Genetic diversity in yield traits and kernel composition of selected Ethiopian sorghum landraces

by

Chalachew Endalamaw Engida

Submitted in fulfilment of the requirements in respect of the of the

Doctoral Degree in Plant Breeding

in the Department of Plant Sciences in the Faculty of Natural and Agricultural Sciences at the University of the Free State

November 2023

Supervisor: Prof Maryke T. Labuschagne

Co-supervisors: Prof Liezel Herselman

Dr Angeline van Biljon

Dr Habte Nida

ABSTRACT

Climate-resilient and high-yielding sorghum cultivars must be developed for sustainable food systems and improved livelihoods in sub-Saharan Africa. Low grain protein, Fe and Zn concentrations are a frequent problem in food crops and they affect the crop's nutritional value, particularly in cereals. In this study, a total of 361 Ethiopian sorghum landraces with four commercial checks were grown under well-watered, wet intermediate and natural drought conditions for two consecutive seasons. The purpose of this research was to assess the genetic diversity and drought tolerance of this Ethiopian sorghum landrace collection using agronomic and grain quality traits and to select promising lines for production and/or breeding. The specific objectives were to (i) assess the level and pattern of genetic variation and drought adaptation in a sorghum landrace collection using phenotypic traits, (ii) evaluate chemical composition, genotype by environment interaction, drought effects and environmental influence and determine performance and stability of sorghum landraces, (iii) identify genomic regions associated with grain quality to determine the genetic basis for starch, protein and mineral concentrations of Ethiopian sorghum landraces.

In stress environments, yield and yield components as well as grain quality traits had low means, range, genetic variance (σ^2g) and heritability values. In terms of grain yield performance under different water regimes, Melkassa (irrigated) had a 54.5% higher mean grain yield than Miesso (drought stress) in 2020 and Melkassa (irrigated) had a 26.9% higher mean grain yield than Miesso (drought stress) in 2021. Therefor drought stress reduced mean grain yield by 44.6% and grain weight by 14%. In all environments, ten genotypes (G189, G210, G125, G324, G325, G353, G319, G219, G187 and G229) had the highest stress tolerance index (STI), geometric mean productivity (GMP), mean productivity (MP) and yield in non-stress conditions (Yp) and stress conditions (Ys). In stress environments, starch content decreased by 5%, while protein content increased by 21% when compared to wet intermediate conditions. The Zn content was 60% and 55% higher in wet intermediate environments than in irrigated and stress environments, respectively. The Fe content was 21% and 13% higher in wet intermediate environments than in irrigated and stress environments, respectively. A genome-wide association study (GWAS) on 365 diverse Ethiopian sorghum landraces identified 209,572 single nucleotide polymorphisms (SNPs) (HB $p \leq 0.05$). Several candidate genes were identified that were significantly associated with grain Zn, Fe, starch,

protein and ash concentrations, such as LEA, CHI, bZIP, ARF12 and WOX6. The Late Embryogenesis Abundant (LEA) and basic leucine zipper (bZIP) family of proteins are involved in various stress responses in plants. The ABC transporter and sugar carrier proteins, Auxin Response Factors 12 (ARF12), Alpha-glucan phosphorylase and Zinc finger CCHC gene, are known for starch synthesis and storage in growing seeds. The WUSCHEL-related homeobox 6 (WOX6) and TatD family of hydrolases genes, acts as a transcription factor and is known to regulate stem cell maintenance and differentiation. Chitinase (CHI), flavonoid glycosides, architecture1 (tga1) and Class III peroxidase 70 and 69 precursors are all engaged in biological processes including plant defence against pathogens. The Zinc finger proteins (ZNFs) family is involved in Zn and Fe ion uptake and transport. These genes are involved in various biological processes such as stress response, starch synthesis, plant development and defence. The identification of loci associated with grain quality provides new insight into the genetic control of the traits, while sorghum landraces with nutritious grains can serve as sources of genes for breeding for good nutritional value**.**

Keywords: candidate genes, environment, grain quality, grain yield, GWAS, sorghum landraces, stress tolerance

DECLARATION

I declare that the thesis hereby submitted for a Doctorate Degree in Plant Breeding at the Department of Plant Sciences (Plant Breeding), Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein, South Africa, is my personal, original and independent work that has not previously been submitted at any other university or institution of higher education. I also grant the University of the Free State copyright to this dissertation.

