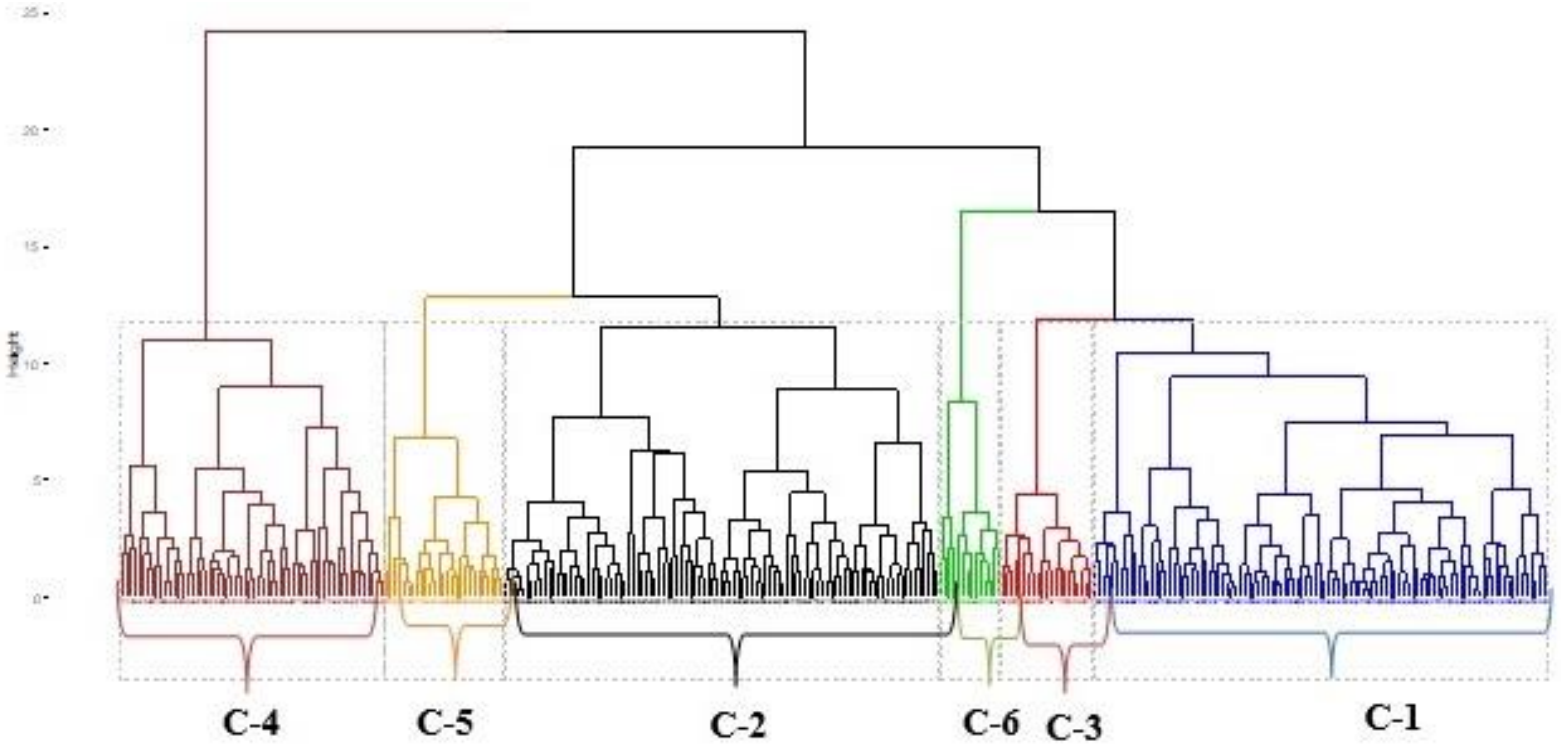

Figure 4.1. Cluster analysis of 324 sorghum accessions based on six morphological quantitative traits.

The principal component analysis of the 324 accessions based on the six agronomic traits resulted in six principal components indicated in Table 4.7. The first four PCs explained 80% of the total variations in the dataset (Table 4.7). The first and second PCs explained 34 and 20% of the total variations, respectively. LN, IL and PH were the major contributors for PC1 while, NET and SD were the major contributors for PC2. SD and DB were the major contributors for PC1 while, LN and DB were the major contributors for PC4 (Table, 4.7).

The first and second principal component cumulatively contributed 54% of the total variation. The variation in the second PC comes from highly contributing traits of number of effective tillers, leaf numbers and dry biomass with positive loading. In the third PC, stalk diameter and dry biomass caused differences positively and the rest caused differences negatively.

Table 4.7. Eigenvalue, proportion and cumulative variances and eigenvectors on the first 6 principal components for 6 traits in 324 sorghum accessions.

Variable	PC1	PC2	PC3	PC4	PC5	PC6
NET	-0.26	**0.67**	-0.14	0.00	0.68	0.07
SD	-0.34	**-0.43**	**0.59**	0.42	0.41	0.04
LN	**-0.44**	**0.35**	-0.03	**0.56**	-0.49	-0.37
IL	**-0.50**	-0.21	-0.43	0.05	-0.14	0.70
PH	**-0.45**	-0.37	-0.37	-0.37	0.18	-0.59
DB	-0.41	0.23	**0.56**	**-0.61**	-0.28	0.11
Eigenvalue	2.07	1.17	0.85	0.72	0.69	0.51
Proportion (%)	0.34	0.20	0.14	0.12	0.12	0.09
Cumulative (%)	0.34	0.54	0.68	0.80	0.92	1.00

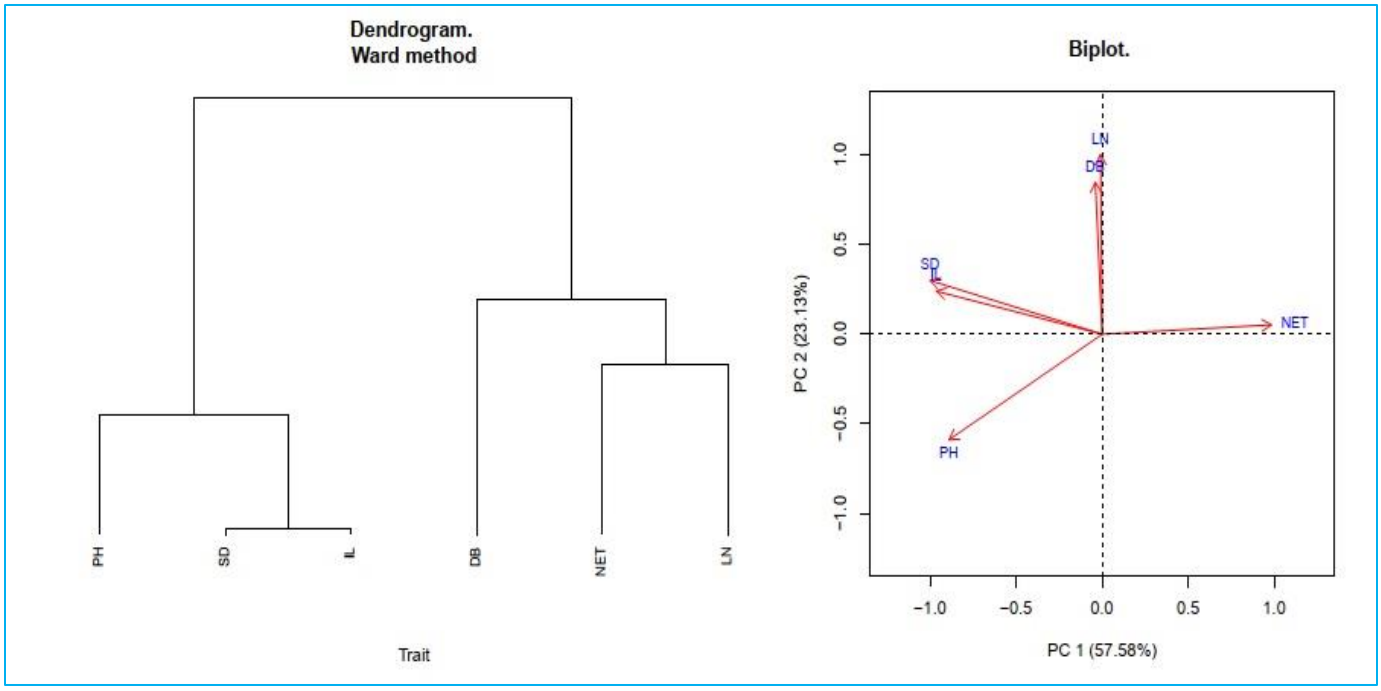

Figure 4.2. Cluster for distance matrix and bi-plot for principal component analysis for six traits across 324 sorghum accessions.

(NET = number of effective tillers, SD = stalk diameter, LN = number of leaf, IL = internode length, PH = plant height, DB = dry biomass).

4.2. Genetic Diversity and Population Structure Analysis of Sorghum Accessions using SSR Markers

4.2.1. Microsatellite markers based polymorphism

The analysis revealed that all 15 SSR loci were polymorphic (Figure. 4.3) and produced 108 alleles with an average of 7.2 alleles per locus among the 100 sorghum accessions (Table 4.8). From the total of 108 alleles, the frequency of 8 (7.41%) is below 0.05 (rare alleles), 100 (92.59%) alleles were between 0.05 and 0.5 (common alleles), and no abundant alleles were recorded with allelic frequency more than 0.5. The number of alleles per locus ranged from four (XTXP021) to twelve (XTXP265) (Table 4.8) with overall mean value of 7.2. Across the entire populations, the major allele frequencies ranged from 0.14 for locus XCUP50 to 0.34 for XTXP136 with an overall mean of 0.25 (Table 4.8). The highest Na, Ne, I, Nm and PIC were recorded for XTXP265. Microsatellite loci XCUP50 (h = 0.90) and XTXP021 (h = 0.73) resulted in the largest and the lowest gene diversity across the entire populations, respectively (Table 4.8).

The Locus XCUP265 showed the least genetic differentiation (Gst = 0.02) and the highest gene flow (Nm = 12.56) among the populations. The same locus also showed the highest observed (H_o = 0.96) and expected heterozygosity (H_e = 0.87) for each locus where as the locus XTXP021 showed the least gene diversity (expected heterozygosity, H_e = 0.83) (Table 4.8). The lowest mean value (0.06) for observed heterozygosity (H_o) was obtained in this experiment (Table 4.8).The informativeness of individual loci as measured by polymorphic information content ranged from 0.68 (XTXP021) to 0.89 (XCUP50) with overall mean of 0.80. All of the loci were highly

informative (PIC > 0.5). All the tested microsatellite loci were significant for deviation from Hardy Weinberg Equilibrium (P<0.01) (Table 4.8).

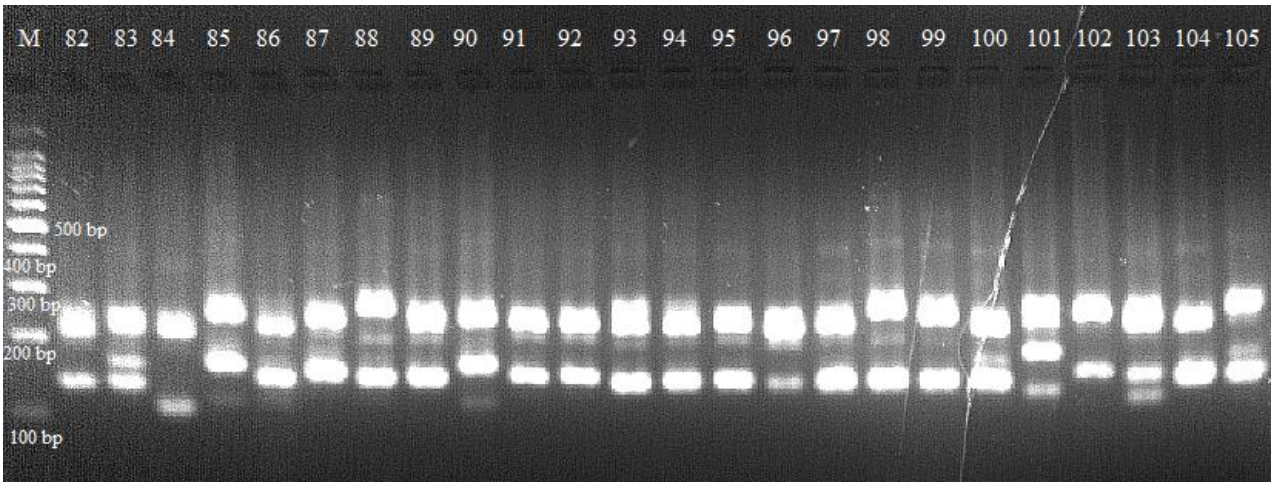

Figure 4.3. PCR products for Primer Xtxp265 with samples 82 - 105. M = marker (100 bp DNA ladder).

Table 4.8. Informativeness and other genetic diversity summary statistics for all 15-microsatellite loci across 100 accessions of sorghum representing five populations in Ethiopia.

