GENETIC DIVERSITY OF SWEET SORGHUM [*Sorghum bicolor* (L.) MOENCH] ACCESSIONS FROM ETHIOPIA USING SIMPLE SEQUENCE REPEAT MARKERS

MSc Thesis

BY

MELKAMU GENET WORKU

JUNE, 2024

JIMMA, ETHIOPIA

Genetic Diversity of Sweet Sorghum [*Sorghum bicolor* (L.) Moench] Accessions from Ethiopia using Simple Sequence Repeat Markers

MSc. Thesis

By

Melkamu Genet Worku

Submitted to the School of Graduate Studies of Jimma University College of Agriculture and Veterinary Medicine Department of Horticulture and Plant Sciences in Partial Fulfillment of the Requirement for the Degree of Master of Science in Plant Biotechnology

Major Advisor: Wosene Gebreselassie (Ph.D)
Co- Advisor: Tsegaye Getahun (Ph.D)

June, 2024
Jimma, Ethiopia

CHOOL OF GRADUATE STUDIES

JIMMA UNIVERSITY

COLLEGE OF AGRICULTURE AND VETERINARY MEDICINE

MSc THESIS APPROVAL SHEET

We, the undersigned, member of the Board of Examiners of the final open defense by **Melkamu Genet** have read and evaluated his/her thesis entitled **"Genetic Diversity among Ethiopian Sweet Sorghum [Sorghum bicolor (L.) Moench] Accessions using Simple Sequence Repeat (SSRs) Markers"** and examined the candidate. This is therefore to certify that the thesis has been accepted in partial fulfillment of the requirements for the degree Master of Science in **Plant Biotechnology**

Dr Esubalew Getachew	.....	14/6/2024
Name of the Chairperson	Signature	Date
Dr Wosene Gebreselassie,	.....	14/6/24
Name of Major Advisor	Signature	Date
Dr. Temesgen Matiwos	.....	14 Jun 24
Name of the Internal Examiner	Signature	Date
Dr. Tewodros Mulualem	.....	.....
Name of the External Examiner	Signature	Date

DEDICATION

This Thesis is dedicated to my beloved mother Tarik Tebabel, my father Genet Worku, my aunt Abayensh Worku and Sister Digise Genet.

STATEMENT OF THE AUTHOR

First, I declare that this thesis is entirely my own under the supervision and guidance of my advisors. In addition, the sources of all materials used in the work are indicated. This thesis was submitted in partial fulfillment of the requirements of MSc degree at Jimma University and has been deposited at the University Library for the use of borrowers in accordance with the conditions of the library. I certify that this thesis has not been submitted to any university for a degree, diploma or certificate.

A brief quotation from the thesis is allowable without special permission provided that accurate acknowledgment of the source is made. Requests for permission to quote or reproduce all or part of this document may be made by the department chair or the dean of the graduate school. If you decide on your own initiative, the use of the resources will benefit the scholarship. However, in all other cases permission must be obtained from the author.

Name: Melkamu Genet Worku Signature: _____

Place: Jimma University, Jimma

Date of submission: May, 2024

BIBILOGRAPHICAL SKETCH

The author was born on June 06, 1995 in Benishangul-Gumuz region, Metekel zone at Dungur woreda Manbuk town from his father Genet Worku and his mother Tarik Tebabel. He attended elementary school at Belaya Elementary and Junior School from 2002-2009 and He went to Manbuk Senior, Secondary and Preparatory School from 2010-2013 for senior classes. Then, he joined Gondar University in 2014 to study biotechnology. After four years of study, he graduated with a BSC degree in biotechnology in 2018. Then, he was employed as a plant protection expert at Metekel Dangur woreda level. Finally, he joined Jimma University College of Agriculture and Veterinary in 2022 to study MSc specialization in plant biotechnology.

ACKNOWLEDGEMENT

First of all, I would like to thank God for giving me physical and mental strength during my study. I am very thankful and grateful to my major research advisor Dr. Wosene Gebreselassie for his encouragement, suggestions and moral support from the very beginning in executing the research work and in the writing of this thesis. I would like to take this opportunity to express my gratitude and appreciation to my research co-advisor Tsegaye Getahun (PhD) and Prof. Tileye Feyissa for his inspiration, moral and experimental support, guidance and suggestions during the preparation of this thesis. We appreciate your patience for close follow up of his work from the very start to finish. Without their dedication, this work would not have been completed on time.

I would like to thank the graduate students of Addis Ababa University Institute of Biotechnology, especially to Mr. Amare Genatu, Mr. Tefera Habtegiorgis and Mesfin Hailemariam for their technical unfailing guidance in each and every necessary procedure of the laboratory activities. I would like to express my deepest and heartfelt gratitude to Addis Ababa university plant genetics department of microbial, cellular and molecular biology (MCMB) staffs, especially Dr. Kassahun Tesfaye for helping me to work on his laboratory (to use Gel docs). I am extremely grateful to my mother, father and sisters for their understanding and help in using whatever resources they have. I will never forget their care for me that brought me to this success. I thank my aunt, Abaynesh Worku for her general support during my study period.

Finally, we would like to thank everyone who directly or indirectly participated in the progress of the research.

LIST OF ABBREVATIONS

AMOVA	Analysis of Molecular Variance
CTAB	Cetyl Trimethyl Ammonium Bromide
EDTA	Ethylene Diamine Tetra Acetate
FAOSTAT	Food and Agriculture Organization Statistics
He	Expected Heterozygosity
Ho	Observed Heterozygosity
MARC	Melkassa Agricultural Research Center
PCoA	Principal Coordinates Analysis
PCR	Polymerase Chain Reaction
PIC	Polymorphic Information Content
RFLP	Restriction Fragment Length Polymorphism
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
UPGMA	Unweighted Pair Group Method with Arithmetic Average

TABLE OF CONTENTS

PAGE

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF APPENDICES

Page

Genetic Diversity of Sweet Sorghum [*Sorghum bicolor* (L.) Moench] Accessions from Ethiopia using Simple Sequence Repeat Markers

Melkamu Genet Worku

June 2024, Jimma University, Jimma, Ethiopia

ABSTRACT

Sweet sorghum is a cereal crop that can grow for multiple purpose uses. Despite its global significance and potential, it faced genetic erosion, due to its low productivity and low farmers' preferences over grain sorghum, which is not considered as one of the most important cereal crops in Ethiopia. To utilize and popularize this crop, understanding the genetic diversity and population structure is a pre-request. Therefore, *this study aimed to assess the genetic diversity and population structure of 82 Ethiopian sweet sorghum accessions that represents seven geographic regions of Ethiopia using 15 simple sequence repeat markers. The study revealed a total of 116 alleles with a mean of 11.6 alleles per locus being amplified. Ten microsatellite loci were highly polymorphic with polymorphic information content (PIC) ranging from 0.75 to 0.90 and an average of 0.82. They showed high gene diversity ranged from 0.59 to 0.81 with a mean of 0.70. There was moderate genetic differentiation (FST = 0.21), showing the presence of high gene flow where 90% of the total variation was accounted for among accessions and 5% accounted for genetic variability between populations. The clustering, principal coordinate analysis (PCoA) and population structure did not cluster the studied populations into separate groups according to their geographical origin due to the presence of high gene flow (Nm = 5.033). In conclusion, from the evaluated loci the highest private alleles were observed among populations in North Wollo and Kaffa, and hence these areas can be considered hot spots for the development of improved varieties with unique traits that are well-suited to the local agricultural practices.*

Keywords: Ethiopia, Genetic Differentiation, Gene Flow, Population Structure, Sweet Sorghum

1. INTRODUCTION

Sorghum (*Sorghum bicolor* (L.) Moench) belongs to the grass family *Poaceae* (Venkateswaran *et al*., 2019). It is a diploid (2n = 20) self-pollinating crop with a genome size of 730 Mb (Ciampitti and Prasad, 2020). It is a short-growing day and a C4 crop that better tolerates diverse agroecology's. It can be easily grown in tropical, subtropical, temperate, and semi-arid regions with poor soil quality (Mathur *et al*., 2017). It has both cultivated and wild races and it has a high genetic variation for traits of agronomic importance (Hart *et al*., 2001). Sweet sorghum was originally proposed by Clayton in 1961 as cultivated *Sorghum bicolor* (L.) Moench, which is at present more selective due to the high production of biomass and sugar in stem juice (Srinivasa *et al*., 2013; López-Sandin *et al*., 2021). Like grain sorghum, it is cultivated by seeds. But it is taller than grain sorghum and known by the name *tinkish* in Ethiopia.

Sweet sorghum is both a grain and stalk crop that can be used to make alcohol, syrup, jaggery, fodder, fuel, bedding, roofing, gardens, paper, and chewing (Ratnavathi *et al*., 2011; Jemal *et al*., 2019). In addition, it produces more fermentable raw sugar under marginal conditions that are unsuitable for sugar cane or beet production (Whitfield *et al*., 2012). It accumulates these edible sugars up to 20–30% in its stalk juice (Waniska *et al*., 2004). The cultivation costs are also three times less than those of sugarcane (Reddy *et al*., 2005), making it an ideal raw material for bioenergy. In Ethiopia, sorghum is used for making injera, bread, feed, local beer production, and animal fodder, but sweet sorghum is mostly chewed for pleasure and nutrition (Disasa, 2016). Its leaves and stalks are used for animal feed, composting, and the building of houses.

According to Jiang *et al*. (2019), the global marginal land suitable for sweet sorghum cultivation is 4 .8 billion hectares. It is grown in all regions of the world and in a variety of climate conditions, making it a very adaptable and tolerant plant (Smith and Frederiksen, 2000; Srinivasa *et al*., 2013; Bruinsma, 2017). But it is widely cultivated in America, Brazil, India, China, Mexico, Sudan, Argentina, Asia, Europe, and Ethiopia (Mathur *et al*., 2017). Depending on cultivar, climate, location, and production practices, sweet sorghum yields can range from 32 to 112 Mg/ha of fresh biomass and from 15 to 25 Mg/ha of dry biomass annually (Ekefre *et al*., 2017). According to Yali and Began (2022), Ethiopia is the center of sweet sorghum origin and diversity, like grain sorghum. Grain sorghum is the third largest in area coverage next to maize and teff, and fourth in production after maize, teff and wheat (Welderufael, 2024). Its total area coverage was estimated to be around 1.66 million hectares,

whereas the estimated annual national production is about 4.2 million tons (FAOSTAT, 2022). However, the precise hectares of sweet sorghum covering the area have not been determined in Ethiopia, even if in the world.

Despite the importance of this crop, it produces low grain yields (Aruna *et al*., 2018), which discourage farmers interest in planting sweet sorghum. Moreover, sweet sorghum stems are more susceptible to insect attacks than grain sorghum due to sugar accumulation (Nasidi *et al*., 2019) and the diversity of sweet sorghum in Ethiopia is poorly understood. This leads severe yield and diversity losses. Sweet sorghum improvement is one of the alternative ways to overcome these constraints. Genetic diversity study for sweet sorghum is a critical step for improving sweet sorghum and conservation (Bouargalne *et al*., 2022). It can be determined using morphological and molecular differences existing among the populations (Enyew *et al*., 2021). Morphological assessment is a simple method and a relatively straightforward process that evaluates the physical characteristics of organisms. However, it can be labor-intensive, time-consuming, highly affected by environmental factor, subject to temporary phenotypic changes, reveals only limited polymorphism and also it is difficult to assess at any stage of plant development (Hillis, 1987; Govindarajet *et al*., 2015; Chen *et al*., 2020; Mamo *et al*., 2023). On the other hand, molecular assessment is the analysis of genetic information at the molecular level. It is unaffected by environmental factors, can be applied at any stage of plant development, and is also effective in distinguishing very closely related genotypes (Garcia *et al*., 2004; Rao *et al*., 2013). Therefore, it is more selective than the morphological method. Ritter

Earlier investigations on the sweet and grain sorghum genetic diversity were conducted using random amplified polymorphism DNA (RAPD) analysis (Ruiz-Chután *et al.*, 2019), Amplified fragment length polymorphism (AFLP) markers (Ritter *et al*., 2007), simple sequence repeat (SSR) markers (Disasa *et al*., 2016), express sequence tags (EST) SSR markers (Ramu *et al*., 2013), and SNP markers (Wondimu *et al*., 2021). From those types of markers, SSR markers are well-suited for assessing the genetic diversity of sorghum because they are reproducible, require a small DNA sample, exhibit high polymorphism, and can provide additional genetic information even for closely related cultivars, multi allelic nature, co-dominant nature, large number, and presence throughout the genome (Mangena *et al*., 2018; Mamo *et al*., 2023). In addition, it is available or less costly, which makes it more selectable.

In Ethiopia, the Ethiopian Biodiversity Institute (EBI) has conserved 11,353 sorghum accessions collected from diverse agro-ecologies across the country (Tesfaye, 2017). However, sweet sorghum is not well characterized, and is currently difficult to distinguish from the grain sorghum in EBI. Therefore, it is difficult to utilize (Disasa *et al*., 2016), and the genetic diversity of most of its accessions in the collection remains molecularly uncharacterized. The diversity of sweet sorghum is also threatened by genetic erosion because of the limited of targeted conservation efforts and increase in grain sorghum production. Farmers prefer the production of grain sorghum due to its low productivity, and sweet sorghum was not considered to be among the important crops in Ethiopia. In addition, most studies in Ethiopia have focused on grain sorghum diversity, and no intensive study of sweet sorghum germplasm collected in Ethiopia has been reported. Furthermore, the genetic relationships and diversity of sweet sorghum are poorly studied under Ethiopian condition. Therefore, this study was aimed to quantify the genetic diversity present among 82 sweet sorghum accessions collected from different parts of Ethiopia using SSR molecular markers for breeding and conservation.

Objective:

- To quantify the genetic diversity and population structure of sweet sorghum accessions collected from different parts of Ethiopia using SSR molecular markers

2. LTERATURE REVIEW

2.1. Taxonomy of Sorghum

Sorghum [*Sorghum bicolor* (L.) Moench] belongs to the grass family Poaceae (Gramineae), subfamily Panicoideae, tribe Andropogoneae, and subtrbe Sorghinae (Kowal, 2017). The genus Sorghum was thought to have 20–30 known species so far, divided into five sections: Stiposorghum, Chaetosorghum, Eusorghum, Heterosorghum and Parasorghum. Three species are known in the section *Eu-sorghum: S. propinquum* (Kunth) Hitchc, *S. halepense* (L.) Pers, and *S. bicolor* (L.) Moench (Olani *et al*., 2021). Sweet sorghum belongs to the domesticated species [*Sorghum bicolor* (L.) *Moench*], and their chromosome numbers are 2n = 2x = 20 in diploids and 2n = 4x = 40 in tetraploids (Mathur *et al*., 2017; Bednářová *et al*., 2021).

Wild Sorghum bicolor is diverse in its morphology and ecology. It could serve as a source of new genes and offer an untapped wealth of novel traits against new pests or changing abiotic conditions. In addition, S. bicolor has three subspecies, including Subsp. Bicolor (all cultivated sorghums), S. bicolor verticilliflorum (cultivated sorghum-wild stem cells), and S. bicolor drummondii (Murray, 2009; Mathur *et al*., 2017). Furthermore, based on the grain shape, glume, and panicle, cultivated varieties of Sorghum bicolor have been classified into five basic races, including bicolor, guinea, caudatum, kafir, and durra (Mathur *et al*., 2017). Except Kaffir, all other sorghum races are found in Ethiopia (Olani *et al*., 2021). Farmers in Ethiopia use several folk names to differentiate the many types of sorghum growing there. The name system is based on their physical appearance (Regasa, 2021).