Chalachew Endalamaw Engida

November 16, 2023
Date

DEDICATION

This work is dedicated to my lovely wife, Dagmawit Tsegaye and my sons, Henok and Betselot, as well as all well-wishers.

ACKNOWLEDGMENTS

I sincerely thank the Almighty God for His abundant grace, sustenance, leadership, protection, love and favour. I would like to express my heartfelt gratitude to all the people and institutions who helped me finish my studies successfully. First, I want to express my gratitude to my promoter, Prof Maryke Labuschagne, who assisted me in obtaining a fellowship to pursue this study. Prof Maryke, thank you so much for your dedication, guidance and kindness throughout my studies. Prof Maryke, you are not only a supervisor and a scientist, but you are also a mother. By thinking like a mother and taking care of and facilitating all the necessary things, you are the one who motivated me to complete this study successfully. Special thanks to my co-promoters: Prof Liezel Herselman for your observations and constructive criticism, as well as your significant contribution to the completion of my thesis, for which I am grateful. Dr Angeline van Biljon, please accept my heartfelt gratitude for your technical assistance in carrying out the laboratory studies. Dr Habte Nida, I am truly grateful for your technical and courteous help on this study and your contributions to the thesis are greatly appreciated.

I am indebted to the University of the Free State and the Department of Plant Sciences for providing financial support for this study. I would like to thank all the lecturers and staff members in the Department of Plant Sciences, especially Orpah Taylor and Johnica Vlotman, for their assistance with the registration and logistical process at the University of the Free State.

I am grateful to Dr Diriba Gelete, Dr Taye Tadesse and Dr Bedru Beshir as well as the Ethiopian Institute of Agricultural Research and Melkassa Agricultural Research Centre management and administrative staff for their support with financial and administrative issues during my study. Special appreciation to the national sorghum breeding programme team, for supporting and managing the trials conducted at different locations in Ethiopia. A special thanks to Dr Alemu Tirfessa for his support in my research career during this study.

Finally, I want to express my heartfelt gratitude to my mother (Emuhaye Aleganesh Kebede), who invested in my life and education up to my first degree. Last but not least, I would like to thank my wife, Dagmawit Tesgaye, for her patience, support, understanding, encouragement and care of our lovely kids.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AEA	Average environment axis
AEC	Average environment coordinate
AM	Association mapping
AMY	Amylose
ANOVA	Analysis of variance
Av. P	Available phosphorus
BLINK	Bayesian information and linkage disequilibrium iteratively nested keyway
BLUP	Best linear unbiased predictor
Chr	Chromosome
CL	Cluster
CMLM	Compressed mixed linear model
DBE	Starch de-branching enzymes
DNA	Deoxyribonucleic acid
DTF	Days to 50% flowering
DTM	Days to 90% maturity
EC	Electrical conductivity
ENVT	Environment
FarmCPU	Fixed and random model circulating probability unification
Fe	Iron
G	Genotype
GBS	Genotyping-by-sequencing
GBSS	Granule-bond starch synthase
GEI	Genotype by environment interaction
GGE	Genotype main effects with genotype by environment interaction
GLM	General linear model
GMP	Geometric mean productivity
GWAS	Genome-wide association study
GY	Grain yield
H^2	Heritability
HB *p*-value	Holm-Bonferroni test

HLP	Head length
HWD	Head width
HWT	Head weight
ILP	Internode length per plant
JM20	Jimma 2020
JM21	Jimma 2021
kb	kilobase
LD	Linkage disequilibrium
masl	Meter above sea level
MAF	Minor allele frequency
MET	Multi-environment trial
MGE	Mega environment
MK20	Melkassa 2020
MK21	Melkassa 2021
MLM	Mixed linear model
MP	Mean productivity
MS20	Miesso 2020
MS21	Miesso 2021
MT	Metric tons
NIP	Number of internodes per plant
NIRS	Near-infrared spectrometer
Nobs	Number of observations
OM	Organic matter
PHT	Plant height per plant
bp	base pair
Pos	Position
QQ	Quantile-quantile
QTL	Quantitative trait loci
QTN	Quantitative trait nucleotide
R^2c	Coefficient for calibration
R^2cv	Coefficients for cross-validations