Locus	MAF	Na	Ne	I	h	H_o	Fst	Nm	PIC	P value	P_{HWE}
SBAGBO02	0.22	7.00	6.43	1.90	0.84	0.00	0.06	4.11	0.83	<0.001	<0.001
XCUP74	0.21	7.00	6.54	1.91	0.85	0.00	0.09	2.70	0.83	<0.001	<0.001
XTXP136	0.34	5.00	4.33	1.54	0.77	0.00	0.08	2.95	0.73	<0.01	<0.001
XCUP50	0.14	10.00	6.54	1.28	0.90	0.00	0.06	4.18	0.89	<0.001	<0.001
XCUP67	0.33	5.00	4.29	1.53	0.77	0.00	0.05	5.02	0.73	<0.001	<0.001
XCUP14	0.25	5.00	4.69	1.58	0.79	0.00	0.06	4.07	0.75	<0.01	<0.001
XGAP084	0.18	10.00	8.14	2.17	0.88	0.00	0.05	5.26	0.87	<0.01	<0.001
XTXP010	0.23	6.00	5.77	1.78	0.83	0.00	0.04	5.73	0.80	<0.01	<0.001
XTXP012	0.19	9.00	7.99	2.14	0.88	0.00	0.06	3.90	0.86	<0.001	<0.001
XTXP021	0.33	4.00	3.74	1.35	0.73	0.00	0.02	10.67	0.68	<0.01	<0.001
XTXP141	0.29	6.00	5.29	1.73	0.81	0.00	0.05	4.63	0.79	<0.01	<0.001
XTXP145	0.32	8.00	5.32	1.85	0.81	0.00	0.05	4.63	0.79	<0.001	<0.001
XTXP265	0.18	12.00	8.87	2.33	0.89	0.96	0.02	12.56	0.88	<0.01	<0.01
XTXP273	0.25	6.00	5.48	1.75	0.82	0.00	0.03	8.77	0.79	<0.01	<0.001
XTXP320	0.23	8.00	6.38	1.93	0.84	0.00	0.04	6.20	0.82	<0.05	<0.001
Mean	0.25	7.2	6.19	1.85	0.83	0.06	0.05	4.81	0.80		

MAF = Major allele frequency, Na = Observed number of alleles, Ne = Effective number of alleles (Kimura and Crow, 1964), I = Shannon's Information statistic, h = Nei's gene diversity, H_o = Observed heterozygosity for each locus, h/H_e = Expected heterozygosity for each locus, Gst = Genetic differentiation statistics by locus, Nm = Estimate of the number of migrants (gene flow) from Fst where Nm = [(1/Fst) - 1]/4, Fst = Inbreeding coefficient within subpopulations relative to total, PIC = Polymorphic information content, *P*-value and P_{HWE} for deviation from Hardy Weinberg Equilibrium highly significant at ($p<0.001$), highly significant at ($p<0.01$) and significant at ($p<0.05$)].

4.2.2. Genetic variability within and among the populations

A summary of the different genetic diversity estimates over all the loci across populations is presented in Table 4.9. The overall genetic diversity estimates within the populations showed a mean number of effective alleles of 5.01 with a range of 4.25 to 5.71, a mean number of private alleles (number of alleles unique to a single population) of 0.03 with a range of 0.00 to 0.13, mean Shannon's information index of 1.66 ranging from 1.47 to 1.80 and gene diversity value of 0.78 ranging from 0.74 to 0.81. The population of Oromia showed the highest Na, Ne, NPA, I, and genetic diversity (Table 4.9).

The sorghum accessions from Amhara region ranked second in terms of Na, Ne and I and similar to Oromia population in NLCA and gene diversity index (Table 4.9). Except for the population of Oromia where NPA = 0.13, there was no private allele unique to a single population. NLCA was the highest in both Oromia and Amhara populations. The lowest Nei's gene diversity was observed in the populations of SNNP while collections from Oromia and Amhara regions showed the highest gene diversity. Sixty percent of the study populations showed a genetic diversity greater than the mean value of 0.78. The percentage of polymorphic loci per population showed 100% polymorphism across all populations (Amhara, Dire Dawa, Oromia, SNNP, and Tigray) (Table 4.9).

Table 4.9. Allelic patterns and diversity indices[a] across populations averaged over the 15 SSR loci.

Population[b]	Size	Na	Ne	NPA	NLCA	I	h
Amhara	31	6.93	5.45	0.00	0.20	1.76	0.81
Dire Dawa	14	5.73	4.62	0.00	0.00	1.58	0.76
Oromia	30	7.13	5.71	0.13	0.20	1.80	0.81
SNNP	7	4.80	4.25	0.00	0.00	1.47	0.74
Tigray	18	6.47	5.01	0.00	0.13	1.69	0.79
Mean	20	6.21	5.01	0.03	0.11	1.66	0.78

[a] Na = Observed number of alleles, Ne = Number of effective alleles, NPA = Number of Private Alleles (i.e., the number of alleles unique to a single population), I = Shannon's information statistic, h = Nei's genetic diversity and NLCA = Number of Locally Common Alleles found in populations. N.B. PPL = 100% across all populations, [b]SNNP = Southern Nations, Nationalities, and People's Region.

4.2.3. Genetic relationships between the populations

Table 4.12 presents pairwise population measures of Nei's genetic distance (below diagonal) and gene flow (Nm) (above diagonal). The pairwise Nei's standard genetic distance of each population from the other populations ranged from 0.13 to 0.39 (Table 4.10). The lowest genetic distance (0.13) was observed between the sorghum collections from Oromia and Amhara with the highest gene flow of 21.66, which was followed by the relationship between Amhara and Tigray and Oromia and Tigray populations that scored a distance of 0.16 and gene flow of 17.18 and 17.91, respectively. The highest measure of genetic distance (0.39) with relatively low gene flow (Nm = 5.27) was observed between the populations of Dire Dawa and SNNP (Table 4.10). The second highest genetic distance was observed between the populations of SNNP and Tigray and SNNP and Oromia (0.32) with gene flow of 7.92 and 8.30, respectively (Table 4.10).

Table 4.10. Pairwise Nei's genetic distance (below diagonal) and gene flow (Nm) (above diagonal) among the five sorghum populations from Ethiopia.

Population [a]	Amhara	Dire Dawa	Oromia	SNNP	Tigray
Amhara	--	5.67	**21.66**	8.10	17.48
Dire Dawa	0.30	--	7.10	**5.27**	7.63
Oromia	**0.13**	0.26	--	7.92	17.91
SNNP	0.31	**0.39**	0.32	--	8.30
Tigray	0.16	0.26	0.16	0.32	--

-- = not applicable. [a] SNNP = Southern Nations, Nationalities, and People's Region.

The pairwise coefficient of genetic differentiation between the populations ranged from 0.10 (between Dire Dawa and Oromia) to 0.91 (between SNNP and Tigray) (Table 4.11). The highest and statistically significant genetic differentiation (PhiPT = 0.91, $p < 0.05$) was observed between populations of Tigray and SSNP, implying the lowest gene flow between them. The second highest genetic differentiation with a gene flow rate of 0.86 was observed between the populations of Tigray and Oromia. The genetic differentiation between populations of Dire Dawa and SNNP was not statistically significant ($p = 0.05$) (Table 4.11).

4.2.4. Population genetic differentiation

The highest Population genetic differentiation was observed between Tigray and Southern Nations, Nationalities, and People's Region and the lowest Population genetic differentiation was between Dire Dawa and Oromia region (Table 4.11).

Table 4.11. Population genetic differentiation measured by PhiPT (above the diagonal) between the five sorghum populations with p-values (below the diagonal).

PhiPT/p value	Amhara	Dire Dawa	Oromia	SNNP	Tigray
Amhara	--	0.18	0.76	0.81	0.84
Dire Dawa	0.04	--	**0.10**	0.65	0.42
Oromia	**0.01**	0.03	--	0.81	0.86
SNNP	0.04	**0.05**	0.04	--	**0.91**
Tigray	0.02	0.03	0.02	0.04	--

-- = not applicable, PhiPT = Genetic differentiation, SNNP = Southern Nations, Nationalities Peoples

4.2.5. Analysis of molecular variance

Analysis of molecular variance (AMOVA) was computed without grouping for Amhara, Dire Dawa, Oromia, SNNP and Tigray and with grouping for North, Central, East and South Ethiopia of the populations according to their geographical locations and major agro-ecological adaptation zones (highland, intermediate and lowland). AMOVA based on Fst values revealed that 97.68% of the genetic variation occurred within populations. The variability among populations, among geographical regions and major agro-ecological adaptation zones only accounted for 2.32%, 1.94% and 0.72%, respectively, of the total genetic diversity. The overall genetic differentiation

coefficient among the populations was low (Fst = 0.023, $P < 0.0001$) with high gene flow (≥ 10.53) (Table 4.12).

Table 4.12. AMOVA showing the partitioning of genetic variation within and among populations and geographical regions of sorghum in Ethiopia.

Source of variation	DF	SS	MS	Estimate of variance components	Percent of variation	[a]Genetic differentiation	[b]P-value
Among populations	4	46.50	11.62	0.15	2.32		
Within Populations	195	1192.54	6.12	6.12	97.68	Fst : 0.023	Fst < 0.0001
Total	199	1239.04		6.26	100%		
Nm(diploid) [c] populations	10.53						
Among Geographical Regions	3	34.16	11.39	0.12	1.94		
Within Geographical Regions	196	1204.88	6.15	6.15	98.06	Fst: 0.019	Fst < 0.0001
Total	199	1239.04		6.27	100%		
Nm(diploid)[d] Geographical regions	12.64						
Among Altitudes	2	18.33	9.16	0.05	0.72		
Within Altitudes	197	1220.71	6.20	6.20	99.28	Fst: 0.0072	Fst < 0.0001
Total	199	1239.04		6.24	100%		
Nm(diploid) [e] Altitudes	34.47						

[a] Fst = the coefficient of genetic differentiation, calculated as = AP / (WP + AP), where AP = the variance among populations and WP = the variance within populations, [b] significance based on 1000 permutations, Gene flow (Nm) calculated as [(1/Fst) −1]/4; [c] among regions, [d]among geographical regions, [e] among major agro-ecological adaptation zones (altitudes) (lowland, intermediate, highland).

4.2.6. Cluster analysis, PCoA and population structure

Seventy-two percent (72%) of the individuals were assigned into C1, which was followed by C3. The lowest (four) number of accessions were assigned to cluster C3 (Figure. 4.4). None of the major clusters and/or sub-clusters is composed of exclusively accessions from a particular population or collection zone, confirming the existence of considerable intermixing of the accessions. The dendrogram generated using the UPGMA grouped the five populations into three major clusters (I, II and III), where the second cluster was composed of only a single population (Oromia), and both clusters I and III were composed of two populations each, Amhara and Tigray and Dire Dawa and SNNP, respectively (Figure. 4.5).

The scatter plot obtained from PCoA revealed almost a uniform distribution of the sorghum accessions of all populations around the center of the two-dimensional coordinate plane with poor population clustering. The first three principal coordinate axes explained 25.19% of the total variations (Figure. 4.6) with each of the dimensions (1, 2 and 3) contributing only 4.32%, 8.52%, and 12.35%, respectively.

The output of the Structure harvester revealed that the delta K (ΔK) values reached a sharp peak at K = 3 (Figure. 4.7a, Table 4.13), confirming that the sorghum accessions can be clustered into three subpopulations. The Clumpak result (bar plot) detected a greater degree of genetic admixture between the three subpopulations, and hence, there was no clear geographic origin-based structuring in the sorghum populations of Ethiopia (Figure. 4.7b). The result also revealed that each individual sorghum sample shared alleles from the three sub-populations.

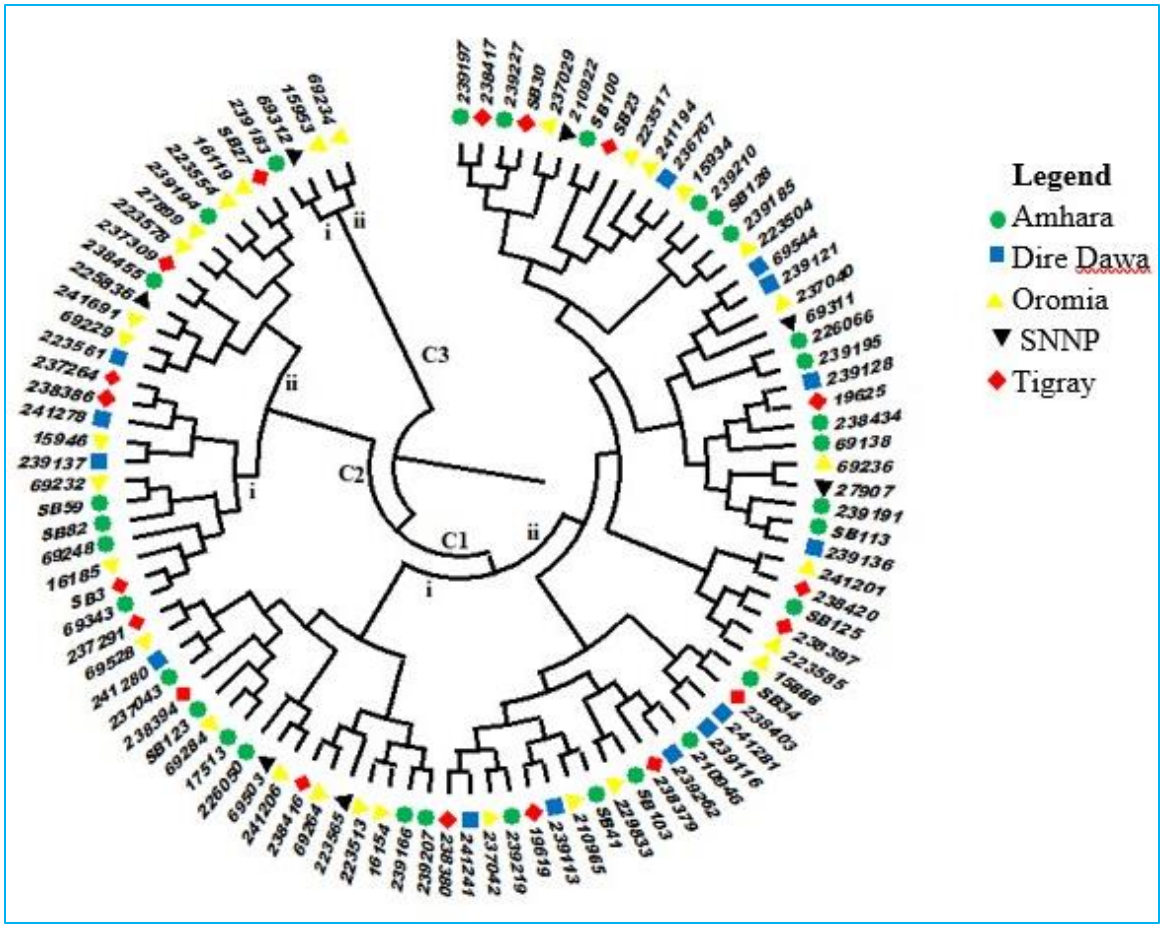

Figure 4.4. Unweighted Neighbor-Joining tree showing relatedness among the 100 sorghum accessions.