2.2. Origin and Distribution of Sorghum

The wild sorghum used as food/domestication was started around 4000–3000 BC in Ethiopia and surrounding countries (Dillon *et al*., 2007). It is believed that Ethiopia is the primary and India is the secondary centre of its diversity. But it was notorious and controversial to know and point out the exact regions for sorghum domestication and cultivation (Venkateswaran *et al*., 2019). Moreover, the presence of a wide range of wild sorghum relatives in Sudan and Ethiopia is further strengthening their centers of origin (Reddy *et al*., 2006). Sorghum was distributed throughout Africa and the globe, mostly by trade routes. Human migration also helped spread sorghum from east Africa to southern and eastern Africa (Ananda *et al*., 2020). Sorghum has gained popularity and economic importance in Indian agriculture (Kleih *et al*.,

2000), leading to India being identified as the second hub for sorghum diversity (Misganaw *et al*., 2023).

Sweet sorghum has a multiphylum origin (more than one ancestor) and is known to be present in all the five major races of *Sorghum bicolor* [L.] Moench: bicolor, guinea, caudatum, kafir, and durra (Ritter *et al*., 2007; Dube and Maphosa, 2022). Due to this, its emergence as a specific variety of S. bicolor is not clear and cannot be clearly separated from grain sorghum lines. Moreover, sweet sorghum was mostly grouped in the race bicolor (Mathur *et al*., 2017). Similarly to grain sorghum, sweet sorghum originated in Africa but is more known in Ethiopia (Srinivasa *et al*., 2013). Then, it was distributed along the trade and shipping routes around the African continent and elsewhere in the world (Dicko *et al*., 2006; Mathur *et al*., 2017). The ecological distribution and major growing regions/zones of Ethiopian sorghum races are presented in Table 1.

Table 1: Ethiopian sorghum races presented according to their ecological distribution (modified from Olani *et al*., 2021).

Ethiopian sorghum races	Ecological distribution	Major growing regions/zones
DURRA	Eastern highland regions and mid-elevation terraces of the north	North Wollo, South Wollo, and East Harergie
CAUDATUM	Hot, dry valleys and lowland savannas in north and west Ethiopia	Metekel, Central and South Tigray, and Gambella
GUINEA	Rift Valley regions of Ethiopia	North Shewa
BICOLOR	Rift Valley regions of Ethiopia	West Showa

2.3. Production and Constraints of Sweet Sorghum in Ethiopia

In Ethiopia, sorghum is the third most important food cereal after maize and teff in terms of the total number of farmers, area, and grain production (Yali and Began, 2022). Sweet sorghum was cultivated in Ethiopia as early as 200 AD. However, large-scale cultivation of sweet sorghum is not common in most regions of Ethiopia. In most cases, a small amount of seed is mixed with grain sorghum or maize and planted on a limited plot of land (Gerrano *et al*., 2014; Disasa *et al*., 2017).

In Ethiopia, biotic and abiotic factors, primarily drought and a parasitic weed called Striga (Striga spp.) in the lowland climate, have hampered the yield and quality of 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 (Derese *et al*., 2018; Degebasa *et al*., 2022*)*. Moreover, the lack of study on the genetic diversity and population structure of sweet sorghum as one of the main problem in sorghum production in Ethiopia to overcome the above constraints and to make sweet sorghum as popular production by farmers.

2.4. Importance of Sweet Sorghum

Sweet sorghum Known as a "camel crop" and also known as a "smart" crop by many farmers and researchers due to its well-adapted to the arid and semi-arid tropics (Mokariya and Malam, 2020). It tolerates drought and uses water very efficiently. It is a multipurpose crop that produces food, animal feed, fodder, chewing, syrup production, and biofuel (Mathur *et al*., 2017; Roger and Willis, 2019; Dube and Maphosa, 2022). It can be an alternative feedstock for biofuels because of its high sugar content in the stalks and low input requirements without affecting food or feed security (Srinivasa *et al*., 2013; Mathur *et al*., 2017). Its grains can also be used as a gluten-free substitute for wheat or corn flour. Its whole part of the crop has economic importance (Mathur *et al*., 2017).

Sweet sorghum is a sustainable cereal crop due to its rapid growth, higher biomass production, sugar accumulation, adaptation to wider abiotic conditions, and cheaper production costs compared to other energy and cereal food sources (Reddy and Reddy, 2003). It can also be used to remove cadmium from polluted soil due to its heavy metal tolerance, offering phytoremediation potential (Xiao *et al*., 2021). Due to its versatility, adaptability, biofuel potential, economic benefits, and sustainability, it is an important crop, particularly for Sub-Saharan Africa, where it can promote resilience in rural livelihoods (Dube and Maphosa, 2022). It is mostly used for chewing in Ethiopia.

2.5. Genetic Diversity of Sweet Sorghum

Sweet sorghum is predominantly self-pollinated C4 grass and possesses extraordinary germplasm diversity (Sudhakar *et al*., 2016; Hazarika *et al*., 2021; Herniwati *et al*., 2024). But it can also accept pollen from other sweet sorghum plants (cross-pollination) by means of

wind or insect transfer in ranges from 5% to 15% with an average of about 6% (Prabhakar *et al.*, 2022).

Genetic diversity refers to the genetic variation that exists within a population or species. There is a broad diversity of landraces mainly grown by farmers and used for sorghum improvement in the region in response to different constraints. Landraces are important genetic resources and possess stress tolerance genes that can be used in sorghum breeding programs because they are adapted to harsh conditions (Disasa *et al*., 2016). The germplasm collection is a priceless genetic resource for various breeding programs. Plant breeders must study genetic diversity within and between collections and accessions to fully utilize these genetic resources in gene banks. This is important for the efficient use and allocation of genetic resources in breeding projects and to minimize the handling of duplicated or closely related accessions (Yahaya *et al*., 2023).

The wide diversity is found within and among the sorghum genotypes at both the phenotypic and genotypic levels. The genetic diversity of crop species is a gift of nature and arises due to geographical separation, genetic barriers to interbreeding, and human-mediated effects (Sinha and Kumaravadivel, 2016; Khalsa, 2013). Progress in plant breeding depends on variation because superior genotypes cannot be obviously selected from homogenous populations, but homogeneity is desirable in the end product of the agricultural variety. Success in improving adaptation requires that the selected population be genetically diverse (Disasa *et al*., 2016). The pattern of genetic variation is an important part of the collection and conservation of genetic resources and the process of crop improvement, including the selection of parents to generate new genetic recombination (Desmae *et al*., 2016).

Assessment of genetic diversity within and between plant populations is performed using various techniques, such as morphological and molecular marker analysis (Govindaraj *et al*., 2015). Any type of marker to be applied in assessing diversity among accessions will depend on the crop species, lab infrastructure and cost, technical expertise, suitability for the specific study, and the desired results (Ebraheim, 2016).

2.5.1. Morphological marker

Morphological marker analysis involves evaluating the physical characteristics of sweet sorghum plants, such as plant height, leaf shape, panicle size, and seed color, and is generally scored visually. It is the easiest and cheapest method for estimating genetic diversity

(Prajapati *et al*., 2018). There are several undesirable factors that are associated with morphological markers. The first is the high dependency on environmental factors. Often, the conditions in which a plant is grown can influence the expression of these markers and lead to false determination (Dido *et al*., 2020). Second, these mutant traits often have undesirable features such as dwarfism or albinism. And lastly, performing breeding experiments with these markers is time-consuming, labor-intensive, and the large populations of plants require large plots of land and/or greenhouse space in which to be grown (Stuber *et al*., 1999).

2.5.2. Molecular Marker

Molecular markers play an important role in the conservation and utilization of sorghum genetic resources and have been used in many sorghum breeding programs. These markers aid in the identification of distinct lines and the mapping of genetic regions that govern desirable traits, enabling marker-assisted breeding. They provide more reliable and consistent information about the genetic diversity of closely related genotypes as compared to morphological and biochemical markers. This is due to their ability to directly assess variations at the DNA level (Disasa *et al*., 2016).

There are different types of DNA markers, such as amplified fragment length polymorphism (AFLP), restriction fragment length polymorphism (RFLP), simple sequence repeats (SSRs), and single-nucleotide polymorphism (SNP). Simple sequence repeats (SSRs) are the most widely used marker for plant variety characterization and diversity analysis, especially in cultivated species that have low levels of polymorphism (Dhaliwal *et al*., 2014). They are short tandem repeat motifs, usually consisting of 1–6 bp of nucleotides. Forward primers amplify exon regions, while reverse primers amplify intron and promoter regions. Their polymorphisms result from variation in the length of their exons, introns, promoters, and spacers, both between individuals and between species (Lekgari and Dweikat, 2014). Most researchers have reported SSR markers to be very useful to analyze the structure of germplasm collection due to the following advantages: ease of detection through polymerase chain reaction (PCR), being reasonably inexpensive, abundant, codominant, multiallelic, highly polymorphic, and chromosome-specific (locus-specific), being easily detected without the use of radioisotope techniques needed with RFLPs, being highly discriminative compared to AFLPs and RFLPs, being highly reproducible, and being highly informative (Schloss *et al*., 2002; Ali *et al*., 2008; Mason, 2015).

2.6. Genetic Diversity of Grain and sweet Sorghum using SSR Markers

Due to the ease of detection through polymerase chain reaction (PCR), it's reasonably cost, and the high level of polymorphism, most researchers have adopted the SSR marker polymorphism (Mangena *et al*., 2018). Up to know a number of studies have confirmed the effectiveness of SSR for the genetic diversity of grain sorghum. The study conducted by Nemera *et al*. (2022) has shown the genetic diversity of eighty sorghum accessions by using 11 SSR markers. The result generates 70 alleles across 80 sorghum accessions with an average of 6.36 alleles per marker. The PIC ranges from 0.50 to 0.86, and the genetic diversity values range from 0.51 to 0.77. Analysis of Molecular Variance (AMOVA) revealed 93.26% of the total genetic variation within populations and 6.74% among populations. Cluster analyses did not show clear grouping of accessions according to their geographical origins, confirming gene flow (Nm =6.65) among populations. In conclusion, the SSR markers used were polymorphic and highly informative, suggesting possible target populations for breeding and conservation.

According to Regasa (2021), Mamo *et al*. (2023), and Misganaw *et al*. (2023), they studied the genetic diversity of 81,100 and 92 Ethiopian sorghum accessions by using 13, 15 and 12 microsatellite loci, respectively. All the used SSR markers were polymorphic and resulted in a total of 70, 10 and 77 alleles with a mean of 5.4, 7.2 and 0.76 alleles per locus and a PIC ranging from 0.49 to 0.78 with an overall mean value of 0.69, from 0.68 to 0.89 with a mean of 0.80, and similarly from 0.66 to 0.82 with a mean of 0.76 on the respective author's order. At the same time, Regasa (2021) reported the presence of moderate genetic differentiation (FST = 0.10) in which 90.46% of the genetic variation was accounted for within populations and only 9.54% was contributed by among populations. And also, Misganaw *et al*. (2023) reported the existence of large genetic differentiation (FST = 0.21), where 90% of the total variation was accounted for within population, leaving only 10% for among population variation. Similarly, Regasa (2021), UPGMA-based clustering, PcoA and structure analysis did not group the individuals sharply into genetically distinct clusters according to their geographical areas, indicating the presence of a higher gene flow (Nm = 2.37). Together, these results provide important insights into the importance of SSR for genetic diversity studies in sorghum improvement programs and conservation measures. Disasa *et al*. (2016) studied the genetic diversity between 202 Ethiopian sweet sorghum accessions and improved accessions from eastern and southern Africa by using 13 SSR markers. A total of 159 alleles

were identified with an average of 12.23 alleles per marker, ranging from 5 alleles to 17 alleles. The mean of the PIC value over the 13 SSR markers averaged 0.69, ranging from 0.37 to 0.85. Ethiopian accessions revealed a higher average number of alleles (9.5) than improved sweet sorghum (7.1). Ethiopian accessions were also significantly richer in private and rare alleles. The 202 accessions from both weighted neighbor joining-based clustering and hierarchical clustering were grouped into three major clusters based on geographical origin. Ethiopian accessions from the north (south Tigray and north Wollo) not only clustered separately from accessions from eastern and western central Ethiopia, but were also distinct from most of the improved genotypes.

Zheng *et al*. (2021) and Tiendrebéogo *et al*. (2022) studied the genetic diversity of 32 sweet sorghum and 34 both sorghum type accessions by using phenotypic and 43 and 15 SSR markers. Average 1.7 and 3 alleles were detected per locus, and PIC varied from 0.06 to 0.49 and 0.12 to 0.63 with a mean of 0.43, respectively). At the same time, the germplasm resources were divided into 4 groups, and group IV exhibited a wide distribution of spike type high grain yields, and a high brix rate.

Nerbewende *et al*. (2018) studied the genetic diversity of 93 Burkina Faso sweet grain sorghum accessions by using 15 polymorphic microsatellite markers. This authors reports a total of 49 alleles were detected with an average of 3 alleles per locus, a moderate Nei diversity of 0.474, a low level of heterozygosity (0.031) in the collection, and a very high Wright's fixation index (Fis) of 0.934. The accessions were organized into three genetic groups: A, B and C. Groups A and B was the farthest, with an Fst and a genetic distance of 0.37 and 0.22, respectively, whereas Groups B and C were the closest, with an Fst (genetic differentiation) of 0.279 and a genetic distance of 0.142.

According to Olweny *et al*. (2014), Lekgari and Dweikat (2014), and Mangena *et al*. (2018), they studied the genetic diversity of 86,142 and 18 sweet sorghum accessions by using 11, 33 and 25 SSR markers. A total of 86, 84 and 106 alleles were detected with a mean of 8, 2.90 and 4.46 per locus, respectively. The mean PIC values of SSR markers were 0.53, 0.52 and 0.56, ranging from 0.09 to 0.89, 0.22 to 0.75 and 0.00 to 0.39, respectively. Neighbor joining clusters of 142 sweet sorghum accessions into five major groups using SSR based on their origin (Lekgari and Dweikat, 2014). Clustering analysis of sweet sorghum genotypes into two distinct clusters based on the genetic similarity (GS) group (Olweny *et al*., 2014) and cluster

analysis using UWPGM grouped the 18 genotypes into three distinct clusters (Mangena *et al*., 2018).

Wang *et al*. (2009), Ali *et al*. (2008) and Nerbewende *et al*. (2018) studied the genetic diversity of 96, 68 and 93 sweet sorghum accessions by using 95, 41 and 15 SSR markers. A total of 705, 132 and 49 alleles were detected with an average of 7.6, 3.22 and 3 alleles per locus with the mean PIC value was 0.54, 0.40 and, respectively. The sorghums were generally clustered by geographic origin using both population structure and nMDS methods into four branches (Wang *et al*. (2009)).

In addition, Ali *et al*. (2008), infers that the genetic similarity coefficient was estimated based on the segregation alleles with clustering analysis based on genetic similarity (GS) grouped the 72 sorghum accessions into 10 distinct clusters.

3. MATERIALS AND METHODS

3.1. Plant Material

A total of 82 sweet sorghum accessions were used in this study. From the total 82 accessions, six were cultivars released in 2019 and obtained from Melkassa Agricultural Research Center (MARC). The rest were collected from different sweet sorghum growing areas of Ethiopian farmers field (Table 2). They were distinguished from grain sorghum based on the sweetness of the stalk. Based on their source of geographic origin, they were categorized into seven populations, such as: North-Wollo, South-Wollo, Oromo special zone, Melkassa, Asossa, Metekel and Kaffa populations.