SBE	Starch branching enzyme
SE	Standard errors
SECV	Standard error of cross-validation
SNNP	Southern Nations, Nationalities and Peoples
SNP	Single nucleotide polymorphisms
SSI	Stress susceptibility index
SSR	Simple sequence repeat
SSS	Starch synthase
STI	Stress tolerance index
TN	Total nitrogen
TOC	Total organic carbon
TOL	Tolerance index
TSW	Thousand seed weight
UPGMA	Unweighted pair group method with arithmetic mean
Yp	Yield in non-stress conditions
$\bar{Y}p$	Mean yield of all non-stress conditions
Ys	Yield in stress conditions
$\bar{Y}s$	Mean yield of all cultivars under stress
YSI	Yield stability index
Zn	Zinc
σ^2en	Environmental variances
σ^2g	Genotypic variances
σ^2gen	Genotype x environment interaction variances

CHAPTER 1

INTRODUCTION

Sorghum (*Sorghum bicolor* (L.) Moench) originated in Ethiopia and is now grown all over the world (Harlan, 1969; Mekbib, 2006). Sorghum is drought tolerant and Ethiopia's most important crop, which is grown in the country's parched lowlands (Doggett, 1991; Kebede, 1991; Amelework et al., 2015). Sorghum can grow in very diverse and harsh environments, which makes it suitable for growing in regions where other major crops cannot be produced (Ananda et al., 2020). Ethiopian sorghum is a source of biotic stress resistance. For instance, wild sorghum species have been identified as resistant to ergot (Reed et al., 2002) and to Striga (Rich and Ejeta, 2008). Furthermore, research in Australia has identified a source of drought tolerance (stay green) genes to be found in four genomic regions (*Stg1, Stg2, Stg3* and *Stg4*) derived from a sorghum line (B35) native to Ethiopia (Borrell et al., 2003). Sorghum is cultivated over a wide range of elevations and rainfall conditions in Ethiopia. It is a particularly important food and feed crop in dry lowland areas where moisture is limited and it is often the only crop grown in these regions (Amelework et al., 2015). Farmers in Ethiopia consume high-lysine sorghum varieties as immature sorghum grain. High-lysine sorghums have higher-quality protein endosperm due to their high lysine content (Tesso et al., 2006).

Sorghum is important for subsistence farming in Africa, Asia and Latin America's semi-arid tropics. The major constraints and challenges of sorghum production are biotic and abiotic stresses, high agricultural inputs and cultural influence. It is the most important staple cereal crop in the semi-arid regions of eastern Africa and South Asia, while it is used as animal feed in the developed world (Mundia et al., 2019). Global sorghum grain production increased from 61.02 million metric tons for 2020 to 62.0 million metric tons in 2023. Ethiopia is among the top ten sorghum producing countries in the world and the third largest producer of sorghum in Africa (USDA, 2023). In 2019, 5.9 million small-holder farmers produced 4.8 million tons of sorghum grain from an area of 1.75 million hectares and the national average productivity was 2.7 tons/ha (CSA , 2022).

In Ethiopia, sorghum grain is liked as much as tef, a small cereal grain crop used for injera (a fermented flat bread), for animal feed, fuel, construction purposes and the preparation of beverages. Preparing injera from sorghum has considerable economic benefits over tef, as sorghum commands a much lower price and preparing tef is time-consuming and expensive. Micronutrient malnutrition, along with under and over nutrition, is among the greatest global challenges of current times. Malnutrition occurs when nutrient intake does not meet the needs for normal body functions and, consequently, leads to alterations in growth and development in children. Fighting malnutrition reduces health-care costs while increasing productivity (Larson-Nath and Goday, 2019; Dukhi, 2020). In Ethiopia, one of the most important public health issues is the high prevalence of child malnutrition that is a major impediment to improving child health outcomes (Mengistu et al., 2013). It is leading to a double burden of malnutrition on nations and more so on developing countries. To address malnutrition, dietary-induced micronutrient supplementation and food fortification are important to ensure and develop low-cost sustainable solutions to address this problem globally (Dukhi, 2020).

Genome-wide association studies (GWAS) is a research approach used to identify genomic variants that are statistically associated with a particular trait. In sorghum, GWAS has been used to identify quantitative trait nucleotides (QTNs) related to a variety of grain quality variables such as polyphenols, proteins and fat, minerals, amylose, starch, crude protein, crude fat and gross energy (Elango et al., 2020; Kimani et al., 2020; Lee et al., 2023).