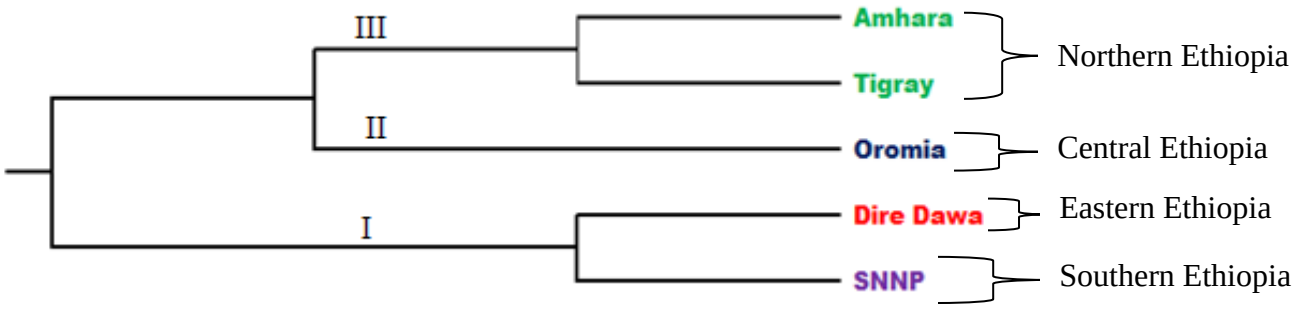

Figure 4.5. UPGMA dendrogram showing genetic relationships among the five populations of 100 sorghum accessions sampled from Ethiopia based on Nei's unbiased genetic distances.

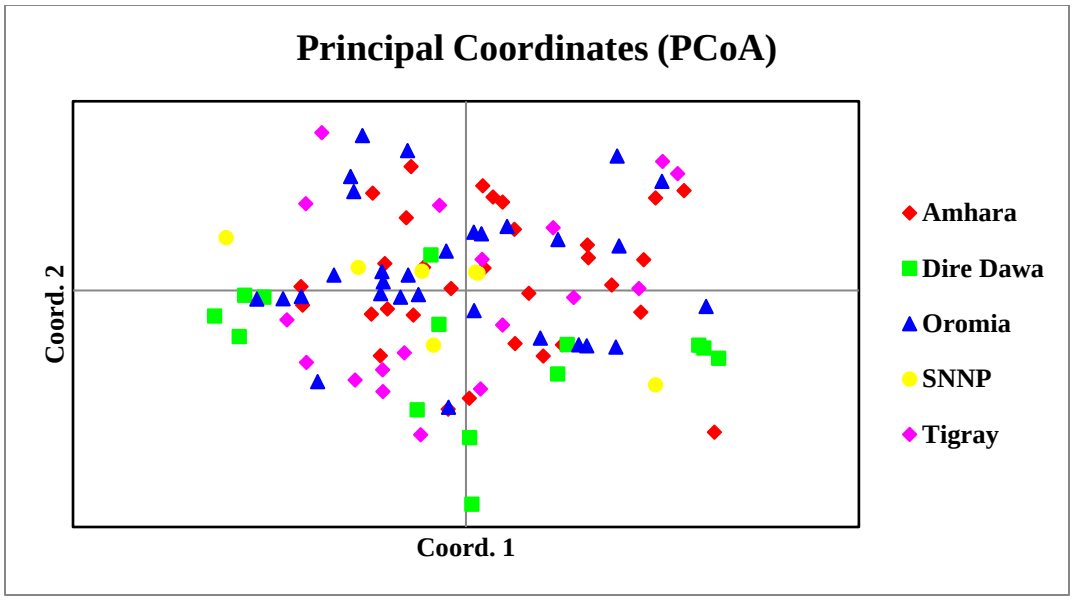

Figure 4.6. PCoA of the 100 sorghum accessions as revealed by 15-microsatellite loci. Samples coded with the same symbol and color belongs to the same population. SNNP = Southern Nations, Nationalities, and People's Region.

Table 4.13. Identification of the true number of clusters (K) based on Delta k (ΔK) from structure harvester.

K	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	-5472.79	0.396398	—	—	—
2	-5302.2	4.385382	170.585	2.435	0.555254
3	-5129.18	2.083671	173.02	42.395	20.3463
4	-4998.56	3.046737	130.625	22.825	7.491622
5	-4890.76	9.766941	107.8	20.49	2.097893
6	-4762.47	17.18862	128.29	35.755	2.080155
7	-4669.93	7.535537	92.535	0.99	0.131378
8	-4578.39	16.18028	91.545	5.41	0.334358
9	-4481.43	19.02622	96.955	1.315	0.069115
10	-4383.16	20.52024	98.27	—	—

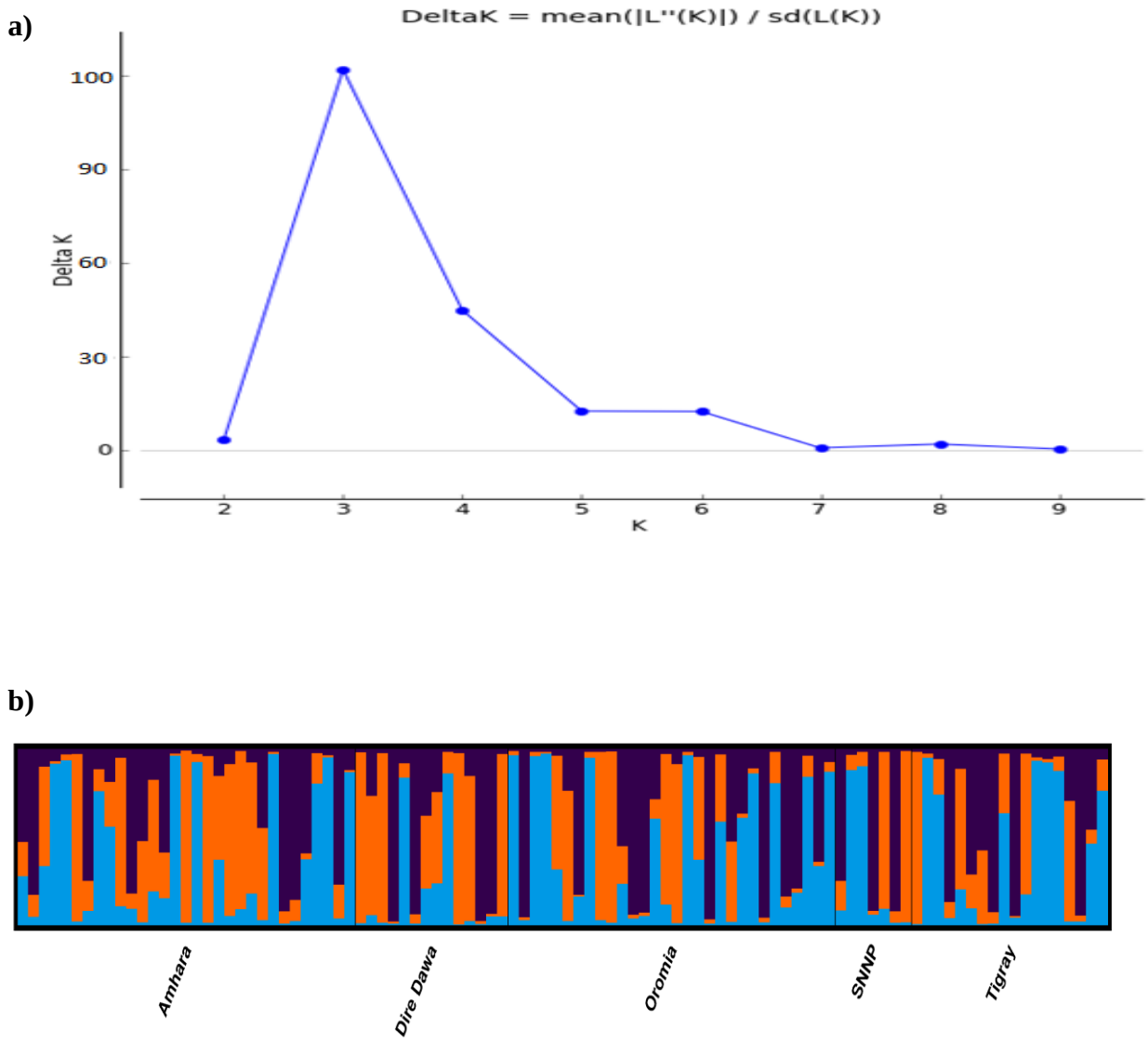

Figure 4.7. Population structure of 100 sorghum accessions representing five populations in Ethiopia.

(a) Best delta K value estimated using the method of (Evanno *et al.*, 2005); and (b) Estimated population structure for K = 3 according to geographical locations. The different (blue, black and orange) colors represent genetic groups or sub-populations designated by structure harvester: the x-axis represents individual samples and y-axis represents the proportion of ancestry to each cluster. SNNP = Southern Nations, Nationalities, and People's Region.

5. DISCUSSION

5.1. Morphological Diversity based on Quantitative Traits

5.1.1. Mean performance and range of sorghum accessions

Agro-morphological characterization is one of the most important steps in maximizing the use of existing genetic resources for genetic improvement and classification. In the present study, the tested sorghum accessions showed highly significant differences for all morphological quantitative traits, implying that the accessions had a large variation useful for sorghum breeding. Analysis at individual locations (Melkassa, Miesso and Mehoni) indicated the presence of considerable genotypic mean performance among the genotypes for all the studied agronomic traits (Tesfaye Disasa, 2016).

The mean performance of number of effective tillers, stalk diameter, leaf numbers and internode length were the highest at Melkassa. Plant height and dry biomass showed higher performances at Miesso. This could be probably due to the favorable environmental factors such as medium and light textured soil types (Abera Deresa *et al.*, 2005), as sorghum crop grows successfully in such circumstances and the location could be better area for the crop to perform better.

The genotypes mean performance for number of effective tillers ranged from 3.4 (for the accession 223517 in Oromia) to 4.54 (for 237038 in Oromia) with an overall mean value of 3.95 (Appendix 1). Stalk diameter ranged from 16.13 mm for genotype 19625 in Tigray to 22.60 mm for 241239 in Oromia with overall mean value of 19.54 mm. The average leaf numbers obtained

in the present study was substantially higher (18.04) than the number of leaves reported in previous study (15.64) by Sihag *et al.* (2019).

The average plant height (272.12 cm) observed in the present study was significantly higher than the mean plant height (174-196.6 cm) reported in earlier studies by Chalachew Endalemaw *et al.* (2017), Solomon Assefa *et al.* (2018) and Tafere Mulualem *et al.* (2018). This increased plant height in the present study lines with the farmers' preference for food and animals feed. This is a promising opportunity for selection in sorghum improvement programs (Solomon Assefa *et al.*, 2018).

The dry biomass yield in this study ranged from 5.47 tonne ha^{-1} (210957 in Oromia) to 23.67 tonne ha^{-1} (69232 in Oromia) with overall mean of 11.85 tonne ha^{-1}. The average biomass yield observed in the present study was lower than the value reported by Vasilakoglou *et al.* (2011) (13.70 tonne ha^{-1}), and significantly higher than the value reported by Tang *et al.* (2018) who described an average of 6.10 and 9.20 tonne ha^{-1} respectively. Significant variations in environmental factors such as climate (precipitation, temperature, and evaporation), as well as soil type and fertility, could explain the substantial differences in biomass yield at different sites in this study. Tang *et al.* (2015) found that precipitation and soil organic matter were important environmental factors influencing sorghum biomass yield.

Generally, the tested sorghum genotypes exhibited significant variability for the studied traits, indicating their scope for improvement through selection. This study revealed that sorghum collections from Oromia region showed a higher performance for all quantitative traits except

plant height indicating that the region is a potential hotspot for sorghum genetic variability studies and identification of candidate genotype carrying desirable traits to be used in breeding programs. Dire Dawa is potential area for sorghum to bear better plant height, having a direct relation to both biomass and grain yield (Solomon Assefa *et al.*, 2018), which is an indication that better sorghum genotypes can be exploited from the area.

5.1.2. Estimation of analysis of variance

ANOVA indicated significant differences among genotypes for all the six traits studied, indicating the presence of wide range of variability. Analysis of variance showed that the mean squares due to genotype were highly significant ($p < 0.01$) for stalk diameter and very highly significant ($p < 0.001$) for the other traits, indicating promising possibility for genetic improvement through exploiting the genetic variations. ANOVA also showed very highly significant effect of location on studied agronomic traits, indicating the relevance of conducting multi-location trials to identify broad adaptive and stable genotypes. Similar variability in location effect in sorghum genotype performance were reported in various sorghum researchers (Tekle Yoseph and Zemach Sorsa, 2014; Tesfaye Disasa, 2016; Adane Gebreyohannes *et al.*, 2018; Tafere Mulualem *et al.*, 2018; Gedifew Gebrie and Tsige Genet, 2019).

5.1.3. Estimation of genetic parameters

5.1.3.1. Genotypic and phenotypic coefficient of variation

The Genotypic Coefficient of Variance (GCV) alongside with heritability estimations give solid estimates of the genetic advance to be expected through phenotypic choice (Burton, 1952). In the present study, the GCV of the quantitative traits ranged from 2.66% for stalk diameter to 22.58% for plant height, while the PCV ranged from 5.39% for leaf numbers to 26.95% for plant height. The high (>20%) GCV observed for plant height, indicates the large genetic control of the trait, and thus promising possibility for improvement through breeding.

Those agronomic traits with lower GCV but higher PCV indicates the large environmental control of the traits, and hence less possibility to improve through selection. The phenotypic variation was relatively greater than the genotypic variation for all traits indicating the immense inherent contribution of environmental variance for the expression of the phenotypic variance of the traits that remains unaltered by environmental conditions among the genotypes, which in turn is more useful for exploitation in selection and hybridization programs. On the other hand, the extent of the difference between GCV and PCV was relatively low for NET, LN and IL indicating the influence of the environment on these traits was very low.

The ratio of PCV and GCV was maximum for plant height indicating that these traits are more influenced by the environment. Low GCV and moderate PCV were observed for internode length. Low estimates of GCV and PCV were observed for NET, SD and LN indicating that these traits contribute low magnitude of heritable (genetic) factor to the next generation, which shows no need for investment to improve these traits aiming for sorghum improvement. Similar

to the present study, Ranjith *et al.* (2017) reported low results for the number of effective tillers, which supports the present findings.

5.1.3.2. Trait heritability

In this study, the agronomic traits showed considerable variation in broad sense heritability ranging from low (H^2 = 7.20%) for stalk diameter to high heritability (H^2 = 70.22%) for plant height (Johnson *et al.*, 1955). Like the previous studies by Seetharam and Ganesamurthy (2013), Ranjith *et al.* (2017) and Gebremedhn Gebregergs and Firew Mekbib (2020) the present study reveald high heritablity for plant height (H^2 = 70.22%), indicating higher and positive response for selection to improve the trait.