Table 2: Lists of Sweet sorghum accessions used in diversity study, collection areas and regions

Entry	Accession Code	Collection year	District	Geographic zone	Region
1	G001	2023	Gubalafito	North Wollo	Amhara
2	K002	2023	Kobo	North Wollo	Amhara
3	G003	2023	Gubalafito	North Wollo	Amhara
4	H004	2023	Habiru	North Wollo	Amhara
5	K006	2023	Kobo	North Wollo	Amhara
6	H007	2023	Habiru	North Wollo	Amhara
7	HOO8	2023	Habiru	North Wollo	Amhara
8	G0010	2023	Gubalafito	North Wollo	Amhara
9	K0012	2023	Kobo	North Wollo	Amhara
10	K0013	2023	Kobo	North Wollo	Amhara
11	G0074	2023	Gubalafito	North Wollo	Amhara
12	H0015	2023	Habiru	North Wollo	Amhara
13	K0016	2023	Kobo	North Wollo	Amhara
14	H0076	2023	Habiru	North Wollo	Amhara
15	RK0019	2023	Kobo	North Wollo	Amhara
16	K0022	2023	Kobo	North Wollo	Amhara
17	H0023	2023	Habiru	North Wollo	Amhara
18	KOO24	2023	Kobo	North Wollo	Amhara
19	GOO52	2023	Gubalafito	North Wollo	Amhara

20	G0033	2023	Gubalafito	North Wollo	Amhara
21	K0072	2023	Kobo	North Wollo	Amhara
22	H0034	2023	Habiru	North Wollo	Amhara
23	H005	2023	Habiru	North Wollo	Amhara
24	G0040	2023	Gubalafito	North Wollo	Amhara
25	H0042	2023	Habiru	North Wollo	Amhara
26	K0044	2023	Kobo	North Wollo	Amhara
27	G0047	2023	Gubalafito	North Wollo	Amhara
28	K0048	2023	Kobo	North Wollo	Amhara
29	H0053	2023	Habiru	North Wollo	Amhara
30	H0055	2023	Habiru	North Wollo	Amhara
31	K0056	2023	Kobo	North Wollo	Amhara
32	G0059	2023	Gubalafito	North Wollo	Amhara
33	H0062	2023	Habiru	North Wollo	Amhara
34	G0064	2023	Gubalafito	North Wollo	Amhara
35	K009	2023	Kalu	South Wollo	Amhara
36	K0011	2023	Kalu	South Wollo	Amhara
37	K0017	2023	Kalu	South Wollo	Amhara
38	K0020	2023	Kalu	South Wollo	Amhara
39	AM0025	2023	Ambasel	South Wollo	Amhara
40	AM0073	2023	Ambasel	South Wollo	Amhara
41	K0031	2023	Kalu	South Wollo	Amhara
42	K0032	2023	Kalu	South Wollo	Amhara
43	K0068	2023	Kalu	South Wollo	Amhara
44	H0035	2023	Kalu	South Wollo	Amhara
45	K0043	2023	Kalu	South Wollo	Amhara
46	AM0045	2023	Ambasel	South Wollo	Amhara
47	AM0046	2023	Ambasel	South Wollo	Amhara
48	K0049	2023	Kalu	South Wollo	Amhara
49	AM0051	2023	Ambasel	South Wollo	Amhara
50	K0054	2023	Kalu	South Wollo	Amhara
51	AM0063	2023	Ambasel	South Wollo	Amhara
52	AM0066	2023	Ambasel	South Wollo	Amhara

53	D0067	2023	Dawachefa	Oromo special zone	Amhara
54	Da0021	2023	Dawachefa	Oromo special zone	Amhara
55	Da 0028	2023	Dawachefa	Oromo special zone	Amhara
56	Da0050	2023	Dawachefa	Oromo special zone	Amhara
57	Da0057	2023	Dawachefa	Oromo special zone	Amhara
58	D0058	2023	Dawachefa	Oromo special zone	Amhara
59	Da0060	2023	Dawachefa	Oromo special zone	Amhara
60	Da0065	2023	Dawachefa	Oromo special zone	Amhara
61	A-2267-2	2022	Melkassa	Melkassa	Oromia
62	Chelenko	2022	Melkassa	Melkassa	Oromia
63	Gamballa 1107	2022	Melkassa	Melkassa	Oromia
64	chiro	2022	Melkassa	Melkassa	Oromia
65	Tilhun	2022	Melkassa	Melkassa	Oromia
66	NTJ-2	2022	Melkassa	Melkassa	Oromia
67	AA0069	2023	Abirham	Asossa	Benishangul-Gumuz
68	AA001	2023	Abirham	Asossa	Benishangul-Gumuz
69	AB002	2023	Bambasi	Asossa	Benishangul-Gumuz
70	AS003	2023	saleg	Asossa	Benishangul-Gumuz
71	AS004	2023	saleg	Asossa	Benishangul-Gumuz
72	MD005	2023	Dangur	Metekel	Benishangul-Gumuz
73	MD006	2023	Dangur	Metekel	Benishangul-Gumuz
74	MD007	2023	Dangur	Metekel	Benishangul-Gumuz
75	MP008	2023	Pawi	Metekel	Benishangul-Gumuz
76	MD009	2023	Dangur	Metekel	Benishangul-Gumuz
77	MD0010	2023	Dangur	Metekel	Benishangul-Gumuz
78	B0070	2023	Bonga	Kaffa	South West Ethiopia
79	B0071	2023	Bonga	Kaffa	South West Ethiopia
80	B0011	2023	Bonga	Kaffa	South West Ethiopia
81	B0012	2023	Bonga	Kaffa	South West Ethiopia
82	B0013	2023	Bonga	Kaffa	South West Ethiopia

3.2. Genomic DNA Extraction

Ten seedlings from each sweet sorghum accession were grown in a pot at the greenhouse of Addis Ababa University. From each accession, healthy young 100 mg leaf samples were collected from two to three week-old plants. The collected leaves were rinsed with distilled water, blot dried, cut into 2cm length and placed in sterile mortar. It was quickly ground using a mortar and pestle in the presence of liquid nitrogen and transferred into a sterile 2ml Eppendorf tube using a sterile spatula. Then, the genomic DNA was extracted following cetyltrimethyl ammonium bromide (CTAB) protocol (Doyle and Dyle, 1987), with some modifications (Mekonnen *et al.*, 2017). The 700µl pre-warmed (65°C) mix of 2% CTAB extraction buffer with 0.2 vol% β-mercaptoethanol was added in sample containing tube and incubated in a water bath at 65°C for 40 min. The samples were then centrifuged at 12000 rpm for 10 minutes at room temperature and supernatant was transferred into a new 1.5ml centrifuge tube. Then 800µl of chloroform/iso-amayl alcohol (24:1) was added and shake vigorously and centrifuged at 12000 rpm for 10 min 26°C. After that, the 600µl of the supernatant was transferred into a new 1.5ml centrifuge tube and 60µl of 3M sodium acetate and 600µl of ice-cold isopropanol were added, then mixed by inverting the tubes 3-5 times and incubated at -20°C refrigerator for 1 hour. Then the solution was centrifuged at 12,000 rpm for 10 min to pellet down DNA at 4°C. The Supernatant was discarded and the pellet was washed with 1ml of cold 75% ethanol with subsequent centrifugation at 12,000 rpm for 10 min 4^{o}c. The supernatant was discarded and the pellet was washed again with 1ml of absolute (100%) ethanol and centrifuged at 12000 rpm for 10 minutes 4°C.These alcohols were discarded and the pellet was air dried for 1:00 to 1:30 hours (inclined the tubes on clean tissue paper) and then the pellet was suspended in 100µl of 1X TE buffer for 1 hour at room temperature.

3.3. Measuring of Genomic DNA Quality and Quantity

The genomic DNA quality and quantity were evaluated on a 1% agarose gel by electrophoresis at 100 voltages (which was primarily mixed about 3 µl of DNA with 2 µl of loading dye) and with a Nanodrop spectrophotometer, respectively. The A260/A280 ratio was used to provide an estimate of DNA purity. The DNA sample with high band intensity and purity with 1.8 was used for further PCR analysis. Prior to PCR the genomic DNA was normalized by diluting the DNA concentration into 100ng/µl.

3.4. Selection of Microsatellite Markers

Fifteen simple sequence repeats (SSRs) used in this study, which were previously mapped by Tirfessa *et al*. (2020), were selected based on their polymorphism and consistent amplification across accessions and genome coverage observed. The list of SSR markers, including primer sequences, information on repeat motif and length are given in Table 3.

3.5. Polymerase Chain Reaction and Fragment Analyses

All primers of annealing temperature were optimized by using gradient PCR. PCR amplification was performed in 11 µl reaction containing 4 µl of Master Mix (supplemented with all PCR reaction components such as MgCl2, PCR buffer, dNTPs, and TaqDNA polymerase, loading dye), 1 µl DNA, 0.5 µl of each of forward and reverse primer and 5 µl of double distilled water. The reaction was carried out using Prima-96™ Thermal Cycler (Applied Biosystems, USA) programmed at an initial denaturation of 94 °C for 4 minutes, followed by denaturation at 94 °C for 30 seconds, annealing at 54 °C up to 60°C (depending on the primer used) for 30 seconds, initial extension at 72 °C for 30 seconds and final elongation at 72 °C for 10 minutes. After amplification, the 100 DNA marker ladder with a known reference band was mixed with 2µl of loading dye loaded in the first well and the PCR product loaded in each remaining well of the PCR plate of 2% (w/v) agarose gel and electrophoresis was run for three hours at 100 volts. The DNA was documented and viewed with a UV transilluminator (Bio-Doc) after staining with ethidium bromide.

Table 3: SSR markers used for study of genetic diversity in Ethiopian sweet sorghum

No	Marker names	Repeat motif	Forward primer			Reverse primer			Expected band size(bp)
			sequence	TM^0	GC%	sequence	TM^0	GC%	
1	Xcup53	(TTTA)5	GCAGGAGTATAGGCAGAGGC	62.8	60	CGACATGACAAGCTCAAACG	64.5	50	182-202
2	Xtxp015	(TC)16	CACAAACACTAGTGCCTTATC	56.4	42.8	CATAGACACCTAGGCCATC	57	52.6	197-273
3	Xisep0310	(CCAAT)4	TGCCTTGTGCCTTGTTTATCT	63.4	42.8	GGATCGATGCCTATCTCGTC	63.5	55	158-219
4	Xcup61	(CAG)7	TTAGCATGTCCACCACAACC	63.5	50	AAAGCAACTCGTCTGATCCC	63.2	50	189-204
5	Xtxp012	(CT)22	AGATCTGGCGGCAACG	63.9	62.5	AGTCACCCATCGATCATC	58.3	50	143-215
6	Xtxp321	(GT)4+(AT)6+(CT)21	TAACCCAAGCCTGAGCATAAGA	64.6	45.4	CCCATTCACACATGAGACGAG	65.6	52.3	180-252
7	Xtxp057	(GT)21	GGAACTTTTGACGGGTAGTGC	64.5	52.3	CGATCGTGATGTCCCAATC	64.1	52.6	213-285
8	Xtxp265	(GAA)19	GTCTACAGGCGTGCAAATAAAA	62.8	40.9	TTACCATGCTACCCCTAAAAGTGG	65.6	45.8	163-246
9	Xtxp12	(CT)22	AGATCTGGCGGCAACG	63.9	62.5	AGTCACCCATCGATCATC	58.3	50	159-234
10	Xtxp3	(CT)8+(CT)36	AGCAGGCGTTTATGGAAG	60.4	50	ATCCTCATACTGCAGGAC C	59.1	52.6	170-236
11	Xtxp114	(AGG)8	CGTCTTCTACCGCGTCCT	63.1	61.1	CATAATCCCACTCAACAATCC	60.7	42.8	196-245
12	Xtxp145	(AG)22	GTTCCTCCTGCCATTACT	56.5	50	CTTCCGCACATCCAC	56.6	60	148-243
13	Xtxp273	(TTG)20	GTACCCATTTAAATTGTTTGCAGTAG	62.4	34.6	CAGAGGAGGAGGAAGAGAAGG	63.3	57.1	148-243
14	Xtxp320	(AAG)20	TAAACTAGACCATATACTGCCATGATAA	61.6	32.1	GTGCAAATAAGGGCTAGAGTGTT	60.3	43.4	180-252
15	mSbCIR262	(CATG)3.	GCACCAAAATCAGCGTCT	61.5	50	CCATTTACCCGTGGATTAGT	60.1	45	208-446

NB: TM^0, melting temperature in centigrade

3.6. Scoring

Amplified products will be scored based on fragment band size compared with 100 DNA ladders (Wang and Hiruki 2001) by using Image Lab 6.1 software (https://www.bio-rad.com/en-et/product/image-lab-software?ID=KRE6P5E8Z). Fragments with the same mobility were considered as identical fragments and treated as a unit character. DNA-fragments that result in different gel patterns between samples or individuals are called polymorphic markers. The visible differences on the gel result from differences at the DNA level.

3.7. Diversity Parameters and Analysis

Analyses of genetic diversity and population structure were performed based on scored bands. Locus based diversity indices including the number of alleles (Na), major allele frequency (MAF) and polymorphic information contents (PIC) were analyzed using Power marker version 3.25 (Liu and Muse, 2005). Effective number of alleles (Ne), Shannon's information index (I), population differentiation (Fst), gene flow (Nm), observed heterozygosity (Ho) and expected heterozygosity (He) were analyzed using GenAlex version 6.5 (Peakall and Smouse, 2015).

3.7.1. Number of observed alleles

Observed number of alleles is a measure of genetic diversity that simply counts the total of different alleles present at a specific locus within a population. It may be affected by the size of the sample. With a larger sample size, there is a higher chance of observing a greater number of alleles. The average number of alleles per locus is the sum of all detected alleles in all loci, divided by the total number of loci. It was calculated as:

$$n = (1/k)\sum_{i=1}^{k} ni$$

Where,

K = the number of loci

ni = the number of alleles detected per locus (De Vicente *et al*., 2004).

3.7.2. Effective number of alleles

The effective number of alleles (Ne) is a crucial parameter to measure genetic diversity. It represents the number of equally frequent alleles required to achieve a specific level of gene

diversity. It allows for the comparison of populations with varying numbers and distributions of alleles. The formula to calculate the effective number of alleles is:

$$Ne = 1/(1 - He)$$

Where, Ne is the effective number of alleles and He is expected heterozygosity or gene diversity (De Vicente *et al*., 2004).

The effective number of alleles (Ne) with high values indicates a high level of genetic diversity, and low values suggest a low level of diversity. The mean Ne is affected by various factors, such as the genetic markers used, the population under study, and the level of genetic diversity present in the samples analyzed (De Vicente *et al*., 2004; Kanaka *et al*., 2023).

3.7.3. Heterozygosity

Heterozygosity refers to the condition of having two different alleles at a locus. It is a common way of measuring genetic diversity (Allendorf, 2014).

The expected heterozygosity (He) or gene diversity (D) is the probability that an individual will be heterozygous at a given locus. It ranges from 0 to 1. It can be calculated as:

Gene diversity (excepted heterozygosity) was estimated using the following formula (Höglund, 2009):

$$He = 1 - \sum (pi)^2$$

Where, pi is the frequency of the i^{th} allele at a locus.

The observed heterozygosity (Ho) is the fraction of heterozygous individuals in a population. It is calculated as the number of homozygotes divided by the number of individuals without missing data. It is represented as:

$$Ho = Nnumber\ of\ \frac{homozygotes}{numbers\ of\ individuals}$$

Where, Ho is observed heterozygosity. He is less susceptible to sample size than observed heterozygosity, and it is estimated to characterize genetic diversity in crops (De Vicente *et al*., 2004).

3.7.4. Wright's F-statistics

F-statistics, also known as fixation indices, are widely used measures in population genetics to characterize population structure in natural populations. F-statistics provide information

about the distribution of genetic variation within and among populations. They quantify the degree of genetic differentiation or subdivision among populations, as well as the extent of inbreeding or relatedness within populations (Wright, 1951).