The overall aim of this study was to understand the diversity and genomics of kernel composition of Ethiopian sorghum landrace accessions, as well as their phenotypic variability and general and drought adaptation.

The specific objectives of this study were to:

(i) Assess the level and pattern of genetic variation and drought adaptation in a sorghum landrace collection using phenotypic traits.

(ii) Evaluate chemical composition, genotype by environment interaction, drought effects and environmental influence and the performance and stability of sorghum landraces.

(iii) Identify genomic regions associated with grain quality to determine the genetic basis for starch and protein concentrations of Ethiopian sorghum landraces.

(iv) Determine the genetic basis of iron (Fe), zinc (Zn) and ash content in Ethiopian sorghum landraces by identifying genomic areas related to these mineral and ash concentrations.

References

Amelework, B., Shimelis, H., Tongoona, P. and Laing, M. (2015). Physiological mechanisms of drought tolerance in sorghum, genetic basis and breeding methods: a review. *African Journal of Agricultural Research,* 10: 3029-3040.

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: 1108.

Borrell, A., Van Oosterom, E., Hammer, G., Jordan, D. and Douglas, A. (2003). The physiology of "stay-green" in sorghum. In: Proceedings of the 11th Australian Agronomy Conference, Solutions for a better environment, Victoria, pp. 2-6.

CSA (2022). Central Statistical Agency. The federal democratic republic of Ethiopia central statistical agency farm management practices (private peasant holdings, Meher season) statistical bulletin. https://www.csa.gov.et (accessed 01/12/2024).

Doggett, H. (1991). Sorghum history in relation to Ethiopia. In: Engels, J.M.M., Hawkes, J.G. and Worede, M. (eds). Plant Genetic Resources of Ethiopia, Cambridge University Press, Cambridge, pp. 140-159.

Dukhi, N. (2020). Global prevalence of malnutrition: evidence from the literature. In: Imran, M., Imran, A. (eds). Malnutrition. *Intech Open*, pp.1-16. Doi: 10.5772/intechopen.92006.

Elango, D., Xue, W. and Chopra, S. (2020). Genome-wide association mapping of epi-cuticular wax genes in *Sorghum bicolor*. *Physiology and Molecular Biology of Plants*, 26: 1727-1737.

Harlan, J.R. (1969). Ethiopia: a center of diversity. *Economic Botany,* 23: 309-314.

Kebede, Y. (1991). The role of Ethiopian sorghum germplasm resources in the national breeding program. In: Plant Genetic Resources of Ethiopia. Engels, J.M.M., Worede, M. (eds). Cambridge University Press, Cambridge, pp. 315-322.

Kimani, W., Zhang, L.M., Wu, X.Y., Hao, H.Q. and Jing, H.C. (2020). Genome-wide association study reveals that different pathways contribute to grain quality variation in sorghum (*Sorghum bicolor*). *BMC Genomics*, 21: 1-19.

Larson-Nath, C. and Goday, P. (2019). Malnutrition in children with chronic disease. *Nutrition in Clinical Practice,* 34: 349-358.

Lee, Y.J., Yang, B., Kim, W.J., Kim, J., Kwon, S.J., Kim, J.H. and Ryu, J. (2023). Genome-Wide Association Study (GWAS) of the agronomic traits and phenolic content in sorghum (*Sorghum bicolo*r L.) Genotypes. *Agronomy*, 13: 1449.

Mekbib, F. (2006). Farmer and formal breeding of sorghum (*Sorghum bicolor* (L.) Moench) and the implications for integrated plant breeding. *Euphytica,* 152: 163-176.

Mengistu, K., Alemu, K. and Destaw, B. (2013). Prevalence of malnutrition and associated factors among children aged 6-59 months at Hidabu Abote District, North Shewa, Oromia Regional State. *Journal of Nutritional Disorders and Therapy,* 1: 2161-0509.

Mundia, C.W., Secchi, S., Akamani, K. and Wang, G. (2019). A regional comparison of factors affecting global sorghum production: The case of North America, Asia and Africa's Sahel. *Sustainability,* 11: 2135.

Reed, J., Ramundo, B., Claflin, L. and Tuinstra, M. (2002). Analysis of resistance to ergot in sorghum and potential alternate hosts. *Crop Science,* 42: 1135-1138.