Agronomic traits that showed high heritability coupled with higher GCV and GAM are effective for selection for ensuring biomass improvement in sorghum as they are under additive genetic control. Some agronomic traits including NET (37.78%) and DB (45.46%) showed moderate heritability in broad sense while SD (7.20%), LN (29.20%) and IL (16.07%) showed low broad sense heritability, indicating less chance to improve such traits through breeding likely due to the large non-additive genetic effect to the phenotypic expression of traits.

5.1.3.3. Genetic advance as percent of mean

When combined with heritability estimates, the estimate of genetic advance is more valuable as a selection tool (Johnson *et al.,* 1955). Genetic advance estimates aid in determining the type of gene activity involved in the expression of various polygenic features. Additive gene action is indicated by high genetic advance, whereas non-additive gene activity is indicated by poor genetic advance (Singh and Narayanan, 1993). As a result, if high genetic progress is present, heritability estimates will be reliable.

In the present study, the genetic advance as percent of the mean ranged from 1.47% for stalk diameter to 38.99% for plant height. Plant height (38.99) and dry biomass (18.68) showed higher GAM (Johnson *et al.,* 1955), indicating that they are under higher additive genetic control where the influence of non-additive genetic effect was minimal. The result indicated that planning to improve sorghum through selection of the above traits is suitable breeding procedure because of its higher estimate of GAM. The present result's higher GAM for plant height and dry biomass was in line with the report of Ranjith *et al.* (2017), Gebremedhn Gebregergs and Firew Mekbib (2020). Generally, the higher broad sense heritability, higher estimate of GCV and GAM for plant height confirmed that the trait is under was additive genetic control, indicating the possibility to improve through selection. A high heritability coupled with high GAM in sorghum genetic variability was reported by Gebremedhn Gebregergs Firew Mekbib (2020) and Sabiel *et al.* (2016) in agreement with the present study.

5.1.4. Correlation analysis

5.1.4.1. Phenotypic and genotypic correlation of dry biomass with other traits

Genotypic and phenotypic correlation analysis indicated that dry biomass showed positive and highly significant ($p < 0.001$) association with all the studied quantitative traits, indicating the possibility to improve dry biomass performance through simultaneous selection for agronomic traits. Positive and significant correlations for dry biomass at phenotypic and genetic levels were reported by Chalachew Endalemaw *et al.* (2017), which could be considered as selection criteria to improve sorghum biomass.

5.1.4.2. Genotypic and phenotypic correlation among other traits

Most of the studied traits had highly significant ($p < 0.001$) phenotypic and genotypic correlation with other traits. However, number of effective tillers showed a non-significant genotypic correlation with stalk diameter and a non-significant phenotypic correlation with stalk diameter and plant height. The genotypic correlation coefficient provides a measure of the genetic relationship between traits that could be used as a selection criterion (Can and Yoshida, 1999). In the present study, most traits showed positive and highly significant ($p<0.001$) phenotypic (rp) and genotypic (rg) correlation such as NET with LN (rp = 0.53, rg = 0.27), DB (rp = 0.53, rg = 0.21) and PH (rp = 0.53), SD with LN (rp = 0.57, rg = 0.14), IL (rp = 0.98, rg = 0.24), PH (rp = 0.68, rg = 0.26) and DB (rp = 0.58, rg = 0.19). This indicates the selection of one trait is directly related to the other. Similar result reported that biomass yield had significant positive correlation with stalk diameter (rg = 0.48) and plant height (rg = 0.39) (Chalachew Endalemaw *et al.*, 2017). Blum *et al.* (1989) also reported positive significant associations between biomass yield and other variables.

5.1.5. Cluster and principal component analysis

Since clustering is based on trait similarities, significant levels of recombination and heterosis can be generated by combining superior individuals from different clusters (Derese Solomon *et al.*, 2018). Based on hierarchical cluster analysis, 324 accessions were classified into six clusters. This study agrees with the report conducted by More *et al.* (2018) who grouped 26 sorghum accessions into six different clusters using hierarchical cluster analysis. Similar to the present result, Kassahun Tesfaye (2017) reported 13 cluster groups for 119 sorghum accessions. Sihag *et al.* (2019) also reported seven gropus for 96 accessions based on 13 traits. Though the number of accessions and type of characters used in these experiments are different, the clustering pattern indicated the existence of a significant amount of variability among the sorghum genotypes.

Intra and inter cluster distance is measured to determine the degree of divergence between the clusters. The maximum intra cluster distance was recorded for cluster 6 (2.46) followed by cluster 4 (2.01), cluster 3 (1.81), cluster 5 (1.66), and cluster 2 (1.65), while minimum distance was observed for cluster 1 (1.55) indicating that the traits were less divergent. The maximum inter cluster distance was perceived between cluster 3 and cluster 4 (70.25) followed by cluster 2 and cluster 4 (60.23). So the genotypes of cluster 3 and cluster 4 can be used as parents in yield improvement programs. The genotypes within the same cluster considered to have the similar phenotypic characters and are more related as compared to the genotypes of the other clusters whereas the genotypes between the clusters are more diverse. The evaluated accessions showing diverse clusters in this study could be valuable source of parents for hybridization programs

(Solomon Assefa *et al.*, 2018) to develop high yielding sorghum varieties. The finding of the present study was in line with the reports of Sihag *et al.* (2019).

The principal component analysis result showed that the first two principal components have an Eigenvalue greater than one suggesting that traits such as number of effective tillers, stalk diameter, number of leaves, internode length, plant height and dry biomass are useful traits and had higher contribution to the total variation of the accessions.

5.2. Genetic Diversity and Population Structure of Sorghum Accessions as revealed by SSR markers

5.2.1. Polymorphism level of the microsatellite markers

In this study, fifteen polymorphic markers that are distributed across the sorghum genome were used to determine the extent of diversity among hundred sorghum accessions collected from different parts of Ethiopia. The observed simple-sequence repeat (SSR) fragment sizes were within the range of the sizes in the previously published reports in sorghum (Ng'uni *et al.*, 2012; Tesfaye Disasa *et al.*, 2017; Alemu Tirfessa, 2020).

Each locus had showed four to twelve alleles per locus with overall number of 7.2 per locus. The average number of alleles observed in the present study is significantly higher than the level reported by Alemu Tirfessa *et al.* (2020), who described an average of 6.7 alleles per locus among 200 sorghum accessions. However , the number of alleles reported in our study was lower than the mean of alleles per locus (19) obtained from genotyping of large number of sorghum accessions (3367) using 41 SSR markers collected around the world (Billot *et al.*, 2013). The observed variation is likely due to the diverse nature of accessions used for genotyping and the polymorphism level of the selected SSR markers.

All the SSR markers in this study were highly polymorphic, confirming their usefulness for genetic analysis. The PIC ranged from 0.68 (XTXP021) to 0.89 (XCUP50), which was comparable to a similar recent study of sorghum collections in Ethiopia (Alemu Tirfessa *et al.*, 2020) in which a PIC range of 0.06 to 0.81 was observed using 39 SSRs across 200 accessions, and 0.05 - 0.78 (Danquah *et al.*, 2019). The average PIC value of the current study (0.80) was

higher than some of the previously reported values on sorghum (Mofokeng *et al.*, 2014; Danquah *et al.*, 2019; Alemu Tirfessa, 2020). This could be due to the polymorphic nature of the selected SSR markers used for genotyping and the type of the accessions studied. This increased and highly polymorphic PIC value is an indication for the high-resolution potential of the markers to detect differences among the sorghum genotypes (Salih *et al.*, 2016); implying that the primer are useful for molecular characterization and marker assisted breeding.

In the current study, the gene diversity (h) ranged from 0.73 to 0.90 with an overall mean of 0.83. This average gene diversity was greater than the reports of Billot *et al.* (2013). The increased genetic diversity implies the presence of significant genetic diversity among the studied sorghum accessions (Mofokeng *et al.*, 2014) that provides breeders options to select genotypes for crossing and conservation purposes.

5.2.2. Population genetic diversity

The presence of high genetic diversity of sorghum populations in Ethiopian was evidenced by the statistical analysis of population genetic variability across all loci. The five populations had a high average number of observed alleles (6.21), number of effective alleles (5.01), Shannon's information index (1.66), Nei's gene diversity (0.71) and percentage of polymorphic loci (100%). The gene diversity observed in the current studied populations (0.71) is higher than the diversity value (0.57) reported by Tesfamichael Abraha *et al.* (2014) in sorghum. The high levels of gene diversity of SSR markers observed in this study was probably due to the presence of genetic diversity in these sorghum accessions that represented different geographic regions.

In the present study, all the five populations showed a highest percentage of polymorphic loci (100%) and the number of private alleles were 0.13 (in Oromia). Similarly, Salih *et al.* (2016) reported 1.24 private alleles in East and Central Africa including Ethiopia. The presence of these private alleles in the population may indicate the gene flow from wild to cultivated sorghum accessions as Ethiopia is believed as center of diversity for sorghum; having abundant wild accessions that grow in close proximity with cultivated accessions resulting in regular hybridization between the different gene pools. The private alleles could have unique agricultural importance.

The molecular diversity analysis in the present study populations showed a high genetic diversity in Oromia collections with the highest number of effective alleles (5.71), Shannon's Information Index (1.80), Percentage of Polymorphic Loci (0.20) and gene diversity (0.81). This may resulted due to the presence of greater genetic diversity of sorghum genotypes in the region, suggesting the region is hot spot for genetic analysis of sorghum in Ethiopia for breeding and conservation purposes.

5.2.3. Population genetic structure

AOMVA showed high genetic diversity within populations than among populations It was observed that the within genetic diversity in geographical regions and altitudes in the major agro-ecologies (Amsalu Ayana, 2001) is higher than the diversity among them. AMOVA showed that studied sorghum accessions showed low genetic differentiation where the among populations variation accounted for only 2.32% of the total genetic variation leaving the most for within population variations. Similar lower proportions of among population genetic variations of

0.38%, 2.75%, 8% and 9.56% were reported for *S. bicolor* populations in Kenya (Muui *et al.* , 2016), South Africa (Motlhaodi *et al.*, 2017), Ethiopia (Alemu Tirfessa, 2020) and China (Zhu *et al.*, 2020) in the same order. A higher proportion of among population genetic variations of 41.2% was reported in Ethiopia (Asfaw Adugna *et al.*, 2013). This could be due to the diverse nature of sorghum accessions since Ethiopia is within the center of origin for sorghum (Doggett, 1988).

For the same crop, variable rates of genetic differentiation were reported from different parts of the world. A statistically significant low to moderate genetic differentiation (0.03 to 0.08) was reported for *S. bicolor* populations in Ethiopia (Adugna *et al.*, 2013), 0.08 to 0.082 in China (Zhu *et al.*, 2020) and Zambia (Ng'uni *et al.*, 2011), respectively. In the present study, the high within-population genetic variability (97.68%) could be attributed to sexual recombination and spontaneous mutation (Berraies *et al.*, 2013).

Genetic diversity is influenced by gene flow, which encompasses several mechanisms of gene exchange among populations (Slatkin, 1987). In this study, seed exchange practices between communities and marketing (Dalvand *et al.*, 2018) as well as socio-economic conditions have facilitated a high gene flow (Nm = 10.53) among populations from different seed collection regions leading to reduced among-populations genetic differentiation (Fst: 0.023) (Dalvand *et al.*, 2018). The same supportive result was observed for geographical regions and major agro-ecologies (Nm = 12.64, Fst: 0.019 and Nm = 34.47, Fst: 0.0072) with significant ($p < 0.0001$) magnitude of genetic variation, respectively. In line to this study, Beyene Seboka and van

Hintum (2006) reported that Ethiopia's significant influx of sorghum genotypes to farmer's seed systems are characterized by widespread germplasm sharing and gene flow dynamics.

The existence of greater genetic variation within populations than among populations was indicated by PCoA where the individuals of different groups failed to form distinct clusters (Pandian *et al.*, 2020), rather mixed up along the axis. This could be due to the presence high genetic materials exchange rate among the various sorghum-growing regions of the country.

The 100 accessions collected from the five different regions were classified into three distinct sub-populations (Delta K = 3) using the structure software. The sub-populations were complementary to the Unweighted Neighbor-Joining tree analysis showing relatedness among the 100 sorghum accessions. UPGMA dendrogram also showed presence of genetic relationships among the five sorghum populations and clustered them into three clusters. The present study was in line with the report of Pandian *et al.* (2020) in India and Weerasooriya *et al.* (2016); Tesfaye Disasa *et al.* (2017) in Ethiopia. The structure analysis further substantiated the presence of three sub-populations with significant admixture in the study populations.

6. CONCLUSION AND RECOMMENDATIONS

6.1. Conclusion

The present study was conducted aiming at the overall objective of exploring the genetic diversity in sorghum landraces using quantitative morphological traits and microsatellite markers. Wide variability was detected among the accessions collected from the major growing regions of Ethiopia for morphological characters. The significant mean squares due to genotypes indicated the existence of variation among the genotypes, which is useful to exploit for the improvement of respective traits. Collections from Oromia region showed better performance and variability for almost all traits implying that the region has potential in morpho-agronomic traits improvement through selection and conservation of the sorghum genetic resource.

The significant genotype by environment interaction for pooled data for biomass related agronomic traits such as leaf number, internodes length, plant height, dry biomass and number of effective tillers, indicate the significance of multi-locations germplasm evaluation to identify high biomass yielding and broad adaptive accessions. Heritability for the six quantitative traits ranged from 7.20% (stalk diameter) to 70.22% (plant height). The high heritability coupled with a higher estimate of genetic advances as percent mean and genetic coefficient of variation for the biomass related trait: the plant height indicates that it is largely under additive genetic control, and thus, it could be improved through breeding to fit Ethiopian farmers demand for animals feed and also other energy demands. The study revealed that some agronomic traits like dry biomass showed significant positive correlation with number of effective tillers, stalk diameter, leaf number, internodes length and plant, indicating that dry biomass could be simultaneously improved through indirect selection of any of these traits. Quantitative traits based grouping

clustering grouped the accessions into six clusters, which does not correlate with their geographical areas of sampling. Microsatellite marker based profiling also revealed the presence of considerable genetic variation in sorghum accessions in Ethiopia.