There are several F-statistics, including FST, FIS, and FIT, which are commonly used in population genetics studies. (1) FST measures the proportion of genetic variation that is due to differences among populations relative to the total genetic variation. It indicates the degree of genetic differentiation among populations. A high FST value suggests significant genetic differentiation, while a low value indicates genetic similarity among populations. (2) The inbreeding coefficient (Fis) is a measure that represents the probability that two alleles, randomly selected from an individual, are identical by descent (IBD) relative to a base population. It quantifies the level of inbreeding within a population or an individual. A higher inbreeding coefficient indicates a higher probability of inheriting identical alleles from common ancestors, which can increase the risk of deleterious traits and homozygosity but reduce genetic diversity (Zhang *et al*., 2015; Villanueva *et al*., 2021; Kanaka *et al*., 2023). (3) FIT (Global Inbreeding Coefficient): FIT represents the overall levels of inbreeding in a population, taking into account both within-population inbreeding (FIS) and among-population differentiation (FST). It measures the reduction in heterozygosity compared to the expected heterozygosity under random mating.

According to Wright (1951), the F statistics parameters were performed as follows:

$$Fst = (FIT - Fis)/(1 - Fis)$$

Where,

Fst is the fixation index that describes the correlation of genes among different individuals in the same population.

Fis is the inbreeding coefficient, which describes the correlation of genes within individuals in the population.

FIT is the overall inbreeding coefficient, describing the correlation of genes within individuals relative to the total population. If Fis is 0, it implies there is neither inbreeding nor outbreeding, i.e., random mating exists, but 1 indicates that it is completely inbred, and FIS = − 1 indicates completely outbred. Fst = 0 indicates no structure, thus no differentiation of the population, while FST = 1 indicates that it is complete differentiation of the population (Wright, 1965). The Fst value between 0 and 0.05 implies the existence of small genetic differentiation; 0.05 to 0.15 indicates the presence of moderate genetic differentiation; and 0.15 to 0.25 indicates the presence of large genetic differentiation (Wright, 1951).

3.7.5. Shannon's information indices

Shannon information index is a measure of diversity that quantifies the uncertainty in predicting the species identity of an individual taken at random from a dataset. It is calculated as follows:

$$I = \sum pi\ log2\ pi$$

Where,

I is Shannon information index, and

pi is the frequency of a marker. As the Shannon information index value increases, the diversity also increases. But it depends on the sample size (Kanaka *et al*., 2023).

3.7.6. Gene flow

Gene flow (Nm) is a genetic diversity parameter that represents the amount of genetic exchange between populations. It is calculated by multiplying the effective population size (Ne) by the rate of migration (m). The effective population size refers to the size of an idealized population that experiences the same amount of genetic drift as the actual population under consideration. The rate of migration represents the movement of individuals between populations, which can facilitate gene flow. Directly measuring the rate of gene flow by tracking individuals can be challenging. Instead, often gene flow is estimated indirectly by assessing the level of genetic differentiation among populations using F-statistics, such as FST. Gene flow can be calculated as:

$$Nm = (1/Fst - 1)$$

Where,

Nm is gene flow.

Fst is measures the genetic variation between populations relative to the total genetic variation within populations. It takes into account both genetic diversity within populations and genetic differences between populations (McDermott and McDonald, 1993). A lower Fst value indicates higher gene flow, while a higher Fst value suggests lower gene flow. Generally, higher levels of gene flow between sub-populations indicate a greater exchange of genetic material, suggesting a lower level of population isolation and higher dispersion

capacity. Conversely, lower levels of gene flow suggest limited genetic exchange, indicating a higher level of population isolation and reduced dispersion capacity (Kanaka *et al*., 2023). According to Slatkin (1985) and Waples (1987), the Nm values are between 0.000–0.249 low, 0.25–0.99 intermediate, and Nm > 1.00 high.

3.7.7. Polymorphic information content

Polymorphic information content (PIC) measures the ability of a marker to detect polymorphisms and is used to select markers for genetic studies (Serrote *et al*., 2020). It was calculated using the method of Botstein *et al*. (1980) with a formula:

$$PIC = 1 - \left(\sum_{i=1}^{K} p^2\right) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2pi^2 2j^2$$

Where,

PIC is polymorphic information content.

n is the number of alleles.

pi and pj are the frequencies of alleles i and j, respectively.

PIC values range from 0 (monomorphic) to 1 (very highly informative). The markers with PIC values greater than 0.5 are very informative, between 0.25-0.50 are informative, and <0.25 are uninformative.

3.8. Analysis of Molecular Variance

Analysis of molecular variance (AMOVA) is a statistical method used to study molecular variation within and among populations (De Vicente *et al*., 2004). The AMOVA approach compares the genetic variation observed within and between populations to determine the proportion of genetic variation that can be attributed to different factors, such as geographic distance, migration rates, or environmental factors. It is performed by calculating pairwise genetic distances or similarities between individuals or populations and then using statistical methods (De Vicente *et al*., 2004; Kanaka *et al*., 2023). AMOVA uses estimated F-statistics such as genetic differentiation (Fst), fixation index or an overall fixation index (Fis) to compare the genetic structure among and within populations. The AMOVA procedures were done using GeneAlex version 6.5 (Peakall and Smouse, 2015) software package.

3.9. Genetic Distance and clusters

Genetic distance is the basic measure of genetic diversity. It measures genetic differences (allelic variation) between two populations or closely related species (Mohammadi and Prasanna, 2003). It is often expressed as a dissimilarity measure because it captures the differences or discrepancies in the genetic makeup of populations or individuals. It takes into account the changes that have occurred over time and due to various evolutionary processes such as mutation, migration, and genetic drift. If the distance between two individuals is scored as 0, it indicates there is a match, and if it is scored as 1, there is a mismatch between the individuals. A large genetic distance indicates a distant relationship, whereas a small genetic distance indicates a close relationship (Beaumont *et al*., 1998; Mohammadi and Prasanna, 2003).The commonly used distance metrics are Rogers' distance, Bhattacharyya's distance, FST distance, and Nei's distance. But Nei's distance is mostly used as a standard genetic distance, it takes account of the effects of mutation and genetic drift, and its estimated value is proportional to evolutionary time. It is given as:

$$D = \frac{-\ lnJ_{XY}}{\sqrt{J_X J_Y}}$$

Where,

x and y are two populations and J_x and J_y are the mean of allele frequencies in these populations (Kanaka *et al*., 2023).

Genetic diversity can be assessed by clustering individuals or populations based on their genetic profiles. There are several methods commonly used for clustering genetic data, including: neighbor-joining (NJ) or UPGMA clustering, principal coordinate analysis (PCoA), and population structure.

3.9.1. Neighbor-joining (NJ) and UPGMA clustering

The NJ and UPGMA clustering methods are commonly used for genetic diversity analysis. These methods are based on constructing phylogenetic trees or dendrograms, using genetic distances between individuals or populations. UPGMA is the technique of constructing a rooted phylogenetic tree while neighbor joining tree is the technique of constructing an unrooted phylogenetic tree (Dega and Ercal, 2015).The resulting distance matrix is clustered

by the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) based Neighbor-Joining tree implemented in PowerMarker v3.25 and visualized in MEGA11.

3.9.2. Principal Coordinate Analysis (PCoA)

Principal Coordinate Analysis (PCoA) is a technique used for visualizing and studying the similarity or difference of data by mapping the relative similarities or differences between samples onto a two-dimensional plane. It projects the distances among samples onto a set of coordinate axes, selecting the first two axes that best preserve the original distribution of distances for data representation. PCoA is based on distances other than Euclidean distance and finds the potential principal components of the overall difference through dimensionality reduction. It is a non-binding data dimensionality reduction analysis method that can be used to study the similarity or difference of sample composition and observe the differences between individuals or groups (Morrissey *et al*., 2017). It was performed using the genetic distance indices to confirm the complementarity of the clustering pattern revealed by the dendrogram using the software package GenAlex version 6.5 (Peakall and Smouse, 2015).

3.9.3. Population structure

Population structure refers to the non-random distribution of genotypes among individuals within a population, i.e., mating does not occur randomly in a structured population. This is due to the unequal distribution of alleles between different ancestral subpopulations. Correction of population structure is critical in genotype collection, especially in a breeding program, where genetic relationships among breeding lines are highly variable (Sukumaran and Yu, 2014). The population structure of S. bicolor was studied using SSR (Wang *et al*., 2009; Billot *et al*., 2013, Mamo *et al*., 2023). It was computed using STRUCTURE 2.1 software program, an admixture model with correlated allele frequencies was used (Pritchard et al., 2000). Estimated proportions of each individual's genotype originating from each of the K ancestral populations are set between 1 and 12 ancestral populations, with fifteen runs/ iterations for each value [50,000 burn-in periods and 10,000 Monte Carlo Markov Chain (MCMC)] (Disasa *et al*., 2016). The optimal K value was calculated using the online STRUCTURE HARVESTER ver. 0.6.92 http://taylor0.biology.ucla.edu/structureHarvester (Earl and vonHoldt, 2012).

4. RESULTS AND DISCUSSION

4.1. SSR Markers Polymorphism and Level of Diversity

Among the 15 simple sequence repeat markers, five (Xtxp114, Xtxp145, Xtxp273, Xtxp320, and mSbCIR262) showed a monomorphic pattern and were hence excluded from the analysis, while the rest of the 10 simple sequence repeat markers showed a polymorphic pattern and were retained in the analysis. The polymorphic markers revealed a total of 116 alleles (Table 4). The numbers of alleles per locus generated by each marker varied from eight (Xcup53 and Xtxp015) to seventeen (Xtxp012) with a mean of 11.6 alleles. The mean number of alleles per SSR locus (Na = 11.6) detected on 82 sweet sorghum accessions was similar to the report of Disasa *et al*. (2016) on 202 sweet sorghum by using 13 SSR markers (12.23 alleles per marker). The levels observed in this investigation are far below those observed in sweet sorghum by Zhang *et al*. (2011) and sorghum by Cuevas and Prom (2013), who reported a mean of 15 and 14 alleles per locus, respectively. But our study has greater mean numbers of alleles than Wang *et al*. (2009); Pei *et al*. (2010); Lekgari and Dweikat (2014); Nerbewende *et al*. (2018), who reported 7.6, 2.90, 4.96 and 3, respectively on same crop. These variations could be due to the diverse nature of the sweet sorghum accessions, the number of accessions used in the study, the level of difference in the polymorphism and numbers of the SSR markers used.

The number of effective alleles (Ne) detected varied from 2.91 to 5.67 with a mean of 4.02 (Table 4). Markers Xtxp012 and Xtxp015 showed the highest and the lowest number of effective alleles, respectively. A higher number of effective alleles indicate greater genetic diversity within the population at that locus. A possible explanation for this might be that these multiple alleles are contributing to the genetic variation, with a more even distribution across different alleles. This can be beneficial for the long-term survival and adaptability of a population, as it provides a larger pool of genetic variation for natural selection to act upon, but a lower number of effective alleles indicates reduced genetic diversity at that locus. It suggests that a smaller subset of alleles is contributing to the observed genetic variation, potentially due to genetic drift, selection, or other factors limiting the introduction or maintenance of new alleles within the population (De Vicente *et al*., 2004; Kanaka *et al*., 2023).

Sweet sorghum is a self-pollinating species that is expected to show a high level of homozygosity but a low level of heterozygosity. The results, as shown in Table 4, indicate that the overall observed heterozygosity (Ho) over loci varied from 0.00 (Xtxp015, Xtxp057, and Xtxp3) to 0.12 (Xtxp321) with an average of 0.05. The mean observed heterozygosity of 82 sweet sorghum accessions was 0.05, signifying that 95% of the loci attained an acceptable level of homozygosity (Mangena *et al*., 2018). However, the gene diversity index (expected heterozygosity, He) ranged from 0.59 (Xtxp015) to 0.81 (Xtxp012) with a mean of 0.70. The mean value of He (0.70) in these studies indicates the presence of high allelic variation in the marker loci and their distribution in the sweet sorghum. The gene diversity or expected heterozygosity observed in this study was consistent with that of Regasa (2021) in sorghum, who reported that the He value ranged from 0.61 to 0.82 with a mean of 0.73. In accordance with the present results, previous studies have demonstrated that Pei *et al*. (2010) and Disasa *et al*. (2016) reported similar average values of expected heterozygosity (0.73 and 0.72, respectively). Moreover, the average gene diversity of this study was greater than most of the previously reported values of sweet sorghum by Ali *et al*. (2008), Olweny *et al*. (2014), and Mangena *et al*. (2018), who reported 0.58, 0.4 and 0.56, respectively. However, these comparisons were not fair as the number of samples and the sampling strategy were different.

Inbreeding coefficient (Fis) varied from 0.84 to 1.00. The overview of the inbreeding coefficient was observed highest value (1) from markers Xtxp015, Xtxp057 and Xtxp3, and the lower value (0.84) from marker Xtxp321with a mean value of 0.94. The presence of a high degree of inbreeding *coefficient* of alleles with low heterozygosity is indicated as a result of successive selfing and the presence of appreciable levels of homozygosity among the accessions that were used in the present study (Zhang *et al*., 2015; Villanueva *et al*., 2021; Kanaka *et al*., 2023).

Another important finding is that the Shannon’s information (I) index ranged from 1.04 to 1.78. The highest value of Shannon’s information (I) index was recorded for marker Xtxp012, whereas the lowest value of Shannon's information index was noted for marker Xtxp015 with a mean value of 1.40. A higher value of Shannon's information index indicates a greater diversity of alleles within the population. A value closer to 1.78 suggests a higher level of genetic diversity, meaning that the population contains a wide range of different alleles at the studied loci. These results have implications for a higher degree of variation and potential for adaptation within the population, but the lower value (1.04) suggests reduced genetic diversity.

Table 4 provides an overview of the results of the PIC values for the SSR loci, which ranged from 0.75 to 0.90 with an average of 0.82 (Table 4). The highest PIC value of 0.90 was obtained for marker Xtxp012, whereas the lowest PIC value of 0.75 was recorded for marker Xtxp3. According to Botstein *et al.* (1980), all of the SSR markers used in the present study revealed PIC value > 0.5 imply markers had a strong discriminative and informative potential for detecting genetic diversity in sweet sorghum. These markers are highly applicable in the genetic diversity study of sweet sorghum in the future. Over all, marker Xtxp012 showed the highest Na, Ne, He and PIC values, implying more significant and high genetic diversity in the present study accessions at Xtxp012 locus. The average PIC value observed in the present study was comparable similar to recent study on sorghum by Mamo *et al.* (2023), who reported average PIC value of 0.8 in 100 sorghum using 15 SSR markers; Adugna *et al.* (2013) who reported an average PIC value of 0.81 and Ceuvas and Prom, (2013), who reported 0.78. Overall, the average PIC value of the present study (0.82) was higher than the previously reported values using both grain and sweet sorghum (Agrama and Tuinstra, 2003; Caniato *et al*., 2007; Ali *et al*., 2008; Wang *et al*., 2009; Ramu *et al*., 2013; Lekgari and Dweikat, 2014; Zheng *et al*., 2021 and Tiendrebéogo *et al*., 2022). This difference could be due to the selection of molecular markers, the number and type of the study populations, and the statistical methods used to analyze the PIC values.