Rich, P.J. and Ejeta, G. (2008). Towards effective resistance to Striga in African maize. *Plant Signaling and Behavior,* 3, 618-621.

Tesso, T., Ejeta, G., Chandrashekar, A., Huang, C.P., Tandjung, A., Lewamy, M., Axtell, J.D. and Hamaker, B.R. (2006). A novel modified endosperm texture in a mutant high-protein digestibility/high-lysine grain sorghum (*Sorghum bicolor* (L.) Moench). *Cereal Chemistry,* 83: 194-201.

USDA (2023). United States Department of Agriculture. World Agricultural Production. Foreign Agricultural Service (accessed 25/09/2023).

CHAPTER 2

LITERATURE REVIEW

2.1 Biology of sorghum

Sorghum is a naturally self-pollinated crop propagated mainly through seed (Acquaah, 2012). Because the flower of sorghum lacks petals and sepals, it is classified as both incomplete and perfect. The floral panicle forms vary by race (Kumar, 2016), as some are very open or lax (*bicolor* and *guinea*), with considerable space between the panicle branches, while others are densely packed (*caudatum* and *durra*). The inflorescence is a panicle, which has a length ranging from 75 to 50 cm. The spikelet is borne in pairs on branches arranged in whorls. Where one spikelet is sessile, bisexual and fertile, others are pedicelled and male-sterile. The sessile spikelet contains two florets, one perfect and fertile, while the other is infertile. The fertile floret has a membranous lemma, a palea, two lodicules, three stamens and an ovary with two long styles with plumose stigmas (Acquaah, 2012).

2.2 Sorghum diversity

Sorghum is genetically diverse and can vary widely in colour, size, structure and shape. The phenotypic diversity of wild and weedy sorghum was found to be high across 30 populations spread across five geographical regions and between eight different species (Muraya et al., 2010). In many countries, including Ethiopia, different studies have been done on the genetic diversity in cultivated sorghum phenotypic traits (Adugna and Bekele, 2013; Adugna, 2014; Girma et al., 2020; Mengistu et al., 2020; Menamo et al., 2021). Heritable diversity can be studied using phenotypic traits, biochemical traits and molecular markers. Adugna and Bekele (2013) observed that phenotypic traits do not give a reliable estimate of genetic diversity due to environmental effects as the morphological variations are largely governed by environmental factors. However, molecular diversity data may be computed for genomic areas associated with major phenotypic features (Girma et al., 2019).

Germplasm collections should aim to explore as many geographically and climatically diverse areas as possible (Ayana and Bekele, 1999; Girma et al., 2019; Birhan et al., 2022). Analyses of

morphological traits, which are inherited according to Mendelian genetic principles were the first methods for estimating genetic diversity in most crops (Doggett, 1991). When no information about the population structure of a collection is available, the synthesis and categorisation of morphological data into morphological and presumably genetic similarity groups is useful (Marshall and Brown, 1975). Phenotypic diversity indexes of morphological traits and/or multivariate analysis of qualitative and/or quantitative characters have been used to quantify genetic relationships within sorghum (Ayana and Bekele, 1999; Abdi et al., 2002; Geleta and Labuschagne, 2005). Sorghum genetic and geographic diversity cause large variation in both phenotypic and grain quality parameters (Figure 2.1). Based on the genetic diversity resulting from morphological features of their inflorescence, grain and glumes, sorghum is grouped into five basic races (*bicolor, guinea, caudatum, kafir* and *durra*) as well as ten intermediate races that are combinations of the five races (Harlan and de Wet, 1972). Races were geographically separated in their origin and/or domestication, leading to their unique appearances (Figure 2.2). Some studies have begun to classify sorghum based on various methods. Casa et al. (2008) examined nine sorghum sub-populations using an association mapping panel. Brown et al. (2011) studied genetic relationships in comparison to phenotypic racial groupings and sorghum nutritional grain attributes were discovered in five genetic sub-populations that corresponded to race groupings (Sukumaran et al., 2012).

Figure 2.1 Sorghum diversity a) at grain filling and b) at maturity stages.