The high average number of alleles per locus (7.2) and PIC value of 0.80 in the present study confirmed the presence of high genetic diversity among sorghum genotypes, indicating the promising possibility to improve desired agronomic traits through breeding. The observed high (97.68%) within population genetic variation in the studied sorghum accessions could be due to sexual recombination. The sorghum populations showed low genetic differentiations (F_{ST}= 0.023), likely due to the presence of high gene flow as a result of seed exchange, marketing and socio-economic networks among the sorghum collection regions. SNNP showed the highest genetic distance with sorghum collections from Dire Dawa, implying that these populations could be appropriate for selection of diversified sorghum genotypes for further improvement through breeding.

The highest gene diversity was detected in the sorghum populations of Amhara and Oromia regions (h = 0.81), indicating that these regions are suitable for genetic diversity study, important genes in sorghum improvement programs and conservation purposes. Therefore, the present study which integrated morphological and molecular marker technology had generated valuable information for sorghum improvement through breeding, wise utilization and conservation of the available genetic resource.

6.2. Recommendations

The present study covered limited number of sorghum accessions for both quantitative traits and microsatellite marker sections collected from major growing regions of the crop in Ethiopia. However, future studies should include more accessions in order to cover other accessions of sorghum landraces. Though better genotypes for effective number of effective tillers, stalk diameter, leaf numbers, internode length, plant height and biomass yield were identified from the quantitative traits point of view, studies in the future should focus on other agro morphologically important traits to indicate a green light for sorghum improvement strategies.

Since the present piece of work revealed the abundance of private alleles and the presence of high gene diversity in Oromia region, it is good to focus on germplasm sources collected from the region for sorghum improvement programs and conservation plans in the future.

7. REFERENCES

Abera Deresa, Tesfaye Tesso and Asfaw Adugna. (2005). Sorghum Production and Research Experiences. *Afr. J. Biotechnol.* **8** (5): 18-29.

Adane Gebreyohannes, Taye Tadesse, Amare Seyoum, Habte Nida, Amare Nega, Tsegau Senbetay and Chalachew Endalemaw. (2018). Genetic variability in agronomic traits and associations in sorghum [(*Sorghum bicolor* (L.) Moench)] genotypes at intermediate agro-ecology sorghum growing areas of Ethiopia. *Afr. J. Agric. Res.* **13** (49): 2780-2787.

Asfaw Adugna, Snow, A. A., Sweeney, P. M., Endashaw Bekele and Mutegi, E. (2013). Population genetic structure of in situ wild *Sorghum bicolor* in its Ethiopian center of origin based on SSR markers. Genet. Resour. *Crop Evol.* **60** (4): 1313-1328.

Aldrich, P. and Doebley, J. (1992). Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild Sorghum bicolor. Theor. Appl. Genet. **85** (2-3): 293-302.

Alemu Tirfessa, Tesfaye Tesso, Asfaw Adugna, Mohammed, H. and Kiambi, D. (2020). Genetic diversity among Ethiopian sorghum [*Sorghum bicolor* (L.) Moench] gene bank accessions as revealed by SSR markers. *Afr. J. Biotechnol.* **19** (2): 84-91.

Alemu Tirfessa, Asfaw Adugna, Hussein Mohammed, Kiambi, Dan. (2020). Genetic diversity among Ethiopian sorghum [*Sorghum bicolor* (L.) *Moench*] gene bank accessions as revealed by SSR markers. *Afr. J. Biotechnol.* **19** (2): 84-91.

Alemu Tirfessa. (2018). Identifying sorghum [*Sorghum bicolor* (L.) *Moench*] plant types adapted to moisture stress areas of Ethiopia. The University of Queensland,

Ali, M., Rajewski, J., Baenziger, P., Gill, K., Eskridge, K. and Dweikat, I. (2008). Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. *Mol. Plant Breed.* **21** (4): 497-509.

Ali, M. A., Jabran, K., Awan, S., Abbas, A., Zulkiffal, M., Acet, T., Farooq, J. and Rehman, A. (2011). Morpho-physiological diversity and its implications for improving drought tolerance in grain sorghum at different growth stages. *Aust. J. Crop Sci.* **5** (3): 311.

Allard, R.W. (1960) Principles of Plant Breeding. John Willey and Sons Inc., New York.

Allard, R. W. (1999). *Principles of plant breeding*: John Wiley & Sons. **71**.

Amaducci, S., Monti, A. and Venturi, G. (2004). Non-structural carbohydrates and fibre components in sweet and fibre sorghum as affected by low and normal input techniques. *Ind. Crops Prod.* **20** (1): 111-118.

Ameen, A., Yang, X., Chen, F., Tang, C., Du, F., Fahad, S. and Xie, G. H. (2016). Biomass Yield and Nutrient Uptake of Energy Sorghum in Response to Nitrogen Fertilizer Rate on Marginal Land in a Semi-Arid Region. *Bioenergy Res.* **10** (2): 363-376.

Amsalu Ayana. (2001). Genetic diversity in sorghum (Sorghum bicolor (L.) Moench) germplasm from Ethiopia and Eritrea. Addis Ababa University,

Amsalu Ayana, Endashaw Bekele and Bryngelsson Tomas. (2000). Genetic variation in wild sorghum (Sorghum bicolor ssp. verticilliflorum (L.) Moench) germplasm from Ethiopia assessed by random amplified polymorphic DNA (RAPD). *Hereditas*. **132** (3): 249-254.

Amsalu Ayana and Endashaw Bekele. (1998). Geographical patterns of morphological variation in sorghum (Sorghum bicolor (L.) Moench) germplasm from Ethiopia and Eritrea: qualitative characters. *Hereditas*. **129** (3): 195-205.

Antonopoulou, G., Gavala, H. N., Skiadas, I. V., Angelopoulos, K. and Lyberatos, G. (2008). Biofuels generation from sweet sorghum: fermentative hydrogen production and anaerobic digestion of the remaining biomass. *Bioresour. Technol.* **99** (1): 110-119.

Arya, L., Verma, M. and Lakhanpaul, S. (2014). Diagnostic set of microsatellite markers for hybrid purity testing and molecular identification of hybrids and parental lines in sorghum. *J. Plant Sci Res*. **1** (1): 103.

Asfaw Adugna. (2014). Analysis of in situ diversity and population structure in Ethiopian cultivated Sorghum bicolor (L.) landraces using phenotypic traits and SSR markers. *SpringerPlus*. **3** (1): 1-14.

Asfaw Adugna, Snow, A. A., Sweeney, P. M., Endashaw Bekele and Mutegi, E. (2013). Population genetic structure of in situ wild Sorghum bicolor in its Ethiopian center of origin based on SSR markers. Genet. Resour. *Crop Evol.* **60** (4): 1313-1328.

Assar, A. A., Uptmoor, R., Abdelmula, A., Salih, M., Ordon, F. and Friedt, W. (2005). Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). *Crop Sci.* **45** (4): 1636-1644.

Barbanti, L., Grandi, S., Vecchi, A. and Venturi, G. (2006). Sweet and fibre sorghum (Sorghum bicolor (L.) Moench), energy crops in the frame of environmental protection from excessive nitrogen loads. *Eur. J. Agron.* **25** (1): 30-39.

Berndes, G., Hoogwijk, M. and Van den Broek, R. (2003). The contribution of biomass in the future global energy supply: a review of 17 studies. *Biomass Bioenerg.* **25** (1): 1-28.

Berraies, S., Gharbi, M. S., Belzile, F., Yahyaoui, A., Hajlaoui, M. R., Trifi, M., Jean, M. and Rezgui, S. (2013). High genetic diversity of Mycospaherella graminicola (Zymoseptoria tritici) from a single wheat field in Tunisia as revealed by SSR markers. *Afr. J. Biotechnol.* **12** (12).

Billot, C., Ramu, P., Bouchet, S., Chantereau, J., Deu, M., Gardes, L., Noyer, J.-L., Rami, J.-F., Rivallan, R. and Li, Y. (2013). Massive sorghum collection genotyped with SSR markers to enhance use of global genetic resources. *PloSone.* **8** (4): e59714.

Blum, A., Golan, G., Mayer, J., Sinmena, B., Shpiler, L. and Burra, J. (1989). The drought response of landraces of wheat from the northern Negev Desert in Israel. *Euphytica.* **43** (1): 87-96.

Bohn, M., Utz, H. F. and Melchinger, A. E. (1999). Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. *Crop sci.* **39** (1): 228-237.

Burton, G. W. (1952). Quantitative inheritance in grasses. Pro. VI Int. Grassl. Cong. *1952*: 277-283.

Burton, G. W. and Devane, d. E. (1953). Estimating heritability in tall fescue (Festuca arundinacea) from replicated clonal material. *J. Agron.* **45** (10): 478-481.

Can, N. D. and Yoshida, T. (1999). Genotypic and phenotypic variances and covariances in early maturing grain sorghum in a double cropping. *Plant Prod. Sci.* **2** (1): 67-70.

Caniato, F., Guimaraes, C., Schaffert, R., Alves, V., Kochian, L., Borém, A., Klein, P. and Magalhaes, J. (2007). Genetic diversity for aluminum tolerance in sorghum. *Theor. Appl. Genet.* **114** (5): 863-876.

CGIAR. (2015). The Biology of Sorghum bicolor (L.) Moenchsubsp. bicolor (sorghum). Australian Government Department of Health, Office of Gene Technology Regulator. OGTR, Australia.

Chalachew Endalemaw, Asfaw Adugna and Hussein Mohammed. (2017). Correlation and path coefficient analysis of agronomic and quality traits in a bioenergy crop, sweet sorghum [*Sorghum bicolor* (L.) Moench]. *Afr. J. Biotechnol.* **16** (47): 2189-2200.

Clayton, W. D. and Renvoize, S. A. (1986). Genera graminum. Grasses of the world. *Genera graminum. Grasses of the World.* **13**.

Collard, B. C., Jahufer, M., Brouwer, J. and Pang, E. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica.* **142** (1): 169-196.

CSA. (2020). Ethiopian statistical abstract for the year 2016/17. Report on area and production of major crops. **1**.

Dalvand, M., Zafari, D., Soleimani Pari, M., Roohparvar, R. and Tabib Ghafari, S. (2018). Studying Genetic Diversity in Zymoseptoria tritici, Causal Agent of Septoria Tritici Blotch, by Using ISSR and SSR Markers. *J. Agric. Sci. Technol.* **20** (6): 1307-1316.

Danquah, A., Galyuon, I. K., Otwe, E. P. and Asante, D. K. (2019). Genetic diversity in some Ghanaian and Malian sorghum [*Sorghum bicolor* (L) Moench] accessions using SSR markers. *Afr. J. Biotechnol.* **18** (27): 591-602.

DArT. (2019). Plant DNA extraction protocol for DArT. Diversity array technology.

De Wet, J. (1978). Special paper: systematics and evolution of sorghum sect. Sorghum (Gramineae). *Am. J. Bot.* **65** (4): 477-484.

Derese Solomon, A., Shimelis Hussein, Mwadzingeni, L. and Laing, M. (2018). Agro-morphological characterisation and selection of sorghum landraces Section B Soil and plant science.

Deu, M., Rattunde, F. and Chantereau, J. (2006). A global view of genetic diversity in cultivated sorghums using a core collection. *Genome*. **49** (2): 168-180.

Deu, M., Sagnard, F., Chantereau, J., Calatayud, C., Hérault, D., Mariac, C., Pham, J.-L., Vigouroux, Y., Kapran, I. and Traore, P. (2008). Niger-wide assessment of in situ

sorghum genetic diversity with microsatellite markers. *Theor. Appl. Genet.* **116** (7): 903-913.

Dillon, S. L., Lawrence, P. K. and Henry, R. J. (2005). The new use of Sorghum bicolor-derived SSR markers to evaluate genetic diversity in 17 Australian Sorghum species. *Plant Genet. Resour*. **3** (1): 19-28.

Dillon, S. L., Shapter, F. M., Henry, R. J., Cordeiro, G., Izquierdo, L. and Lee, L. S. (2007). Domestication to crop improvement: genetic resources for Sorghum and S accharum (Andropogoneae). *Ann. Bot.* **100** (5): 975-989.

Djè, Y., Heuertz, M., Ater, M., Lefèbvre, C. and Vekemans, X. (2004). In situ estimation of outcrossing rate in sorghum landraces using microsatellite markers. *Euphytica*. **138** (3): 205-212.

Doggett, H. (1965). The development of the cultivated sorghums. *Genet. Resour. Crop Evol.* 50-69.

Doggett, H. (1988). Sorghum. London Harlow (GB). In: Longman Scientific Technical, (2 ème edition).

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4** (2): 359-361.

EIAR. (2014). Ethiopian strategy for sorghum. Addis Ababa, Ethiopia.

EIAR, E. I. o. A. R. (2021). Ethiopian strategy for sorghum 2014-2024. 1-10.

Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences*, **95** (25): 14863-14868.

Ellstrand, N. and Foster, K. (1983). Impact of population structure on the apparent outcrossing rate of grain sorghum (Sorghum bicolor). *Theor. Appl. Genet.* **66** (3-4): 323-327.

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14** (8): 2611-2620.

Excoffier, L. and Lischer, H. (2006). An integrated software package for population genetics data analysis. *Institute of Zoology, University of Berne, Switzerland*.

Falconer, D. and Mackay, T. (1996). Introduction to Quantitative Genetics 4th edBenjamin Cummings. *Harlow*.

FAOSTAT. (2015). Food and Agriculture Organization of the United Nations Database of agricultural production.

FAOSTAT. (2019). Food and Agriculture Organization of the United Nations (FAO). *FAOSTAT database.*

FAOSTAT. (2021). Food and Agriculture Organization of the United Nations Database of agricultural production.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution,* **39** (4): 783-791.

Frankel, O. H., Brown, A. H. D. and Burdon, J. J. (1995). The Conservation of Plant Biodiversity: Cambridge University Press.

Gabisa Ejeta and Grenier, C. (2005). Sorghum and its weedy hybrids. Crop ferality and volunteerisim. *Boca Raton.* 123-135.

Gebrekidan Brhane. (1981). Utilization of germplasm in sorghum improvement. Paper presented at the Proceedings of the International Symposium on Sorghum,(House LR, Mughogho LK and Peacock JM, eds.).