Table 4: Genetic parameters of the 10 SSR markers used in the study of 82 sweet sorghum accessions obtained from seven sources (populations

No	locus	Genetic diversity parameters							
		MAF	**Na**	**Ne**	**He**	**Ho**	**I**	**Fis**	**PIC**
1	Xcup53	0.32	8	2.96	0.60	0.03	1.10	0.95	0.76
2	Xtxp015	0.28	8	2.91	0.59	0.00	1.04	1.00	0.79
3	Xisep0310	0.15	14	4.07	0.70	0.05	1.40	0.93	0.88
4	Xcup61	0.31	10	4.05	0.74	0.10	1.49	0.86	0.80
5	Xtxp012	0.16	17	5.67	0.81	0.07	1.78	0.92	0.90
6	Xtxp321	0.33	10	3.57	0.71	0.12	1.37	0.84	0.80
7	Xtxp057	0.24	9	3.50	0.66	0.00	1.27	1.00	0.80
8	Xtxp265	0.17	13	4.93	0.77	0.03	1.64	0.96	0.88
9	Xtxp12	0.20	13	4.65	0.74	0.05	1.55	0.94	0.87
10	Xtxp3	0.39	14	3.86	0.72	0.00	1.40	1.00	0.75
Mean		0.25	11.6	4.02	0.70	0.05	1.40	0.94	0.82

MAF, major allele frequency; Na, number of observed alleles; Ne, number of effective alleles; I, Shannon's information index or Shannon's diversity index; Ho, observed heterozygosity; He, expected heterozygosity or gene diversity; FIS, inbreeding coefficient; PIC, polymorphic information content (%)

4.2. Genetic Diversity within the Populations

A summary of the different genetic diversity estimates over all the loci across populations is presented in Table 5. The overall genetic diversity estimates within the populations showed a mean number of alleles (Na) of 5.16 with a range of 3.6 to 9, a mean number of effective alleles (Ne) of 4.02 with a range of 2.97 to 5.75, a mean number of private alleles (number of alleles unique to a single population) (NPA) of 0.39 with a range of 0.1 to 0.8, a mean Shannon's information index (I) of 1.40 ranging from 1.10 to 1.88 and a gene diversity (He) value of 0.70 ranging from 0.61 to 0.81. The population of North Wollo showed the highest Na, Ne, NPA, I, and genetic diversity (Table 5), while the populations of South Wollo ranked second in terms of Na, Ne, He and I. The lowest mean of Na, Ne, I, and genetic diversity were detected in the Metekel populations. This might be happening due to the different number of samples incorporated or considered in the seven population's viz., North Wollo population contains 34, South Wollo contains 18, Oromo special zone contains 8, Melkassa contains 6, Asossa contains 5, Metekel contains 6 and Kaffa contains 5 samples and also due to the selection of different SSR markers and the difference in the genetic diversity level of the experimental material. This result is supported by Adugna (2014), who reported the lower genetic diversity indices of the Metekel populations than those of the North and South Wollo populations and these might be due to some level of genetic drift. Tesfaye *et al*. (2016) also reports same results, who reports north wollo population was high genetic dversty next to southern Tigray. This might be due to environmental factors in North Wollo and South Wollo that contribute to high gene diversity include: (1) diverse Climatic Conditions: The regions experience variations in temperature, rainfall patterns, and elevation, which can influence the adaptation and survival of different plant species. (2) Topography: The diverse topography of these regions provides different ecological niches and habitats, which further promotes gene diversity by providing opportunities for natural selection and genetic exchange. (3) Soil Characteristics: The varied soil characteristics in these regions also contribute to gene diversity by providing different environments that favor the growth and survival of different plant species (Kemal and Hailu, 2019). The comparison between almost similar sample size populations such as Melkassa, Asossa, Metekel and Kaffa, Kaffa population is highest Na, Ne, NPA, NLCA, I, and He but Metekel population was lowest Na, Ne, I, He and high fixation index. This supports above comparison that showed Metekel population have low genetic variation than other population used in the present study. This indicates that Metekel populations are limited gene flow due to low climate divergence and

almost same soil properties in Metekel. Its high fixation index further implies the Metekel populations high inbreeding/selfing and homozygous. On other hand, Kaffa showed highest genetic variation than Melkassa, Asossa, Metekel. This might be due to low fixation index or inbreeding within Kaffa population than Metekel population.

The highest number of different allele frequency > = 5% (Table 5) was noted in the sweet sorghum accessions of North Wollo (6.4) and South Wollo (6.5). It indicates the existence of multiple common alleles present in the population, which implies a larger pool of genetic variation within the population. This can help to increase the population's overall genetic flexibility and resilience. The lowest (3.5) number of different allele frequency >= 5% was observed in Metekel, which indicates the presence of minor or less common alleles in the population.

North Wollo populations showed the highest number (0.8) of private alleles (Table 5), indicates the populations have relatively high numbers of alleles that are unique to them and not shared with other populations. The observation of a high number of private alleles (NPA) may indicate gene flow from wild to cultivated sorghum (Regasa, 2021). These populations could be used as a source of important traits in future sweet sorghum breeding programs in Ethiopia because the private alleles provide unique genetic variability in certain loci. It implies that the North Wollo region is very appropriate for further assessments of the genetic variability of sweet sorghum, germplasm conservation and source of useful alleles for breeding purposes. Kaffa was also ranked secondly based on highest NPA next to North Wollo. The smallest number (0.10) of private alleles was observed in the Oromo special zone and Melkassa. Ethiopia was a center of sweet sorghum diversity, which contains abundant wild accessions that grow side by side with cultivated accessions, resulting in regular hybridization between the different gene pools. It might be expected that unimproved sweet sorghum collections from the different region would have a high proportion of private alleles when compared with improved ones in Melkassa.

The important results in Table 5 is the numbers of local common allele frequency (NLCA) < = 50%, which ranged from 0.5 to 3.30. The North Wollo populations possessed the highest (3.30) local common allele frequency <= 50, which indicates the existence of common allele frequency in the populations. It leads to the presence of genetic variation in the populations, while the Asossa population exhibited the lowest (0.5) local number of allele frequency < = 50%. These findings are somewhat surprising given the fact that other research shows that

North Wollo populations showed an observable variation in loci carrying private alleles, implying the existence of high genetic uniqueness. In this study, the polymorphic loci per population showed high polymorphism across all populations (North Wollo, South Wollo, Melkassa, Asossa, Metekel, Oromo special zone and Kaffa). It indicates high genetic variation across populations.

Table 5: Allelic patterns and diversity indices across populations averaged over the 10 SSR loci

Diversity indices	**NW**	**SW**	**Om.S**	**Mk**	**Asos**	**Met**	**Kaf**	**Mk**
Number of accession	34	18	8	6	5	6	5	11.7
Number of alleles	9	7.2	5	3.6	3.6	3.5	4.2	5.16
Number of effective alleles	5.75	5.2	4.15	3.05	3.25	2.97	3.75	4.02
Na Freq. >= 5%	6.4	6.5	5	3.6	3.6	3.5	4.2	4.69
Number of Private Alleles	0.8	0.4	0.1	0.1	0.3	0.4	0.6	0.39
NLCA < = 50%	3.3	3	1.4	0.6	0.5	0.7	0.9	1.49
Shannon's information index(I)	1.88	1.74	1.49	1.13	1.17	1.1	1.32	1.4
Observed heterozygosity(Ho)	0.04	0.06	0.06	0.03	0.02	0.02	0.08	0.04
Nei's genetic diversity (He)	0.81	0.79	0.75	0.63	0.64	0.61	0.69	0.7
Fixation index(F)	0.96	0.93	0.92	0.96	0.97	0.98	0.9	0.94

NW = North Wollo, SW = South Wollo, Om.S = Oromo special zone, Mk = Melkassa, Asos = Asossa, Met = Metekel, Kaf = Kaffa, M = mean, Na Freq. >= 5% = number of different allele frequency > = 5%, NLCA < = 50% = *numbers of local common allele frequency (NLCA) < = 50%.*

In this experiment, sweet sorghum from Kaffa exhibited the highest number of observed heterozygosity (0.08), while the lowest mean number (0.02) of observed heterozygosity was observed from Asossa and Metekel. The differences might be due to inbreeding, non-random mating, low gene flow or seed exchange and number of self-generated generations from Asossa and Metekel sources, but sweet sorghum accessions from Kaffa might be due to high seed exchange and few cross-pollination events. The overall mean value of observed heterozygosity (Ho) of 0.04 was lower than the corresponding average of expected heterozygosity of 0.7, indicating a high amount of homozygosity. According to Robotham and Chapman (2017), the low observed heterozygosity shows the inbreeding index characteristic of a predominantly self-fertilizing crop with little outcrossing. Moreover, differences existed in the populations, indicating the presence of genetic variation. Similar

results have been previously reported in sorghum (Adugna, 2014; Nemera *et al*., 2022), who report an overall mean of expected heterozygosity of 0.67.

4.3. Genetic Distance between Populations

The pairwise Nei's standard genetic distance of each population from the other populations ranged from 0.18 to 1.12 (Table 6). Pair-wise population genetic distance analysis confirmed the existence of relatively low genetic distances among the various sweet sorghum populations likely due to the presence of higher gene flow among the populations. The relatively highest genetic distance (1.12) was observed between sweet sorghum collections from Melkassa and Metekel, which was followed by the relationship between Asossa and Kaffa and Metekel and South Wollo populations that scored a distance of 1.09 and 1.05, respectively. The greatest genetic dissimilarity was recorded between Melkassa and Metekel, likely due to geographical isolation that hindered or limited the gene flow (Nm = 0.84) between these two places (Gemechu, 2017). There appears to have been limited sharing or exchange of sweet sorghum germplasm between the two regions historically. This lack of germplasm mixing allowed the genetic differences between the populations to accumulate over generations. According to Asfaw (2014), over the last four decades, over 40 varieties have been introduced for various agroecologies, with the exception of Ethiopia's rainy lowlands, especially the Metekel zone. But Melkassa verities are improved one; this makes both Melkassa and Metekel populations have high distant populations. The smallest genetic distance (0.18) was observed between sweet sorghum collections from South Wollo and North Wollo. This may be due to the geographical closeness of the two zones where high possibility to exchange the seed between farmers, get the seeds from Market, migration or through river and the high gene flow (Nm = 12.20). The current investigation found the existence of a wider genetic distance between sweet sorghum accessions obtained from Melkassa and Metekel populations. Therefore, it could be important for the genetic improvement of the crop by selecting parental lines from these two distant populations (Adugna, 2014).

Table 6: Pair-wise Nei's genetic distances among the seven sweet sorghum populations from Ethiopia

Populations	North Wollo	South Wollo	Oromo special zone	**Melkassa**	**Asossa**	**Metekel**	**Kaffa**
North Wollo	-						
South Wollo	**0.18**	-					
Oromo special zone	0.35	0.434	-				
Melkassa	0.64	0.68	0.51	-			
Asossa	0.84	0.88	0.64	0.45	-		
Metekel	0.99	1.05	0.95	**1.12**	0.85	-	
Kaffa	0.82	0.82	0.70	0.63	1.09	0.88	-

4.4. Population Genetic Differentiation

The results of pair-wise genetic differentiation (below diagonal) and gene flow (Nm) (above diagonal) was presented in (Table 7). The pair-wise genetic differentiation of each population from the other populations ranged from 0.02 to 0.23 (Table 7). The highest population genetic differentiation was observed between sweet sorghum collections from Melkassa and Metekel population (Fst: 0.23) with lowest gene flow (Nm = 0.84), which was followed by Asossa and Metekel populations that scored Fst of 0.18 and gene flow of 1.14. The highest population genetic differentiation might be due to low gene flow between these two populations. Because, gene flow can influences genetic diversity, which encompasses several mechanisms of gene exchange among populations (Slatkin, 1987). The lowest population genetic differentiation was between South Wollo and North Wollo (Fst: 0.02). This low genetic differentiation among populations may be due to high gene flow (Nm = 12.20) that resulted from seed exchange practices between communities, marketing and the closest geographic region between two populations. According to Seboka and Hintum (2006), sorghum accession of farmer's seed systems are characterized by widespread germplasm sharing and gene flow dynamics in Ethiopia. So, high gene flow causes the presence of identical alleles and reduces genetic differentiation between these two populations. In this study, there was gene flow between in all seven populations. This implies there was seed exchange between population by different factors, such as water, animals, human activities or migration of human and market (Wichmann *et al*., 2009; Nazareno *et al*., 2021).

Table 7: Pair-wise genetic differentiation (FST) below the diagonal and pair-wise estimates of gene flow (Nm) above the diagonal among 82 sweet sorghum populations from Ethiopia

Population	North Wollo	South Wollo	Oromo special zone	Melkassa	Asossa	Metekel	Kaffa
North Wollo	-	**12.20**	5.82	1.97	1.75	1.33	2.11
South Wollo	**0.02**	-	4.43	1.84	1.64	1.23	2.06
Oromo special zone	0.04	0.05	-	2.59	2.27	1.28	2.53
Melkassa	0.11	0.12	0.09	-	2.57	**0.84**	1.93
Asossa	0.13	0.13	0.10	0.09	-	1.14	1.25
Metekel	0.16	0.17	0.16	**0.23**	0.18	-	1.27
Kaffa	0.11	0.11	0.09	0.11	0.17	0.17	-

4.5. Analysis of Molecular Variance

The analysis of molecular variance was used to determine the extent of overall variation among populations, within accession and among accession (Table 8). Analysis of molecular variance revealed that there were highly significant molecular variances ($P<0.001$) among populations and within populations. Further analysis showed that the highest genetic variability (90%) was attributable to variation among accession, while 5% was attributed to variation among populations and 5% was attributed to variation within accession. This result is in agreement with Regasa (2021) and Nemera *et al*. (2022), who reported 6.74% and 9.54% variation among sorghum population and 93.26% and 90.46% variation within population, respectively, in Ethiopia. The lowest variation among sorghum population was also reported in Kenya (Muui *et al.*, 2016), Ethiopia (Mamo *et al.*, 2023), and South Africa (Motlhaodi *et al.*, 2017), who reported 2.75%, 2.32% and 9.56% variation among population, respectively.

According to Wright (1951), this analysis revealed the presence of moderate genetic differentiation (Fst = 0.093) with high gene flow (Nm = 5.033) between populations in the present study (Table 8). This moderate genetic differentiation might be due to gene flow through continuous gene exchange due to sharing seeds from different populations have been collected, this cause increase of gene flow. The high variation among accession observed in this study could be due to seed exchange in breeding programs, sexual recombination and spontaneous mutation (Nemera *et al.*, 2022; Mamo *et al.*, 2023).

The overall observed gene flow (Nm) in this study was 5.033, which showed high gene flow from one population to the other. Seed exchange among sweet sorghum growing regions and high gene flow (Nm = 5.033) could also contribute to the observed low variation among populations. The sweet sorghum accessions in the current study revealed high genetic diversity and low population divergence. Generally, it is very important that the variation can be used to increase the genetic improvement of the crop in the future.

Table 8: Analysis of molecular variance of seven population using 10 polymorphic SSR markers

Source	df	SS	MS	Est.	% of variation	Fst	P
Among Pops	6	73.162	12.194	0.204	5%	0.09	0.001
Among accession	75	599.814	7.998	3.889	91%		
Within accession	82	18.000	0.220	0.220	5%		
Total	163	690.976	20.411	4.313	100%		
NM	5.033						

df = degree of freedom, SS=sum of squares and MS=mean squares, Fst = the coefficient genetic differentiation, Est.var- estimated variance, Nm = [(1 / Fst) - 1]/ 4.

4.6. Clustering Analysis

The Unweighted Neighbor-joining analysis divided the 82 accessions into three major clusters (I, II and III) (Figure 2). The first cluster (I) consists of 35.37% from different populations (14.64% from North Wollo, 17.07% from South Wollo, 2.44% from Oromo special zone and 1.22% from Kaffa). The second cluster (II) consists of 21.95% from different populations (12.19% from North Wollo, 6.10% from South Wollo, 2.44% from Oromia special zone, 1.22% from Kaffa). Both the first cluster and second clusters do not contain sweet sorghum accession from Metekel, Asossa and Melkassa. The third cluster (III) consists of 42.68% (10.97% from North Wollo, 2.44% from South Wollo , 4.88% from Oromia special zone, 7.32% from Metekel, 6.10 from Asossa, 7.32% from Melkassa , 3.65% from Kaffa). This cluster contains accessions from all 7 populations, but the number of accessions in the distribution was not homogeneous from each population. All Melkassa and Asossa accessions are also found in 3rd clusters.