Gebremedhn Gebregergs and Firew Mekbib. (2020). Estimation of genetic variability, heritability, and genetic advance in advanced lines for grain yield and yield components of sorghum [Sorghum bicolor (L.) Moench] at Humera, Western Tigray, Ethiopia. *Cogent food agric.* **6** (1): 1764181.

Gedifew Gebrie and Tsige Genet. (2019). Morphological characterization and evaluation of sorghum [*Sorghum bicolor* (L.) Moench] landraces in Benishangul Gumuz, North-western Ethiopia. *Greener J. Agric. Sci.* **9** (1): 37-56.

GRIN. (2012). USDA-ARS National Genetic Resources Program, Germplasm Resource Information Network (GRIN). Online database, National Germplasm Resources Laboratory.

Hariprasanna, K. and Rakshit, S. (2016). Economic importance of sorghum. *The sorghum genome*. 1-25.

Harlan, J. R. and De Wet, J. M. (1972). A simplified classification of cultivated sorghum. *Crop sci.* **12** (2): 172-176.

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

ICRISAT. (2015). Sorghum (Sorghum bicolor L. Moench). International Crops Research Institute for the Semi-Arid Tropics.

Iezzoni, A. F. and Pritts, M. P. (1991). Applications of principal component analysis to horticultural research. *HortScience*. **26** (4): 334-338.

Jayaramachandran, M., Kumaravadivel, N., Kandasamy, G. and Eapen, S. (2011). Comparison of genetic variability induced by γ radiation and tissue culture in sorghum. *Int. j.* bio-*resour. stress manag.* **2** (3): 329-333.

Johnson, H. W., Robinson, H. and Comstock, R. (1955). Estimates of genetic and environmental variability in soybeans. *J. Agron.* **47** (7): 314-318.

Kaiser, H. F. (1961). A note on guttman's lower bound for the number of common factors. *Br. J. Stat. Psychol.* **14** (1): 1-2.

Kassahun Tesfaye. (2017). Genetic diversity study of sorghum (Sorghum bicolor (L.) Moenc) genotypes, Ethiopia. *Acta Univ. Sapientiae, Agric. Environ*. **9** (1): 44-54.

Kidanemaryam Wagaw. (2019). Review on Mechanisms of Drought Tolerance in Sorghum (Sorghum bicolor (L.) Moench) Basis and Breeding Methods. *Res. J. Agric. Sci.* **7** (2): 87-99.

Kimber, C. T., Dahlberg, J. A. and Kresovich, S. (2013). The gene pool of Sorghum bicolor and its improvement. In Genomics Inform. 23-41.

Kong, W., Jin, H., Goff, V. H., Auckland, S. A., Rainville, L. K. and Paterson, A. H. (2020). Genetic Analysis of Stem Diameter and Water Contents To Improve Sorghum Bioenergy Efficiency. *G3-Genes Genom Genet.* **10** (11): 3991-4000.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15** (5): 1179-1191.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35** (6): 1547.

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

Liu, K. and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *J. Bioinform.* **21** (9): 2128-2129.

Mace, E. S., Rami, J.-F., Bouchet, S., Klein, P. E., Klein, R. R., Kilian, A., Wenzl, P., Xia, L., Halloran, K. and Jordan, D. R. (2009). A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. *BMC Plant Biol.* **9** (1): 1-14.

Menz, M., Klein, R., Mullet, J., Obert, J., Unruh, N. and Klein, P. (2002). A high-density genetic map of Sorghum bicolor (L.) Moench based on 2926 AFLP®, RFLP and SSR markers. *Plant Mol. Biol.* **48** (5): 483-499.

Mofokeng, A., Shimelis, H., Tongoona, P. and Laing, M. (2014). A genetic diversity analysis of South African sorghum genotypes using SSR markers. *S. Afr. J. Plant Soil.* **31** (3): 145-152.

More, A., Kalpande, H. and Dhutmal, R. (2018). Genetic Divergence Studies in Sorghum (Sorghum bicolor L.) Land Races for Yield and Yield Parameters. *Int. J. Curr. Microbiol. App. Sci*. **6**: 393-399.

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B. and Mitchell, S. E. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences*. **110** (2): 453-458.

Motlhaodi, T., Geleta, M., Chite, S., Fatih, M., Ortiz, R. and Bryngelsson, T. (2017). Genetic diversity in sorghum [Sorghum bicolor (L.) Moench] germplasm from Southern Africa as revealed by microsatellite markers and agro-morphological traits. *Genet. Resour. Crop Evol.* **64** (3): 599-610.

Muui, C. W., Muasya, R. M., Kirubi, D. T., Runo, S. M. and Karugu, A. (2016). Genetic variability of sorghum landraces from lower Eastern Kenya based on simple sequence repeats (SSRs) markers. *Afr. J. Biotechnol.* **15** (8): 264-271.

Nei, M. and Takezaki, N. (1983). Estimation of genetic distances and phylogenetic trees from DNA analysis. *Proc 5th World Cong Genet Appl Livstock Prod*. **21** (21): 405-412.

Ng'uni, D., Geleta, M. and Bryngelsson, T. (2011). Genetic diversity in sorghum (Sorghum bicolor (L.) Moench) accessions of Zambia as revealed by simple sequence repeats (SSR). *Hereditas*. **148** (2): 52-62.

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

Norwood, C. (2015). Chinese toast australian sorghum. *Grains Research and Development corporation*.

O'Hara, I., Kent, G., Albertson, P., Harrison, M., Hobson, P., McKenzie, N., Moghaddam, L., Moller, D., Rainey, T. and Stolz, W. (2013). Sweet sorghum: Opportunities for a new, renewable fuel and food industry in Australia. *RIRDC Publication No. 13/087*

Pandian, S., Satish, L., Shilpha, J. and Ramesh, M. (2020). Genetic Diversity Analysis Reveals Strong Population Structure in Sorghum Germplasm Collection. Proceedings of the National Academy of Sciences, India Section B: *Biol. Sci.* **90** (1): 179-190.

Parikka, M. (2004). Global biomass fuel resources. *Biomass and bioenergy*, **27** (6): 613-620.

Patterson, H. and Williams, E. (1976). A new class of resolvable incomplete block designs. *Biometrika*. **63** (1): 83-92.

Pavel, A. B. and Vasile, C. I. (2012). PyElph-a software tool for gel images analysis and phylogenetics. *BMC Bioinform.* **13** (1): 1-6.

Peakall, R. and Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes.* **6** (1): 288-295.

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*. **155** (2): 945-959.

Purseglove, J. (1972). Taccaceae. Tropical Crops. Monocotyledons 2. *Longman Group Ltd, London*, **607** (2): 517-518.

Rami, J.-F., Dufour, P., Trouche, G., Fliedel, G., Mestres, C., Davrieux, F., Blanchard, P. and Hamon, P. (1998). Quantitative trait loci for grain quality, productivity, morphological and agronomical traits in sorghum (Sorghum bicolor L. Moench). *Theor. and Appl. Genet.* **97** (4): 605-616.

Ramu, P., Kassahun, B., Senthilvel, S., Kumar, C. A., Jayashree, B., Folkertsma, R., Reddy, L. A., Kuruvinashetti, M., Haussmann, B. and Hash, C. T. (2009). Exploiting rice–sorghum synteny for targeted development of EST-SSRs to enrich the sorghum genetic linkage map. *Theor. and Appl. Genet.* **119** (7): 1193-1204.

Ranjith, P., Ghorade, R., Kalpande, V. and Dange, A. (2017). Genetic variability, heritability and genetic advance for grain yield and yield components in sorghum. *Int. J. Farm Sci*. **7** (1): 90-93.

Roa, N.G. Murty and Rana, U. R. (2002). Sorghum. In: Evolution and Adaptation of Cereal Crops. *Science.* 213–238.

Rosenow, D. and Dahlberg, J. (2000). Collection, conversion, and utilization of sorghum. Sorghum: origin, history, technology and production. *Frederiksen:* 471-24237-3.

Ruiz-Chután, J. A., Salava, J., Janovská, D., Žiarovská, J., Kalousová, M. and Fernández, E. (2019). Assessment of genetic diversity in Sorghum bicolor using RAPD markers. *Genetika.* **51** (3): 789-803.

Sabiel, S. A., Noureldin, I., Baloch, S. K., Baloch, S. U. and Bashir, W. (2016). Genetic variability and estimates of heritability in sorghum (Sorghum bicolor L.) genotypes grown in a semiarid zone of Sudan. *Arch. Agron. Soil Sci*. **62** (1): 139-145.

Saidur, R., Abdelaziz, E., Demirbas, A., Hossain, M. and Mekhilef, S. (2011). A review on biomass as a fuel for boilers. *Renew. Sustain. Energy Rev.* **15** (5): 2262-2289.

Salih, S. A., Herslman, L., Labuschange, M. T. and Mohammed, A. H. (2016). Assessment of genetic diversity of sorghum [Sorghum bicolor (L.) Moench] germplasm in East and Central Africa. *World J. Microbiol. Biotechnol.* **1** (03): 113-120.

Beyene Seboka and van Hintum, T. (2006). The Dynamics of on-farm Management of Sorghum in Ethiopia: Implication for the Conservation and Improvement of Plant Genetic Resources. Genet. Resour. *Crop Evol.* **53** (7): 1385-1403.

Seetharam, K. and Ganesamurthy, K. (2013). Characterization of sorghum genotypes for yield and other agronomic traits through genetic variability and diversity analysis. *Electron. J. Plant Breed.* **4** (1): 1073-1079.

Shehzad, T., Okuizumi, H., Kawase, M. and Okuno, K. (2009). Development of SSR-based sorghum (Sorghum bicolor (L.) Moench) diversity research set of germplasm and its evaluation by morphological traits. Genet. Resour. *Crop Evol.* **56** (6): 809-827.

Shehzad, T. I., Hiroyoshi Okuno, Kazutoshi. (2009). Genome-wide association mapping of quantitative traits in sorghum (*Sorghum bicolor* (L.) Moench) by using multiple models. *Breed. Sci.* **59** (3): 217-227.

Shewayrga, H., Jordan, D. R. and Godwin, I. D. (2008). Genetic erosion and changes in distribution of sorghum (Sorghum bicolor L. (Moench)) landraces in north-eastern Ethiopia. *Plant Genet. Resour.* **6** (1): 1-10.

Sihag, N., Madhubala Kurmanchali, D. P. and Pandey, P. (2019). Genetic divergence analysis of sorghum [Sorghum bicolor (L.) Moench] germplasm lines. *J. Pharmacogn. Phytochem.* **8** (5): 1837-1839.

Singh, B. (2001). Organisation for crop improvement in India. *Plant Breed.* 801-830.

Singh, P. and Narayanan, S. (1993). Biometricl Techniques in Plant Breeding. *New Delhi*.

Singh, R. K. and Chaudhary, B. D. (1977). Biometrical methods in quantitative genetic analysis. *Biometrical methods in quantitative genetic analysis.*

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*. **236** (4803): 787-792.

Sneath, P. H. and Sokal, R. R. (1973). Numerical taxonomy. *The principles and practice of numerical classification*.

Snowden, J. (1955). The wild fodder sorghums of the section Eu-sorghum. Bot. J. Linn. Soc. **55** (358): 191-260.

Snowden, J. D. (1936). The cultivated races of sorghum. *The cultivated races of sorghum.*

Solomon Assefa, Shimelis, H., Mwadzingeni, L. and Laing, M. (2018). Agro-morphological characterisation and selection of sorghum landraces. *J. Soil Sci. Plant Nutr.* **68** (7): 585-595.

Sukumaran, S. and Yu, J. (2014). Association mapping of genetic resources: achievements and future perspectives. *Plant Genet. Resour.* 207-235.

Tafere Mulualem, Sentayehu Alamrew, Taye Mindaye and Dagne Wegary. (2018). Genetic Variability, Heritability and Genetic advance for Agronomical Traits of Ethiopian Sorghum [Sorghum bicolor (L.) Moench] Genotypes. *Acad. Res. J. Agri. Sci. Res.* **6** (4): 251-259.

Tang, C., Luo, F., Li, X., Pei, Z., Gao, J. and Sun, S. (2015). Responsiveness of sweet sorghum in yield and quality related traits to environmental factors. *Acta Agronomica Sinica.* **41** (10): 1612-1618.

Tang, C., Yang, X., Chen, X., Ameen, A. and Xie, G. (2018). Sorghum biomass and quality and soil nitrogen balance response to nitrogen rate on semiarid marginal land. *Field Crops Res.* **215**: 12-22.

Taye Mindaye, Mace, E. S., Godwin, I. D. and Jordan, D. R. (2016). Heterosis in locally adapted sorghum genotypes and potential of hybrids for increased productivity in contrasting environments in Ethiopia. *Crop J.* **4** (6): 479-489.

Tesfamichael Abraha, Githiri, S., Kasili, R., Skilton, R., Solomon, M. and Nyende, A. (2014). Genetic diversity analysis of Eritrean sorghum (Sorghum bicolor (L.) Moench) germplasm using SSR markers. *Mol. Plant Breed.* **5**.

Tesfaye Disasa. (2016). Genetic Diversity Analysis and QTL Mapping of Selected Traits in Sorghum *(Sorghum bicolor* (L.) Moench*).* Addis Ababa University,

Tesfaye Disasa, Tileye Feyissa and Belayneh Admassu. (2017). Characterization of ethiopian sweet sorghum accessions for 0 Brix, morphological and grain yield traits. *Sugar Tech*. **19** (1): 72-82.

Tesfaye Tesso, Kapran, I., Grenier, C., Snow, A., Sweeney, P., Pedersen, J., Marx, D., Bothma, G. and Gebisa Ejeta. (2008). The potential for crop-to-wild gene flow in sorghum in Ethiopia and Niger: A geographic survey. *Crop Sci.* **48** (4): 1425-1431.

Thudi, M. and Fakrudin, B. (2011). Identification of unique alleles and assessment of genetic diversity of rabi sorghum accessions using simple sequence repeat markers. *J. Plant Biochem.* **20** (1): 74-83.

Udeh, I. and Ogbu, C. (2011). Principal component analysis of body measurements in three strains of broiler chicken. *Sci. World J.* **6** (2): 11-14.

USDA. (2019). World sorghum production 2019/2020. World agricultural production.com.