This study has been unable to demonstrate that the studied accessions clearly cluster according to their geographical location, but it has revealed admixture. Several factors could explain this observation, Firstly, this might be due to the presence of higher gene flow through the extensive seed exchange existing between sweet sorghum farmers of different regions through marketing and migration of people; second, contentious introduction of new seeds into the respective sweet sorghum growing regions; third, animals movement and other means leading low to moderate genetic differentiation (Regasa, 2021 and Abtew, 2020). This result is consistent and similar to the reports of Motlhaodi *et al*. (2014); Tirfessa *et al*. (2020); Regasa (2021) and Mamo *et al*. (2023), who inferred clustering of Ethiopian sorghum accessions are not reveal clear grouping based on either altitude or geographical origin. Another source of uncertainty is that Sorghums collected from Ethiopia and Eritrea also failed to form clusters according to their adaptation and ecological zones as analyzed by RAPD (Ayana *et al*., 2000). The cluster analysis in the present study is independent of the effect of geographical distribution and low population differentiation with higher genetic diversity.

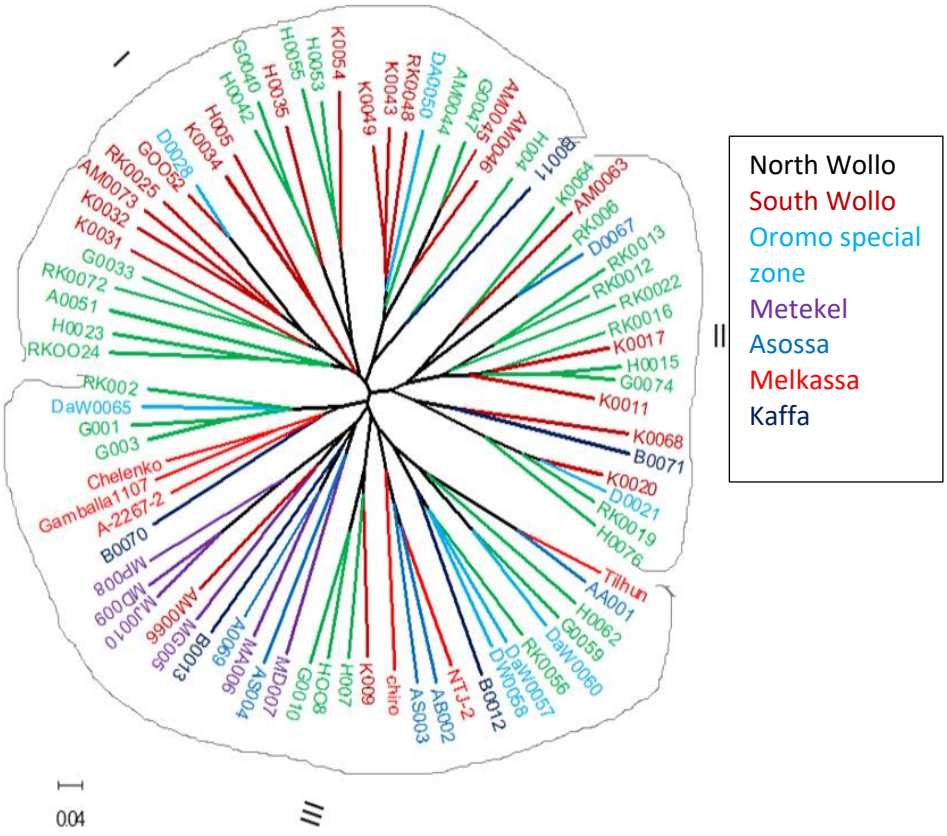


Figure 1. Dendrogram for 82 Ethiopian sweet sorghum accessions based on UWNJ with the pair group method

4.7. Principal Coordinates Analysis

Principal coordinate analysis (PCoA) results showed that the three most informative principal coordinates accounted for about 20.35% of the genetic variation in the SSR molecular data obtained from sweet sorghum in the study (Figure 3). The first, second and third principal coordinates explained about 7.51%, 6.96% and 5.89% of the total variation. The present study revealed a scatter plot with poor population clustering. It did not cluster the individuals into distinct groups according to their geographic sampling areas. In this analysis, sweet sorghum accessions scattered in the same coordinate showed the presence of a closer genetic relationship and narrow genetic distance between them, while accessions scattered in different coordinates showed the presence of a wider genetic variation. Based on these, the present study of the principal coordinate analysis result indicated the existence of high genetic variation in the sweet sorghum accessions and importance for further genetic improvement and breeding programs. All the Metekel and Asossa accessions were scattered in the first quadrants.

The presence of greater genetic variability within populations than between them suggested a PCoA, where individuals from different groups did not form distinct clusters (Pandian *et al*., 2020), but it has revealed admixture. This could be due to the presence of a high seed exchange rate among sweet sorghum growing regions of the Ethiopia and high gene flow. This PCoA result supports dendogram clustering (Fig 3) that did not form a distinct cluster based on geographic origin.

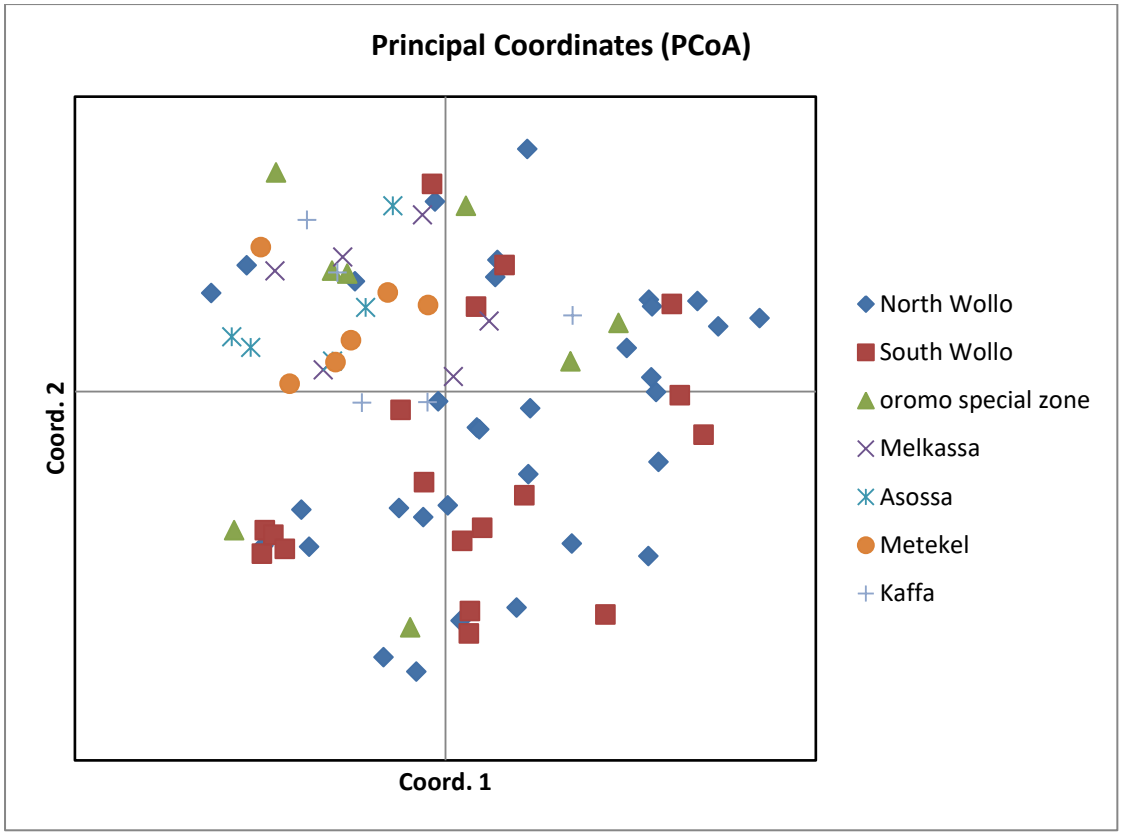
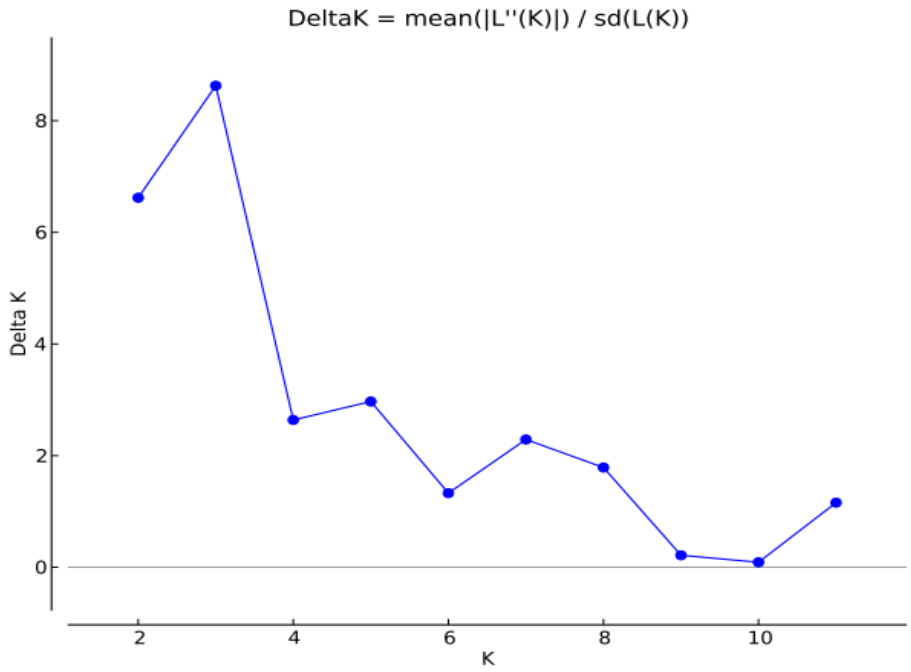


Figure 2. PCoA of the 82 sweet sorghum accessions as revealed by 10 microsatellite loci. Samples coded with the same symbol and color belongs to the same population

4.8. Population Structure

The 82 sweet sorghum accessions collected from different geographic locations were further classified into three distinct sub-populations (delta K = 3) using STRUCTURE 2.1 software program (Figure 4). The sub-populations generated were complementary to the Neighbor-Joining (NJ) tree analysis that also classified 82 sweet sorghum accessions into three clusters. This further validates the representative nature of the markers selected for this study. The present study on genetic diversity was in line with the reports of Disasa *et al.* (2017); Mamo *et al.* (2023); Regasa (2021); Misganaw *et al.* (2023). The structure analysis further substantiated the presence of three sub-populations with significant admixture in the study populations. It revealed the existence of weak sub-structuring (K = 3) of the seven populations of sweet sorghum. This finding may indicate seed exchange between farmers, leading to high gene flow across adjoining populations.

A.



B.

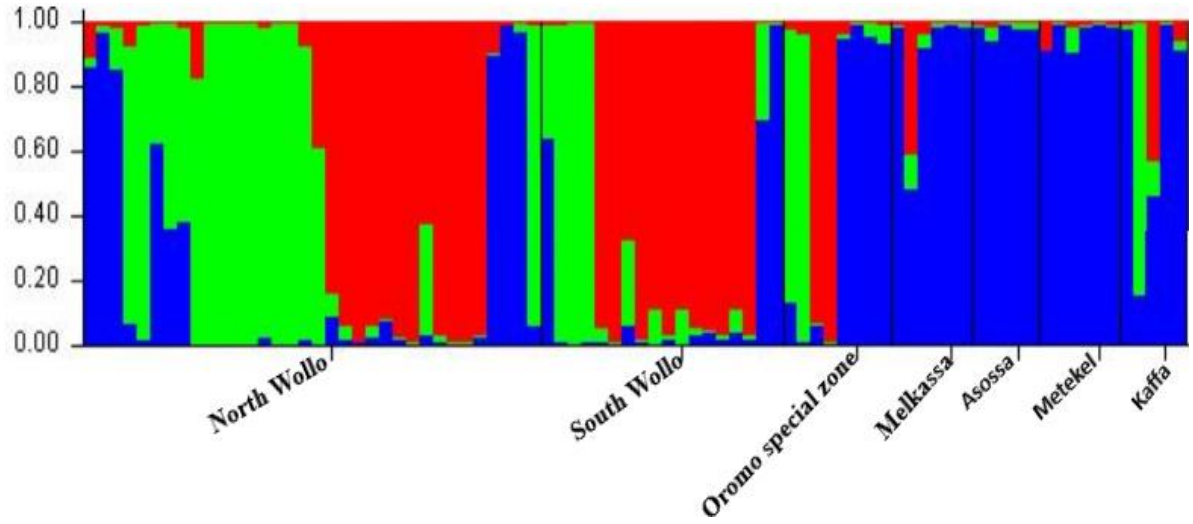


Figure 3: Population structure of 82 sweet Sorghum accessions representing Sven populations in Ethiopia. A) Evanno's K Statistics K=1 to K=12, and B) Bar plot display of STRUCTRE output of the 82 individuals from administrative regions formed three sub-populations (K =3). The x-axis represents individual samples and y-axis represents the proportion of ancestry to each cluster.

5. CONCLUSION AND RECOMMENDATIONS

Sweet sorghum (*Sorghum bicolor* L.) is types of sorghum with a sugar-rich stalk. Its sweet stalks and juice make it suitable for sugar production and biofuel production, while its drought tolerance and adaptability to diverse environments contribute to its agricultural significance. But sweet sorghum production is limited due to low productivity and not considered as an important crop in Ethiopia. Successful molecular characterization of this crop can help breeders to focus on the available genetic resources and implement improvement programs. The abundance of different alleles observed among the populations provides evidence for novel alleles that can be efficiently exploited through future breeding programs for the trait(s) of interest. The current study showed the presence of high genetic diversity among Ethiopian sweet sorghum accessions with lowest population differentiation due to high gene flow. This indicates that there is a wide range of genetic variation present among the sweet sorghum accessions in Ethiopia. So, focus conservation efforts within populations, not between them. Among tested populations, Metekel sweet sorghum collections showed the highest genetic distance with sweet sorghum collections from Melkassa. These populations could be useful for the appropriate selection of diversified parental lines for further improvement through breeding. On the other hand, sweet sorghum populations from North Wollo and Kaffa showed the highest private allele and thus, this region suggests their potential value for future breeding programs and the development of improved varieties with unique traits that are well-suited to the local agricultural conditions. Further studies need to focus on the following points: Genetic diversity study should be supported by SNP markers is recommended, and Considering other sweet sorghum growing areas and taking equal numbers of samples would increase the resolution and accuracy of the diversity analysis for breeders to exploit the genetic potential of the crop to improve its production and productivity through breeding.

6. REFERENCES

Adugna, A., 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), p.212.

Adugna, A., Snow, A.A., Sweeney, P.M., Bekele, E. and Mutegi, E., 2013. Population genetic structure of in situ wild Sorghum bicolor in its Ethiopian center of origin based on SSR markers. *Genetic resources and crop evolution*, ***60***, pp.1313-1328.

Agrama, H.A. and Tuinstra, M.R., 2003. Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *African journal of biotechnology*, ***2***(10), pp.334-340.

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

Allendorf, Fred W., 2014. Heterozygosity. obo in Evolutionary Biology

Ananda, G.K., Myrans, H., Norton, S.L., Gleadow, R., Furtado, A. and Henry, R.J., 2020. Wild sorghum as a promising resource for crop improvement. Frontiers in Plant Science, 11, p.558804.