Van Beuningen, L. and Busch, R. (1997). Genetic diversity among North American spring wheat cultivars: III. Cluster analysis based on quantitative morphological traits. *Crop sci.* **37** (3): 981-988.

Vasilakoglou, I., Dhima, K., Karagiannidis, N. and Gatsis, T. (2011). Sweet sorghum productivity for biofuels under increased soil salinity and reduced irrigation. *Field Crops Res.* **120** (1): 38-46.

Venkateswaran, K., Elangovan, M. and Sivaraj, N. (2019). Origin, domestication and diffusion of *Sorghum bicolor*. In *Breeding Sorghum for diverse end uses*. 15-31.

Wang, D., Bean, S., McLaren, J., Seib, P., Madl, R., Tuinstra, M., Shi, Y., Lenz, M., Wu, X. and Zhao, R. (2008). Grain sorghum is a viable feedstock for ethanol production. *J. Ind. Microbiol. Biotechnol.* **35** (5): 313-320.

Wang, M. L., Dean, R., Erpelding, J. and Pederson, G. (2006). Molecular genetic evaluation of sorghum germplasm differing in response to fungal diseases: Rust (Puccinia purpurea) and anthracnose (Collectotrichum graminicola). *Euphytica*. **148** (3): 319-330.

Wang, M. L., Zhu, C., Barkley, N. A., Chen, Z., Erpelding, J. E., Murray, S. C., Tuinstra, M. R., Tesso, T., Pederson, G. A. and Yu, J. (2009). Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. *Theor. Appl. Genet.* **120** (1): 13-23.

Weerasooriya, D. K., Maulana, F. R., Bandara, A. Y., Tirfessa, A., Ayana, A., Mengistu, G., Nouh, K. and Tesso, T. T. (2016). Genetic diversity and population structure among sorghum (Sorghum bicolor, L.) germplasm collections from Western Ethiopia. *Afr. J. Biotechnol.* **15** (23): 1147-1158.

Xiong, S., Zhang, Q.-G., Zhang, D.-Y. and Olsson, R. (2008). Influence of harvest time on fuel characteristics of five potential energy crops in northern China. *Bioresour. Technol.* **99** (3): 479-485.

Yeh, F. C. (1999). POPGENE (version 1.3. 1). Microsoft Window-Bases Freeware for Population Genetic Analysis. *http://www. ualberta. ca/~ fyeh/.*

Tekle Yoseph and Zemach Sorsa. (2014). Evaluation of sorghum (Sorghum bicolor (L.) Moench) varieties, for yield and yield components at Kako, Southern Ethiopia. *Plant Sci. J.* **2** (4): 129-133l.

Zhu, M., Chen, J., Yuyama, N., Luo, L., Xiao, X., Lv, Y., Liu, Y. and Cai, H. (2020). Genetic Diversity and Population Structure of Broomcorn Sorghum Investigated with Simple Sequence Repeat Markers. *Trop. Plant Biol.* **13** (1): 62-72.

8. APPENDICES

Appendix 1. Passport description of the studied sorghum accessions.

S/No.	Accession	Region	Latitude	Longitude	Altitude
1	9137	Oromia	09-15-04-N	42-02-02-E	1851
2	9142	Oromia	09-23-58-N	41-16-55-E	1626
3	9160	Oromia	09-25-55-N	41-34-47-E	2033
4	15823	Oromia	08-17-57-N	39-19-27-E	1680
5	15836	Harari	09-20-17-N	42-06-22-E	2037
6	15888	Oromia	09-10-60-N	41-05-17-E	1985
7	15909	Oromia	08-49-10-N	40-30-09-E	1796
8	15914	Oromia	08-56-29-N	40-50-17-E	1976
9	15934	Oromia	09-08-19-N	40-49-25-E	1647
10	15946	Oromia	09-10-00-N	40-42-19-E	1516
11	15953	Oromia	08-56-49-N	40-50-26-E	1919
12	16096	Amhara	12-08-30-N	37-45-59-E	2216
13	16097	Amhara	12-08-56-N	34-44-20-E	2335
14	16119	Oromia	08-56-00-N	37-03-59-E	1681
15	16154	Oromia	09-29-40-N	36-23-52-E	1396
16	16168	Oromia	09-01-05-N	37-00-36-E	1614
17	16171	Oromia	09-57-58-N	35-00-36-E	1639
18	16185	Oromia	08-39-28-N	35-00-57-E	1344
19	17508	Amhara	10-07-02-N	39-53-34-E	1250
20	17510	Amhara	10-20-48-N	39-54-48-E	1537
21	17513	Amhara	10-27-01-N	39-56-12-E	1658
22	17517	Oromia	08-50-02-N	37-19-32-E	1760
23	17538	Amhara	09-56-34-N	39-27-14-E	2603
24	19619	Tigray	14-16-30-N	38-57-44-E	1983
25	19620	Tigray	14-17-43-N	38-57-38-E	1955
26	19624	Tigray	14-19-16-N	38-58-40-E	1670
27	19625	Tigray	14-19-8 -N	38-57-47-E	1945
28	19626	Tigray	14-19-28-N	38-57-10-E	1946
29	19628	Tigray	14-19-10-N	38-57-8 -E	1952
30	19629	Tigray	14-07-25-N	38-04-59-E	1276
31	23655	Benishangul Gumuz	11-16-05-N	35-17-17-E	858
32	27897	Gambella	07-13-15-N	35-18-08-E	1263
33	27899	Oromia	07-13-14-N	13-16-72-E	1216
34	27900	Gambella	07-15-89-N	35-13-50-E	1006
35	27902	Gambella	07-16-39-N	35-19-16-E	1258

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
36	27905	Gambella	07-13-11-N	35-20-03-E	1430
37	27907	SNNP	06-46-50-N	35-12-14-E	1367
38	27909	SNNP	06-46-60-N	35-12-15-E	1372
39	27912	SNNP	06-49-22-N	35-16-37-E	1160
40	27913	SNNP	06-52-27-N	35-32-08-E	1775
41	27917	SNNP	06-61-46-N	35-31-13-E	1214
42	27920	SNNP	06-47-41-N	35-32-08-E	1511
43	27921	SNNP	06-53-45-N	35-33-56-E	228
44	69034	Oromia	08-23-00-N	39-44-00-E	2275
45	69038	Oromia	08-01-00-N	39-50-00-E	2660
46	69039	Oromia	08-01-00-N	39-50-00-E	2660
47	69138	Amhara	12-03-00-N	36-57-00-E	2060
48	69171	SNNP	06-59-00-N	37-53-00-E	2050
49	69173	SNNP	06-43-00-N	37-44-00-E	1760
50	69216	Amhara	11-48-00-N	39-35-00-E	1970
51	69227	Oromia	09-05-00-N	40-51-00-E	1750
52	69229	Oromia	09-04-00-N	40-50-00-E	1710
53	69230	Oromia	09-06-00-N	40-49-00-E	1760
54	69232	Oromia	09-05-00-N	40-48-00-E	2090
55	69234	Oromia	09-03-00-N	40-44-00-E	1840
56	69236	Oromia	09-04-00-N	40-42-00-E	1640
57	69237	Oromia	09-04-00-N	40-42-00-E	1640
58	69238	Oromia	09-14-00-N	40-37-00-E	1470
59	69248	Amhara	11-31-00-N	39-15-00-E	1870
60	69261	Oromia	09-09-00-N	41-02-00-E	2190
61	69264	Oromia	09-13-00-N	41-07-00-E	1900
62	69265	Oromia	09-17-00-N	41-05-00-E	1965
63	69266	Oromia	09-16-00-N	41-25-00-E	2200
64	69284	Oromia	08-59-00-N	34-58-00-E	
65	69291	Amhara	09-57-00-N	39-51-00-E	
66	69292	Amhara	09-57-00-N	39-51-00-E	
67	69294	SNNP	06-13-00-N	37-34-00-E	
68	69305	Amhara	11-55-00-N	37-47-00-E	2030
69	69307	Amhara	13-21-00-N	37-54-00-E	1480
70	69310	Amhara	13-20-00-N	37-52-00-E	1380
71	69311	SNNP	07-50-00-N	39-06-00-E	2450
72	69312	SNNP	06-24-00-N	37-38-00-E	2360
73	69343	Amhara	10-27-00-N	38-12-00-E	2570
74	69344	Oromia	10-01-00-N	38-14-00-E	2540
75	69493	SNNP	05-25-00-N	37-24-00-E	1150

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
76	69496	SNNP	05-25-00-N	37-24-00-E	1150
77	69500	SNNP	05-25-00-N	37-24-00-E	1300
78	69503	SNNP	05-25-00-N	37-24-00-E	1250
79	69517	Amhara	12-23-00-N	37-23-00-E	2110
80	69520	Amhara	12-33-00-N	37-24-00-E	2200
81	69525	Oromia	08-47-00-N	36-39-00-E	2260
82	69528	Oromia	08-52-00-N	35-04-00-E	1480
83	69529	Oromia	08-52-00-N	35-04-00-E	1480
84	69540	Oromia	08-21-00-N	36-21-00-E	1900
85	69544	Dire Dawa	07-33-00-N	36-37-00-E	1860
86	69546	SNNP	07-10-00-N	36-25-00-E	1780
87	70933	Amhara	09-35-00-N	39-44-00-E	2880
88	75284	Amhara	10-25-00-N	39-56-00-E	1763
89	210911	Amhara	12-27-00-N	37-22-00-E	1960
90	210917	SNNP	05-43-00-N	37-23-00-E	1560
91	210922	SNNP	05-25-00-N	37-24-00-E	1150
92	210945	Amhara	11-12-00-N	40-01-00-E	1670
93	210946	Amhara	11-12-00-N	40-01-00-E	1670
94	210951	Amhara	11-48-00-N	39-35-00-E	1970
95	210953	Amhara	11-02-00-N	39-45-00-E	1830
96	210955	Oromia	09-05-00-N	40-51-00-E	1750
97	210957	Oromia	09-05-00-N	40-51-00-E	1750
98	210958	Oromia	09-05-00-N	40-51-00-E	1750
99	210965	Oromia	09-04-00-N	40-42-00-E	1640
100	210972	Amhara	11-31-00-N	39-15-00-E	1870
101	210976	Oromia	09-09-00-N	41-02-00-E	2190
102	210981	Oromia	09-13-00-N	41-07-00-E	1900
103	212541	Amhara	09-54-00-N	38-54-00-E	1490
104	212542	Amhara	09-56-00-N	38-54-00-E	1580
105	212546	Amhara	10-03-00-N	38-57-00-E	1540
106	223504	Oromia	09-21-00-N	36-38-00-E	1830
107	223505	Oromia	09-02-00-N	36-38-00-E	1830
108	223513	Oromia	09-02-00-N	36-26-00-E	2100
109	223517	Oromia	09-02-00-N	36-14-00-E	1390
110	223554	Oromia	08-04-00-N	36-27-00-E	1610
111	223561	Dire Dawa	08-08-00-N	37-30-00-E	1880
112	223565	SNNP	08-11-00-N	37-47-00-E	1840
113	223573	SNNP	07-18-00-N	37-48-00-E	1950
114	223575	Oromia	07-13-00-N	38-34-00-E	1860

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
115	223576	Oromia	07-13-00-N	38-34-00-E	1860
116	223578	Oromia	07-15-00-N	38-29-00-E	1700
117	223579	Oromia	07-14-00-N	38-38-00-E	1900
118	223581	Oromia	07-21-00-N	38-45-00-E	1930
119	223582	Oromia	07-21-00-N	38-45-00-E	1930
120	223585	Oromia	08-13-00-N	38-52-00-E	1720
121	223586	Oromia	08-13-00-N	38-52-00-E	1720
122	223587	SNNP	08-16-00-N	38-54-00-E	1700
123	225832	SNNP	06-18-00-N	36-54-00-E	2270
124	225836	SNNP	06-38-00-N	37-38-00-E	1270
125	225838	SNNP	06-38-00-N	37-38-00-E	1270
126	226046	Amhara	11-12-00-N	40-00-00-E	1660
127	226050	Amhara	10-47-00-N	38-41-00-E	2445
128	226056	Amhara	10-42-00-N	38-53-00-E	2540
129	226057	Amhara	10-37-00-N	38-48-00-E	2450
130	226062	Amhara	10-46-00-N	38-48-00-E	2535
131	226066	Amhara	11-28-00-N	38-11-00-E	2520
132	226072	Amhara	11-25-00-N	37-57-00-E	2360
133	226074	Amhara	12-30-00-N	37-24-00-E	1905
134	226079	Amhara	12-20-00-N	37-13-00-E	1855
135	229832	Benishangul Gumuz	11-04-00-N	36-33-00-E	1300
136	229833	Oromia	11-05-00-N	36-34-00-E	1350
137	236730	Oromia	08-59-00-N	37-28-00-E	2100
138	236732	Oromia	08-48-00-N	37-54-00-E	2400
139	236733	Oromia	08-48-00-N	37-54-00-E	2270
140	236734	Oromia	08-48-00-N	37-54-00-E	2100
141	236767	Dire Dawa	09-21-13-N	41-54-59-E	1950
142	237029	Oromia	07-30-00-N	39-05-00-E	2390
143	237038	Oromia	08-09-00-N	39-10-00-E	2160
144	237039	Oromia	08-09-00-N	39-21-00-E	2340
145	237040	Oromia	08-70-00-N	39-20-00-E	2650
146	237042	Oromia	07-40-00-N	40-12-00-E	1990
147	237043	Amhara	08-50-00-N	39-20-00-E	1690
148	237044	Amhara	08-50-00-N	39-20-00-E	1700
149	237045	Amhara	09-30-00-N	39-20-00-E	2020
150	237264	Tigray	12-26-00-N	39-56-00-E	1500
151	237281	Tigray	14-12-00-N	38-56-00-E	2000
152	237287	Tigray	14-19-00-N	38-49-00-E	1420
153	237291	Tigray	14-12-00-N	38-56-00-E	1970