Aruna, C., Visarada, K.B.R.S., Bhat, B.V. and Tonapi, V.A. eds., 2018. Breeding sorghum for diverse end uses. Woodhead Publishing.

Ayana, A., Bryngelsson, T. and Bekele, E., 2000. Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). *Genetic Resources and Crop Evolution*, ***47***(5), pp.471-482.

Beaumont, M.A., Ibrahim, K.M., Boursot, P. and Bruford, M.W., 1998. Measuring genetic distance.In *Molecular tools for screening biodiversity* ***16***:315-325.

Bednářová, M., Karafiátová, M., Hřibová, E. and Bartoš, J., 2021. B chromosomes in genus Sorghum (Poaceae). Plants, 10(3), p.505.

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

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics*, ***32***(3), p.314.

Bouargalne, Y., Ben Mrid, R., Bouchmaa, N., Zouaoui, Z., Benmrid, B., Kchikich, A., El Omari, R., Kabach, I. and Mohamed, N., 2022. Genetic diversity for agromorphological traits, phytochemical profile, and antioxidant activity in Moroccan sorghum ecotypes. *Scientific Reports*, ***12***(1), p.5895.

Bruinsma, J., 2017. World agriculture: towards 2015/2030: an FAO study. Routledge.

Caniato, F.F., Guimaraes, C.T., Schaffert, R.E., Alves, V.M.C., Kochian, L.V., Borém, A., Klein, P.E. and Magalhaes, J.V., 2007. Genetic diversity for aluminum tolerance in sorghum. *Theoretical and Applied Genetics*, ***114***, pp.863-876.

Chen, F., Zhao, Z., Liu, J., Cui, X., Zhang, D., Su, X., Zhao, J., Tang, P. and Liao, H., 2020. Automatic estimation of morphological characteristics of proximal tibia for precise plate treatment using model matching. *Computerized Medical Imaging and Graphics*, *81*, p.101714.

Ciampitti, I.A. and Prasad, P.V., 2020. Sorghum: State of the art and future perspectives.Cuevas, H.E. and Prom, L.K., 2013. Assessment of molecular diversity and population structure of the Ethiopian sorghum [Sorghum bicolor (L.) Moench] germplasm collection maintained by the USDA–ARS National Plant Germplasm System using SSR markers. *Genetic Resources and Crop Evolution*, ***60***, pp.1817-1830.

Cuevas, H.E., Knoll, J.E., Prom, L.K. and Vermerris, W., 2023. Genetic diversity, population structure and anthracnose resistance response in a novel sweet sorghum diversity panel. *Frontiers in Plant Science*, ***14***, p.1249555.

De Vicente, M.C., Lopez, C. and Fulton, T., 2004. Genetic diversity analysis with molecular marker data: learning module. *International Plant Genetic Resources Institute (IPGRI)*.

Dega, R.K.Y. and Ercal, G., 2015. A comparative analysis of progressive multiple sequence alignment approaches using UPGMA and neighbor joining based guide trees. *arXiv preprint arXiv*:1509.03530.

Degebasa, L., Tessema, T., Bekeko, Z. and Belete, K., 2022. Prevalence and Socioeconomic Impact of Striga (Striga hermonthica) in Sorghum Producing Areas of East and West Hararghe Zones, Oromia, Ethiopia. *International Journal of Agronomy*.

Derese, S.A., Shimelis, H., Laing, M. and Mengistu, F., 2018. The impact of drought on sorghum production, and farmer's varietal and trait preferences, in the north eastern Ethiopia: implications for breeding. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, ***68***(5), pp.424-436.

Desmae, H., Jordan, D.R. and Godwin, I.D., 2016. Geographic patterns of phenotypic diversity in sorghum (Sorghum bicolor (L.) Moench) landraces from North Eastern Ethiopia. *African Journal of Agricultural* ***Research, 11***(33), pp.3111-3122.

Dhaliwal, M.S., Yadav, A. and Jindal, S.K., 2014. Molecular characterization and diversity analysis in chilli pepper using simple sequence repeats (SSR) markers. *African Journal of Biotechnology*, ***13***(31).

Dicko, M.H., Gruppen, H., Traoré, A.S., Voragen, A.G. and Van Berkel, W.J., 2006. Sorghum grain as human food in Africa: relevance of content of starch and amylase activities. *African journal of biotechnology*, ***5***(5), pp.384-395.

Dido, A.A., Degefu, D.T., Assefa, E., Krishna, M.S.R., Singh, B.J.K. and Tesfaye, K., 2021. Spatial and temporal genetic variation in Ethiopian barley (Hordeum vulgare L.) landraces as revealed by simple sequence repeat (SSR) markers. *Agriculture & Food Security*, ***10***, pp.1-14

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). *Annals of botany*, ***100***(5), pp.975-989.

Disasa, T., Feyissa, T. and Admassu, B., 2017. Characterization of ethiopian sweet sorghum accessions for 0 brix, morphological and grain yield traits. *Sugar Tech*, ***19***, pp.72-82.

Disasa, T., Feyissa, T., Admassu, B., Paliwal, R., De Villiers, S.M. and Odeny, D.A., 2016. Molecular evaluation of Ethiopian sweet sorghum germplasm and their contribution to regional breeding programs. *Australian Journal of Crop Science*, ***10***(4).

Dube, F. and Maphosa, M., 2022. Significance of sweet sorghum as a multi-purpose crop for sub-Saharan Africa. *African Crop Science Journal*, ***30***(2), pp.235-243.

Earl, D.A. and VonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources*, ***4***, pp.359-361.

Ebraheim, E.H.A., 2016. Assessment of Genetic Diversity of some Sweet Sorghum Genotypes (Sorghum bicolor L.) using Agro-Morphological and Molecular Markers (Doctoral dissertation, Sudan University of Science and Technology).

Ekefre, D.E., Mahapatra, A.K., Latimore Jr, M., Bellmer, D.D., Jena, U., Whitehead, G.J. and Williams, A.L., 2017. Evaluation of three cultivars of sweet sorghum as feedstocks for ethanol production in the Southeast United States. *Heliyon*, ***3***(12).

Food and Agriculture Organization Statistics (FAOSTAT)., 2022. World food and agriculture statistical yearbook. Rome. [Available online] https://www.fao.org/faostat/en/#compare

Garcia, A.A., Benchimol, L.L., Barbosa, A.M., Geraldi, I.O., Souza Jr, C.L. and Souza, A.P.D., 2004. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetics and Molecular Biology*, ***27***, pp.579-588.

Gebreselassie, W., 2020. Characterization of genetic variation among Ethiopian barley (Hoerdeum vulgare L.) accessions. PhD dissertation, University of Hohenheim.

Gemechu, O., 2017. *Genetic Diversity Analysis of Sorghum (Sorghum bicolor (L.) Moench) Races in Ethiopia Using SSR Markers.* Msc thesis. Addis Ababa University. Addis Ababa.

Gerrano, A.S., Labuschagne, M.T., van Biljon, A. and Shargie, N.G., 2014. Genetic variability among sorghum accessions for seed starch and stalk total sugar content. *Scientia Agricola*, ***71***, pp.472-479.

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

Hart, G.E., Schertz, K.F., Peng, Y. and Syed, N.H., 2001. Genetic mapping of Sorghum bicolor (L.) Moench QTLs that control variation in tillering and other morphological characters. *Theoretical and Applied Genetics*, ***103***, pp.1232-1242.

Hazarika, K., Goswami, K.R., Bharali, B., Kalita, P., Goswami, J. and Neog, B., 2022. Physiological performance of sweet sorghum (Sorghum bicolor L. Moench) genotypes under Assam situation. *Research on Crops*, ***23***(3), pp.556-561.

Herniwati, H., Pabendon, M.B., Wicaksono, K.P., Waluyo, B. and Widaryanto, E., 2024. Estimation of genetic diversity of sweet sorghum (Sorghum bicolor (L.) Moench) genotypes as a bioethanol source using SSRs markers. *Czech Journal of Genetics and Plant Breeding*, ***60***(2), pp.86-96.

Hillis, D.M., 1987. Molecular versus morphological approaches to systematics. *Annual review of Ecology and Systematics*, ***18***(1), pp.23-42.

Höglund, J., 2009. Evolutionary conservation genetics. Oxford University Press.

Jemal, B. and Rebuma, M., 2019. Evaluation of Sweet Sorghum (Sorghum bicolor L. Moench) Genotypes for Agronomic Traits, Stalk Yield and Brix Value in Central Rift Valley of Ethiopia.

Jiang, D., Ma, T., Ding, F., Fu, J., Hao, M., Wang, Q. and Chen, S., 2019. Mapping global environmental suitability for Sorghum bicolor (L.) Moench. *Energies*, ***12***(10), p.1928.

Kanaka, K.K., Sukhija, N., Goli, R.C., Singh, S., Ganguly, I., Dixit, S.P., Dash, A. and Malik, A.A., 2023. On the concepts and measures of diversity in the genomics era. *Current Plant Biology*, ***33***, p.100278.

Kassahun Tesfaye, K.T., 2017. Genetic diversity study of sorghum (Sorghum bicolor (L.) Moenc) genotypes, Ethiopia.

Kemal, A. and Hailu, F., 2019. Genetic diversity of Safflower (Carthamus tinctorius L.) genotypes at Wollo, Ethiopia using agro-morphological traits. Trop Plant Res, 6(1), pp.157-165.

Khalsa, S.D.K., 2013. Genetic structure and diversity of Sorghum bicolor at three geographical scales in Africa (Master's thesis).

Kleih, U., Ravi, S.B., Rao, B.D. and Yoganand, B., 2000. Industrial utilization of sorghum in India.

Kowal, R.R., 2017. The Biology of Sorghum bicolor (L.) Moench subsp. bicolor (Sorghum). Australian Government, Department of Healt.

Lekgari, A. and Dweikat, I., 2014. Assessment of genetic variability of 142 sweet sorghum germplasm of diverse origin with molecular and morphological markers. *Open Journal of Ecology*, ***4***(07), p.371.

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

López-Sandin, I., Gutiérrez-Soto, G., Gutiérrez-Díez, A., Medina-Herrera, N., Gutiérrez-Castorena, E., Galicia-Juárez, M. and Zavala-García, F., 2021. Biomass and sugar production dynamics in sweet sorghum variety Roger. *Chilean journal of agricultural research*, ***81***(1), pp.92-101.

Mamo, W., Enyew, M., Mekonnen, T., Tesfaye, K. and Feyissa, T., 2023. Genetic diversity and population structure of sorghum [Sorghum bicolor (L.) Moench] genotypes in Ethiopia as revealed by microsatellite markers. *Heliyon*, ***9***(1).

Mangena, P., Shimelis, H.A., Laing, M.D. and Beyene, A.A., 2018. Genetic interrelationship of sweet stem sorghum genotypes assessed with simple sequence repeat markers. *South African Journal of Plant and Soil*, ***35***(5), pp.351-358.

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

McDermott, J.M. and McDonald, B.A., 1993. Gene flow in plant pathosystems. Annual review of phytopathology, ***31***(1), pp.353-373.

Mekonnen, T., Sathelly, K., Sharma, M., Tesfaye, K. and Kaul, T., 2017. Genomic DNA isolation method from fresh wheat leaf samples without liquid nitrogen. *African Journal of Biotechnology*, ***16***(20), pp.1192-1196.

Misganaw, A., Feyissa, T., Mekonnen, T., Desalegne, O. and Disasa, T., 2023. Genetic diversity analysis of sorghum genotypes for sustainable genetic resource conservation and its implication for breeding program in ethiopia. *Genetic Resources and Crop Evolution*, ***70***(6), pp.1831-1852.

Mohammadi, S.A. and Prasanna, B.M., 2003. Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop science*, ***43***(4), pp.1235-1248.

Mokariya, L.K. and Malam, K.V., 2020. Sweet Sorghum: A smart, multipurpose crop. Agriculture and Environment, 1(1), pp.61-68. Mokariya, L.K. and Malam, K.V., 2020. Sweet Sorghum: A smart, multipurpose crop. *Agriculture and Environment*, ***1***(1), pp.61-68.

Morrissey, E.M., Mau, R.L., Schwartz, E., McHugh, T.A., Dijkstra, P., Koch, B.J., Marks, J.C. and Hungate, B.A., 2017. Bacterial carbon use plasticity, phylogenetic diversity and the priming of soil organic matter. *The ISME Journal*, ***11***(8), pp.1890-1899.

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. *Genetic resources and crop evolution*, ***64***, pp.599-610.

Murray, S.C., Rooney, W.L., Hamblin, M.T., Mitchell, S.E. and Kresovich, S., 2009. Sweet sorghum genetic diversity and association mapping for brix and height. *The plant genome*, ***2***(1).

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. *African Journal of Biotechnology*, ***15***(8), pp.264-271.

Nasidi, M., Agu, R., Walker, G. and Deeni, Y., 2019. Sweet sorghum: agronomic practice for food, animal feed and fuel production in Sub-Saharan Africa. Sweet sorghum: characteristics, cultivation and uses.

Nazareno, A.G., Knowles, L.L., Dick, C.W. and Lohmann, L.G., 2021. By animal, water, or wind: can dispersal mode predict genetic connectivity in riverine plant species?. *Frontiers in Plant Science*, *12*, p.626405.

Nei, M. and Chesser, R.K., 1983. Estimation of fixation indices and gene diversities. *Annals of human genetics*, ***47***(3), pp.253-259.

Nemera, B., Kebede, M., Enyew, M. and Feyissa, T., 2022. Genetic diversity and population structure of sorghum [Sorghum bicolor (L.) Moench] in Ethiopia as revealed by microsatellite markers. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, ***72***(1), pp.873-884.

Nerbewende, S., Teyoure, B.J.B., Zakaria, K., Mahamadi, H.O., Wend-Pagnangde, M.S.F.Z., Kiswendsida, R.N., Nebie, B., Pauline, B.K., Renan, E.T., Mahamadou, S. and Jean-Didier, Z., 2018. Assessment of genetic diversity of Burkina Faso sweet grain sorghum using microsatellite markers. *African Journal of Biotechnology*, ***17***(12), pp.389-395.

Olani, G., Feyissa, T. and Disasa, T., 2021. Genetic diversity analysis of sorghum [Sorghum bicolor (L.) Moench] races in Ethiopia using simple sequence repeats (SSR) markers. *African Journal of Biotechnology*, ***20***(11), pp.431-439.

Olweny, C., Jamoza, J., Dida, M.M., Kimani, W., Njuguna, J., Githae, D., Kiawa, B., Yao, N., Kosambo, L., Sally, C. and Okori, P., 2014. High genetic diversity for improvement of sweet sorghum (Sorghum bicolor (L.) Moench) genotypes for sugar and allied products.

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: Biological Sciences*, ***90***, pp.179-190.

Peakall, R.O.D. and Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular ecology notes*, ***6***(1), pp.288-295.

Pei, Z., Gao, J., Chen, Q., Wei, J., Li, Z., Luo, F., Shi, L., Ding, B. and Sun, S., 2010. Genetic diversity of elite sweet sorghum genotypes assessed by SSR markers. *Biologia Plantarum*, ***54***, pp.653-658.

Prabhakar, Madhusudhana, R. and Aruna, C., 2022. Sorghum Breeding. In Fundamentals of Field Crop Breeding (pp. 367-447). *Singapore: Springer Nature Singapore*.

Prajapati, D.K., Pahuja, S.K., Verma, N.K. and Chaudhary, S., 2018. Morphological characterization of sorghum [Sorghum bicolor (L.) Moench] germplasm for DUS traits. *International Journal of Current Microbiology and Applied Sciences*, ***7***(2), pp.2058-2071.

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

Ramu, P., Billot, C., Rami, J.F., Senthilvel, S., Upadhyaya, H.D., Ananda Reddy, L. and Hash, C.T., 2013. Assessment of genetic diversity in the sorghum reference set using EST-SSR markers. *Theoretical and Applied Genetics*, ***126***, pp.2051-2064.

Rao, P.S., Umakanth, A.V., Reddy, B.V., Dweikat, I., Bhargava, S., Kumar, C.G., Braconnier, S. and Patil, J.V., 2013. Sweet sorghum: genetics, breeding and commercialization. Biofuel crops: production, physiology and genetics. CABI, *Nosworthy Way*, pp.172-198

Ratnavathi, C.V., Chakravarthy, S.K., Komala, V.V., Chavan, U.D. and Patil, J.V., 2011. Sweet sorghum as feedstock for biofuel production: a review. *Sugar Tech*, ***13***(4), pp.399-407.

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

Reddy, B.V.S. and Reddy, P.S., 2003. Sweet sorghum: characteristics and potential. *International Sorghum and Millets Newsletter*, ***44***, pp.26-28.

Reddy, V.G., Upadhyaya, H.D. and Gowda, C.L.L., 2006. Current status of sorghum genetic resources at ICRISAT: their sharing and impacts. International Sorghum and Millets Newsletter, 47, pp.9-13.

Regasa, R.S., 2021. *Genetic diversity and Population Structure Analysis of Sorghum [Sorghum bicolor (L.) Moench] using Microsatellite markers in Ethiopia* (Doctoral dissertation, Addis Ababa University).

Ritter, K.B., McIntyre, C.L., Godwin, I.D., Jordan, D.R. and Chapman, S.C., 2007. An assessment of the genetic relationship between sweet and grain sorghums, within Sorghum bicolor ssp. bicolor (L.) Moench, using AFLP markers. *Euphytica*, ***15***7, pp.161-176.

Robotham, O., Chapman, M. Population genetic analysis of hyacinth bean (Lablab purpureus (L.) Sweet, Leguminosae) indicates an East African origin and variation in drought tolerance. Genet Resour Crop Evol 64, 139–148 (2017). https://doi.org/10.1007/s10722-015-0339-y.

Roger, L. and Willis, M., 2019. Sweet sorghum: characteristics, cultivation, and uses. *Nova Science Publishers.*

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), pp.789-803.

Schloss, S., Mitchell, S., White, G., Kukatla, R., Bowers, J., Paterson, A. and Kresovich, S., 2002. Characterization of RFLP probe sequences for gene discovery and SSR development in Sorghum bicolor (L.) Moench. *Theoretical and Applied Genetics*, ***105***, pp.912-920.

Seboka, B. 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. *Genetic Resources and Crop Evolution*, ***53***, pp.1385-1403.

Serrote, C.M.L., Reiniger, L.R.S., Silva, K.B., dos Santos Rabaiolli, S.M. and Stefanel, C.M., 2020. Determining the polymorphism information content of a molecular marker. *Gene*, ***726***, p.144175.

Sinha, S. and Kumaravadivel, N., 2016. Understanding genetic diversity of sorghum using quantitative traits. *Scientifica*, 2016.

Slatkin, M., 1985. Rare alleles as indicators of gene flow. Evolution, ***39***(1), pp.53-65.

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

Smith, C.W. and Frederiksen, R.A. eds., 2000. *Sorghum: Origin, history, technology, and production* (Vol. **2**). *John Wiley & Sons*.

Srinivasa Rao, P., Ganesh Kumar, C. and Reddy, B.V., 2013. Sweet sorghum: from theory to practice. *Characterization of improved sweet sorghum cultivars*, pp.1-15.

Stuber, C.W., Polacco, M. and Lynn, M., 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. *Crop Science*, ***39***(6), pp.1571-1583.

Sudhakar Reddy, P., Srinivas Reddy, D., Sivasakthi, K., Bhatnagar-Mathur, P., Vadez, V. and Sharma, K.K., 2016. Evaluation of sorghum [Sorghum bicolor (L.)] reference genes in various tissues and under abiotic stress conditions for quantitative real-time PCR data normalization. *Frontiers in plant science*, ***7***, p.172935.

Sukumaran, S. and Yu, J., 2014. Association mapping of genetic resources: achievements and future perspectives. *Genomics of Plant Genetic Resources: Volume **1**. Managing, sequencing and mining genetic resources*, pp.207-235.

Tiendrebéogo, J., Sawadogo, N., Kiébré, M., Kiébré, Z., Tuina, S., Sawadogo, T.A., Nanema, K.R., Traoré, R.E., Ouédraogo, M.H. and Sawadogo, M., 2022. Genetic relationship between sweet grain Sorghum and the other Sorghum types cultivated in Burkina Faso assessed with nuclear microsatellite markers SSRs. *American Journal of Plant Sciences*, ***13***(6), pp.872-883.

Tirfessa, A., Tesso, T., Adugna, A., Mohammed, H. and Kiambi, D., 2020. Genetic diversity among Ethiopian sorghum [Sorghum bicolor (L.) Moench] gene bank accessions as revealed by SSR markers.

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

Villanueva, B., Fernández, A., Saura, M., Caballero, A., Fernández, J., Morales-González, E., Toro, M.A. and Pong-Wong, R., 2021. The value of genomic relationship matrices to estimate levels of inbreeding. *Genetics Selection Evolution*, ***53***(1), p.42.

Wang, K. and Hiruki, C., 2001. Use of heteroduplex mobility assay for identification and differentiation of phytoplasmas in the aster yellows group and the clover proliferation group. Phytopathology, 91(6), pp.546-552.

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. *Theoretical and applied genetics*, ***120***, pp.13-23.

Waniska, R. D., Rooney, I. W. and Mcdonough C. M., 2004. Sorghum utilization. In: Wrigley, C., Corke, H. and Walker, C. *(Eds) Encyclopedia of Grain Science. Oxford: Elsevier.*

Waples, R.S., 1987. A multispecies approach to the analysis of gene flow in marine shore fishes. *Evolution*, ***41***(2), pp.385-400.

Welderufael, S., Abay, F., Ayana, A. and Amede, T., 2024. Genetic diversity, correlation and genotype× yield× trait (GYT) analysis of grain yield and nutritional quality traits in sorghum (Sorghum bicolor [L.] Moench) genotypes in Tigray, northern Ethiopia. *Discover Agriculture*, ***2***(1), p.4.

Whitfield, M.B., Chinn, M.S. and Veal, M.W., 2012. Processing of materials derived from sweet sorghum for biobased products. *Industrial Crops and Products*, ***37***(1), pp.362-375.

Wichmann, M.C., Alexander, M.J., Soons, M.B., Galsworthy, S., Dunne, L., Gould, R., Fairfax, C., Niggemann, M., Hails, R.S. and Bullock, J.M., 2009. Human-mediated dispersal of seeds over long distances. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1656), pp.523-532.

Wondimu, Z., Dong, H., Paterson, A.H., Worku, W. and Bantte, K., 2021. Genetic diversity, population structure, and selection signature in Ethiopian sorghum [Sorghum bicolor L.(Moench)] germplasm. *G3*, ***11***(6), p.jkab087.

Wright, S., 1951.The genetic structure of populations. *Annual Eugenics***.15:** 323-354

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

Xiao, M.Z., Sun, Q., Hong, S., Chen, W.J., Pang, B., Du, Z.Y., Yang, W.B., Sun, Z. and Yuan, T.Q., 2021. Sweet sorghum for phytoremediation and bioethanol production. *Journal of Leather Science and Engineering*, ***3***, pp.1-23.

Yahaya, M.A., Shimelis, H., Nebie, B., Ojiewo, C.O., Rathore, A. and Das, R., 2023. Genetic diversity and population structure of African sorghum (Sorghum bicolor L. moench) accessions assessed through single nucleotide polymorphisms markers. *Genes*, ***14***(7), p.1480.

Yali, W. and Begna, T., 2022. Sorghum breeding in Ethiopia: Progress, achievements and challenges. J Agric Sc Food Technol, ***8***(1), pp.045-051.

Zhang, H., Wang, J., Wang, D., Yao, F., Xu, J., Song, G., Guan, Y. and Li, R., 2011. Assessment of genetic diversity in Chinese sorghum landraces using SSR markers as compared with foreign accessions. *Acta Agronomica Sinica*, ***37***(2), pp.224-234.

Zhang, Q., Guldbrandtsen, B., Bosse, M., Lund, M.S. and Sahana, G., 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. BMC *genomics*, **16,** pp.1-16.

Zheng, J., Liu, T., Qian, Y., Li, J., Liu, Y., Cheng, S., Xu, Z. and Zhan, Q., 2021. Assessment of genetic diversity of sweet sorghum collection using phenotypic variation and ssr markers. *Pak J Bot*, ***53***(1), pp.145-153.

7. APPENDIX

Appendix 1. Solution preparation for DNA extraction

1. Tris-HCl (1 M, pH 8,0) 100 ml (Store @ RT)
- Dissolve 12.1 g Tris in approx. 70 ml demineralized water
- Adjust the pH to 8.0 with concentrated HCl
- Mix well and add demineralized water up to 100 ml
- Autoclave at 121°C

2. EDTA (0.5 M, pH 8.0) 100 ml (Store @ RT)
- Dissolve 18.6 g EDTA in approx. 70 ml demineralized water
- Adjust the pH to 8.0 with NaOH (e.g. 1M solution) EDTA dissolves only after the adjustment of the pH
- Mix well and add demineralized water up to 100 ml
- Autoclave at 121°C

3. 10 x TE-buffer (pH 8.0) 100 ml (Store @ 4°C)
- Mix 1 ml Tris-HCl 1M (Stock solution 1) and 200 µl EDTA 0.5M (Stock solution 2)
- Mix well and add demineralized water up to 100 ml
- Autoclave at 121°C

4. 10 x TBE buffer 1 Liter (Store @ RT)
- Dissolve 108 g Tris and 55 g Boric acid in approx. 700 ml sterile demineralized water
- Add 40 ml 0.5 M Na-EDTA (Stock solution 2)
- Adjust pH to 8.57 with 5M NaOH
- Mix well and add demineralized water up to 1 Liter

5. 25 x TAE buffer 1 Liter (Store @ 4°C)
- Dissolve 121 g Tris in approx. 700 ml sterile demineralized water
- Carefully add 57.1 ml glacial acetic acid in a fume hood
- Add 50 ml 0.5 M Na-EDTA (Stock solution 2)
- Mix well and add demineralized water up to 1 Liter

6. 6x Loading Dye without Xylencyanol 50 ml (Store @ 4°C)
- 0.125 g Brome phenol blue
- 1 ml Tris-HCl (Stock solution 1)
- Dissolve in 34 ml demineralized water
- Add 15 ml Glycerol (= 1.26 * 15 = 18.9 g)
- Make aliquots in Sterile Eppendorf caps

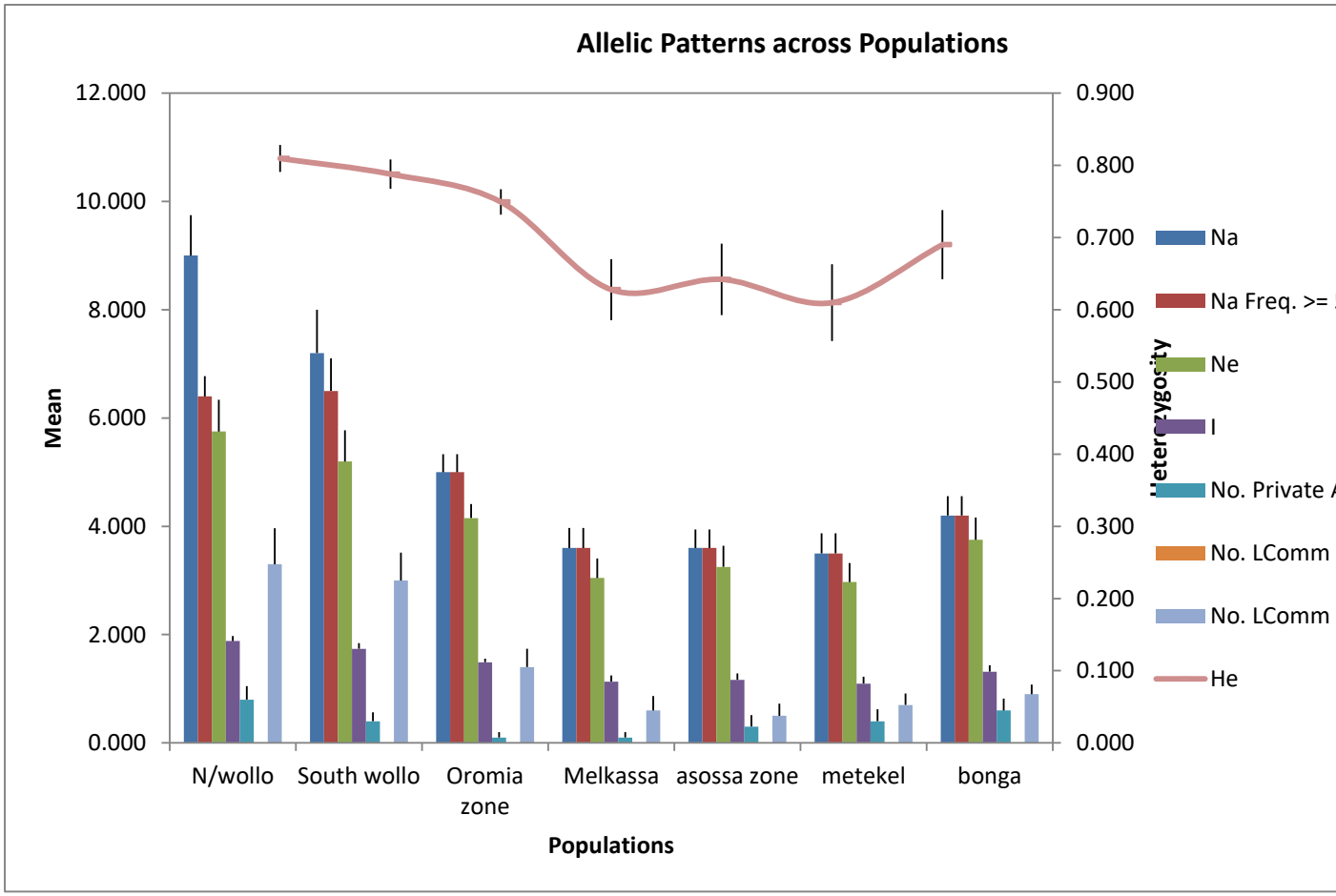
7. 5 M NaCl 100 ml (Store @ RT)
- Dissolve 29.22 g NaCl in 100 demineralized water
- Autoclave at 121°C

8. CTAB 2%, medium salt 100 ml (Store @ RT)
- Add in a 100 ml Flask
- 10 ml 1 M Tris-HCl (Stock solution 1)
- 28 ml 5M NaCl (Stock solution 7)
- 4 ml 0.5 M EDTA (Stock solution 2)
- 2 g CTAB
- 1 g PVP
- Add sterile demineralized water up to 100 ml
- (It is possible to warm the solution up to 50 °C to speed up the process)
- DO NOT AUTOCLAVE THIS SOLUTION!!!

9. 3 M NaAc-solution 250 ml (Store @ RT)
- Use a sterile 500 ml flask
- 61.5 g CH3CO2Na (MW 82,03)
- Add sterile demineralized water up to 250 ml

10. Decontamination Solution (For Ethidium Bromide)
- Weigh in 0.7 g Sodium Nitrite
- Add 3.3 ml Hypo Phosphorous Acid (50%)
- Add demineralized water up to 50 ml

Appendix 2. Allelic Patterns across Populations



Appendix 3 samples of visualized PCR products and DNA ladder using gel documentation UV trans illuminator at 2% agarose gel (Xisep0310 primer with Samples from 16 -39)

A.