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
154	237292	Tigray	14-12-00-N	38-56-00-E	2200
155	237301	Tigray	14-10-00-N	38-45-00-E	2190
156	237308	Tigray	14-00-00-N	38-00-00-E	1520
157	237309	Tigray	11-57-00-N	38-41-00-E	
158	238264	Oromia	09-12-14-N	42-04-72-E	1850
159	238379	Tigray	12-31-00-N	39-31-00-E	2330
160	238380	Tigray	12-5 -00-N	38-02-00-E	2150
161	238386	Tigray	14-04-00-N	38-05-00-E	
162	238390	Tigray	13-04-00-N	39-35-00-E	2000
163	238391	Tigray	13-04-00-N	39-35-00-E	1750
164	238392	Tigray	14-02-00-N	38-04-00-E	2000
165	238394	Tigray	14-02-00-N	38-04-00-E	1950
166	238397	Tigray	14-02-00-N	38-04-00-E	2110
167	238403	Tigray	14-04-00-N	38-05-00-E	2050
168	238406	Tigray	13-09-00-N	39-09-00-E	1920
169	238407	Tigray	13-09-00-N	39-09-00-E	1870
170	238408	Tigray	14-05-00-N	39-06-00-E	1990
171	238411	Tigray	14-05-00-N	38-04-00-E	1950
172	238412	Tigray	14-05-00-N	38-04-00-E	1930
173	238416	Tigray	14-05-00-N	38-04-00-E	1930
174	238417	Tigray	14-05-00-N	38-04-00-E	
175	238420	Tigray	14-06-00-N	38-09-00-E	1500
176	238421	Tigray	14-06-00-N	38-09-00-E	1500
177	238434	Amhara	12-09-00-N	38-06-00-E	2100
178	238455	Amhara	12-08-00-N	38-05-00-E	1890
179	238456	Amhara	10-02-00-N	38-02-00-E	2250
180	239113	Dire Dawa	09-36-26-N	41-06-66-E	1200
181	239114	Dire Dawa	09-35-74-N	41-46-38-E	1220
182	239116	Dire Dawa	09-34-08-N	41-44-45-E	1320
183	239118	Dire Dawa	09-32-35-N	41-52-00-E	1360
184	239121	Dire Dawa	09-31-04-N	41-47-54-E	1440
185	239122	Dire Dawa	09-30-26-N	41-46-13-E	1540
186	239124	Dire Dawa	09-29-64-N	41-44-79-E	1640
187	239126	Dire Dawa	09-30-00-N	41-44-64-E	1650
188	239128	Dire Dawa	09-39-19-N	41-59-17-E	1250
189	239133	Dire Dawa	09-35-83-N	42-06-46-E	1430
190	239134	Dire Dawa	09-35-83-N	42-06-46-E	1430
191	239135	Dire Dawa	09-37-97-N	42-08-66-E	1380
192	239136	Dire Dawa	09-37-97-N	42-08-66-E	1380

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
193	239137	Dire Dawa	09-37-97-N	42-08-66-E	1380
194	239138	Dire Dawa	09-37-97-N	42-08-66-E	1380
195	239141	Dire Dawa	09-38-69-N	42-16-55-E	1510
196	239151	Amhara	10-58-33-N	39-46-20-E	1560
197	239152	Amhara	10-58-33-N	39-46-20-E	1560
198	239161	Amhara	10-58-33-N	39-46-20-E	1560
199	239162	Amhara	10-58-33-N	39-46-20-E	1560
200	239163	Amhara	10-58-33-N	39-46-20-E	1560
201	239164	Amhara	10-58-33-N	39-46-20-E	1560
202	239166	Amhara	10-58-33-N	39-46-20-E	1560
203	239168	Amhara	10-58-33-N	39-46-20-E	1560
204	239170	Amhara	10-58-33-N	39-46-20-E	1560
205	239173	Amhara	10-58-33-N	39-46-20-E	1560
206	239174	Amhara	10-58-33-N	39-46-20-E	1560
207	239175	Amhara	10-58-33-N	39-46-20-E	1560
208	239183	Amhara	10-58-33-N	39-46-20-E	1560
209	239185	Amhara	10-58-33-N	39-46-20-E	1560
210	239191	Amhara	10-58-33-N	39-46-20-E	1560
211	239192	Amhara	10-58-33-N	39-46-20-E	1560
212	239193	Amhara	10-58-17-N	39-46-40-E	1580
213	239194	Amhara	10-58-17-N	39-46-40-E	1580
214	239195	Amhara	10-58-17-N	39-46-40-E	1580
215	239196	Amhara	10-58-17-N	39-46-40-E	1580
216	239197	Amhara	10-58-17-N	39-46-40-E	1580
217	239206	Amhara	10-58-17-N	39-46-40-E	1580
218	239207	Amhara	10-58-17-N	39-46-40-E	1580
219	239209	Amhara	10-21-09-N	39-56-72-E	1420
220	239210	Amhara	10-21-09-N	39-56-72-E	1420
221	239212	Amhara	10-21-09-N	39-56-72-E	1420
222	239214	Amhara	10-21-09-N	39-56-72-E	1420
223	239215	Amhara	10-21-09-N	39-56-72-E	1420
224	239217	Amhara	10-58-17-N	39-46-40-E	1580
225	239219	Amhara	11-12-44-N	40-00-96-E	1670
226	239222	Amhara	10-58-17-N	39-46-40-E	1580
227	239225	Amhara	10-58-17-N	39-46-40-E	1580
228	239240	Amhara	11-50-03-N	39-45-97-E	1510
229	239249	Amhara	11-43-65-N	38-55-73-E	2320
230	239262	Dire Dawa	09-16-41-N	42-06-06-E	1880
231	239761	Somali	09-17-19-N	42-06-43-E	1900
232	240759	Somali	09-29-92-N	42-37-75-E	1900

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
233	241185	Oromia	09-06-30-N	40-50-45-E	1600
234	241187	Oromia	09-03-90-N	40-51-86-E	1750
235	241189	Oromia	09-03-16-N	40-54-31-E	2000
236	241194	Oromia	08-34-02-N	40-18-65-E	1770
237	241196	Oromia	08-36-56-N	40-19-79-E	1720
238	241197	Oromia	08-37-79-N	40-19-78-E	1720
239	241201	Oromia	08-50-17-N	40-35-23-E	1670
240	241206	Oromia	09-05-96-N	40-56-78-E	2130
241	241207	Oromia	09-06-78-N	40-58-34-E	
242	241208	Oromia	09-08-83-N	41-01-29-E	2130
243	241233	Oromia	09-23-02-N	41-30-40-E	2020
244	241236	Dire Dawa	09-30-85-N	41-52-81-E	1380
245	241237	Oromia	09-27-24-N	41-54-77-E	2020
246	241239	Oromia	09-26-35-N	41-47-30-E	
247	241241	Dire Dawa	09-21-99-N	41-46-22-E	1940
248	241254	Somali	09-16-08-N	42-10-49-E	1540
249	241255	Somali	09-16-01-N	42-10-10-E	1520
250	241258	Somali	09-28-19-N	42-40-54-E	1840
251	241278	Dire Dawa	08-59-15-N	41-34-46-E	1680
252	241279	Oromia	09-54-85-N	41-34-62-E	1560
253	241280	Dire Dawa	09-53-97-N	41-34-39-E	1550
254	241281	Dire Dawa	09-53-39-N	41-34-68-E	1510
255	241288	Oromia	09-17-32-N	40-50-55-E	1220
256	241289	Oromia	09-11-36-N	40-39-34-E	1430
257	241689	Oromia	07-17-29-N	38-22-98-E	1490
258	241691	Oromia	07-17-29-N	38-22-99-E	1490
259	695536	Oromia	08-27-00-N	35-38-00-E	1480
260	239227/114	Amhara	10-58-17-N	39-46-40-E	1580
261	239227/305	Amhara	10-58-17-N	39-46-40-E	1580
262	Argiti	Improved			
263	B35	Improved			
264	ESH-1	Improved			
265	Melkam	Improved			
266	SB1	Tigray	12- 42- 25-	12.706966	1650.8
267	SB100	Amhara	11 53 19	11.888786	1896.1
268	SB101	Amhara	11 53 33	11.8925316	1892.8
269	SB103	Amhara	11 53 30	11.891695	1907.3
270	SB104	Amhara	11 53 33	11.8925316	1892.8
271	SB106	Amhara	11 53 33	11.8925316	1892.8
272	SB107	Amhara	11 53 32	11.8923816	1899.2

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
273	SB109	Amhara	11 53 32	11.8923816	1899.2
274	sb11	Tigray	12 45 17	12.754973	1738.9
275	SB110	Amhara	11 53 18	11.8884	1837.3
276	SB113	Amhara	10 56 23	10.9397416	1538
277	SB118	Amhara	10 53 0	10.883526	1469.5
278	SB119	Amhara	10 53 9	10.8859	1473.2
279	SB12	Tigray	12 45 17	12.754973	1738.9
280	SB120	Amhara	10 53 5	10.884765	1492.2
281	SB122	Amhara	10 53 0	10.883526	1469.5
282	SB123	Amhara	10 53 9	10.8859	1473.2
283	SB124	Amhara	10 53 8	10.885629	1490.7
284	SB125	Amhara	10 7 13	10.1202816	1156.7
285	SB127	Amhara	10 8 5	10.1348416	1193.5
286	SB128	Amhara	10 75 6	10.132415	1190
287	SB13	Tigray	12 45 23	12.75664	1748.9
288	SB130	Amhara	10 9 55	10.1650833	1271.9
289	SB14	Tigray	12 45 36	12.760205	1738.9
290	SB15	Tigray	12 45 22	12.756303	1751.4
291	SB16	Tigray	12 45 17	12.754903	1726.9
292	SB17	Tigray	12 45 36	12.760205	1738.9
293	SB18	Tigray	12 45 11	12.753066	1729.5
294	SB19	Tigray	12 45 10	12.7530283	1729.4
295	SB2	Tigray	12 42 25	12.706966	1650.8
296	SB21	Tigray	12 45 27	12.75756	1744.7
297	SB22	Tigray	12 45 36	12.760205	1738.9
298	SB23	Tigray	12 47 46	12.796205	1812.3
299	SB25	Tigray	12 48 17	12.8048383	1831.9
300	SB27	Tigray	12 48 17	12.8048383	1831.9
301	SB29	Tigray	12 48 17	12.804965	1844.8
302	SB3	Tigray	12 42 24	12.7069349	1675.7
303	sb30	Tigray	12 47 43	12.7954283	1807.5
304	SB31	Tigray	12 47 43	12.7954283	1807.5
305	SB34	Amhara	12 37 7	12.04384	1551
306	SB4	Tigray	12 42 24	12.766934939	1675.7
307	SB41	Amhara	12 23 7	12.04384	1551.8
308	SB43	Amhara	12 3 77	12.04384	1551.8
309	SB5	Tigray	12 4224	12.7069349	1675.7
310	SB53	Amhara	11 58 28	11.97503	1615.4
311	SB58	Amhara	12 6 42	12.11117516	1480.3
312	SB102	Amhara	12 6 42	12.11117516	1480.3

Appendix I continued …

S/No.	Accession	Region	Latitude	Longitude	Altitude
313	SB59	Amhara	12 6 42	12.1116816	1456.7
314	SB60	Amhara	12 6 42	12.111716	1488.7
315	SB7	Tigray	12 42 27	12.7076	1679.5
316	SB74	Amhara	11 44 33	11.7426016	1901.9
317	SB78	Amhara	11 32 50	11.47379	1832.9
318	SB79	Amhara	11 32 50	11.547379	1832.9
319	SB81	Amhara	11 32 50	11.547379	1832.9
320	SB82	Amhara	11 32 50	11.547379	1832.9
321	SB83	Amhara	11 32 50	11.547379	1832,9
322	SB84	Amhara	11 32 59	11.549955	1738.6
323	SB9	Tigray	12 42 27	12.70753	1679.6
324	SB98	Amhara	11 50 59	11.849886	1603.3

Appendix 2. Plant DNA extraction protocol for diversity array technology (DArT)

Extraction buffer stock:

- 0.35 M sorbitol
- 0.1 M Tris HCL, PH. 8.0
- 5m M EDTA, PH. 8.0

To make 500 ml:

- 31.9 g sorbitol
- 50 ml 1M Tris HCL, PH. 8.0
- 5 ml 0.5 M EDTA, PH 8.0
- Fill up to 500ml MiliQ H2O

Lysis buffer stock to make 500 ml:

- 0.2M tris HCL, PH 8.0
- 0.05M EDTA, PH. 8.0
- 2 M NaCl
- 2% CTAB

To make 500 ml:

- 100ML IM Tris HCL, PH. 8.0
- 50ml 0.5M EDTA, PH. 8.0
- 200ml 5 M NaCl
- 10 g CTAB
- Fill up to 500 ml with MilliQ H2O

Sarcosyl stock 5% (w/v):

- 0.5% (w/v) sodium disulfide
- 2% (w/v) PVP- 40 (K29-32) sigma
- Dissolved in required volume of extraction buffer stock, add same volume of lysis buffer stock and 0.4 volume of extraction (lysis) buffer stock of sarcosyl stock.

PROTOCOLS:

1. Homogenize leaf tissue in liquid nitrogen to fine powder
2. Weigh 300gmg of homogenized leaf in a 2 ml Eppendorf tube
3. Add 750ul EB and 100 ul SDS (10%)
4. Incubate at 65^0 C for 20min (in water bath)
5. Add 250 ul KAc and mix
6. Put the tubes in ice for at least 30 min (possible to take pause)
7. Centrifuge at 14000 rmp for 15min.
8. Take supernatant to new Eppendorf tube
9. Add 1 volume (same volume as supernatant) of cold isopropanol (freeze at -20 ^{0}C)
10. Centrifuge at 14000rmp for 10min and discarded (pour out) the supernatant
11. Dissolve the pellet in 250 ul TE and after the pellet has dissolved entirely (if not can be put in water bath at 65 ^{0}C to facilitate dissolution) and add 250 ul CTAB
12. Incubate at 65 ^{0}C for 15min (water bath)
13. Extraction with 1 volume of chloroform

- Add 1 volume (same volume of chloroform)
- Centrifuge at 14000 rpm for 15min

- Collect the supernatant in new Eppendorf tube (pool same samples in one tube)

14. Repeat the chloroform extraction
15. Precipitate with one volume of cold isopropanol
16. Centrifuge at 14000 rpm for 15min
17. Discard (pour out) the supernatant
18. Dilute the pellet in 100 ul TE overnight
19. Add 2.5ul RNAse (1mg/ml) in 100 ul of TE- DNA
20. Incubate the samples at 37 ^{0}C for 30min