Bicolor *Guinea* *Caudatum* *Kafir* *Durra*

Figure 2.2 Races of *Sorghum bicolor* - panicle types

As the human population grows, insufficiency in macro- and micronutrients is becoming a severe problem (Mohan Jain and Suprasanna, 2011). In a study by Tesso et al. (2006), sorghum mutants with better protein and starch digestibility and lysine content have been discovered. Deformed protein bodies cause higher protein digestibility in mutants and sorghum as the mutant types exhibit high in vitro raw and cooked protein digestibility (Tesso et al., 2006). Single point mutations and increased lysine content in sorghum grains have been associated with highly digestible sorghum (Wu et al., 2013). Regardless of genotypic background, a mutant allele for the gene encoding pullulanase, an enzyme in the starch metabolic pathway, improved starch digestibility (Mishra et al., 2016). Genetic innovation and environmental variables have combined to produce nutritious grain quality features. Sorghum genetic improvement strategies include increasing the nutritional content and digestibility of sorghum grain proteins (kafirin) (Elkonin et al., 2018). Genetic biofortification of sorghum grain is the most common method for correcting nutritional imbalances in the human diet.

Diverse research has identified a small range of genetic variation for micronutrient such as Fe and Zn (Are et al., 2019). Breeding processes may have significantly reduced genetic diversity in grain minerals. Few studies on the genetic diversity of wild sorghum accessions have been conducted (Aldrich et al., 1992; Cui et al., 1995; Casa et al., 2005). Ayana et al. (2000) examined the level and distribution of genetic variation in 93 wild sorghum individuals representing 11 populations from five geographical regions of Ethiopia, reporting low genetic diversity. Studies have been reported the diversity of wild sorghum in the country (Adugna et al., 2013; Teshome and Feyissa, 2013). In situ population genetic studies using simple sequence repeat (SSR) markers on wild sorghum in Kenya and Mali recently revealed high genetic diversity (Mutegi et al., 2010; Muraya et al., 2011). GWAS is becoming more popular for identifying candidate genes underlying traits of interest. It is becoming a more common approach in trait identification due to its ability to overcome the major limitations of bi-parental populations (Brachi et al., 2011; Mohammadi et al., 2020), particularly with recent advances in high throughput DNA sequencing technologies and large-scale precision phenotyping. Several research studies have explored the extent and distribution of genetic variation in Ethiopian landrace accessions (Girma et al., 2019; Tefera, 2019; Enyew et al., 2022). GWAS discovered loci and single nucleotide polymorphisms (SNPs) linked

to various traits. GWAS identified SNPs with significant association with the various traits studied using the compressed mixed linear model (CMLM) (Girma et al., 2019).

2.3 Utilisation of the grain

Sorghum grain is rich in starch, protein, micronutrients and crude fibre, but low in fat, leading to good functional properties (Abah et al., 2020). Therefore, sorghum is utilised for food, feed and beer manufacturing in Ethiopia. Throughout Africa, sorghum is used to make fermented flatbreads, such as injera, Ethiopia's staple bread and kisra, Sudan's traditional bread. Both injera and kisra are large circular pancake-like flatbreads made with a batter, baked on a clay griddle. Injera features an evenly honeycombed top surface made up of what are referred to as "eyes," each around 4 mm wide and deep and about 50 cm in diameter and 0.7 cm thick (Taylor and Duodu, 2019). Sorghum grain in Ethiopia is used for making injera, which was traditionally made from tef. Fox et al. (2020) identified several Ethiopian sorghum varieties that are better than tef in injera-making quality, based on chemical, image and sensory analysis. Sorghum is utilised for food, feed and alcoholic beverages, malt and beer manufacturing in Ethiopia, with the rest being saved for seed.

Because sorghum grains are an abundant and inexpensive source of starch it can also be used in pharmaceuticals, textiles and paper and it is used to manufacture alcoholic beverages. Traditional alcoholic goods can be classified as beer, "wine," or spirits and come in a wide range of flavours. Alcoholic beverages are made by a fermentation process that produces ethanol and carbon dioxide. Distillation is included as an extra process in the creation of spirits, in which the ethanol is evaporated and then condensed to form high-ethanol liquor (Taylor and Duodu, 2019).

The entire plant is used; sorghum stalks are used as firewood and for building houses, while the leaves are used as animal food. In places impacted by harsh climate conditions, which favour the cultivation of sorghum (as a drought-tolerant crop) over other cereals, per capita consumption of sorghum has increased (Sirany et al., 2022; Khalifa and Eltahir, 2023). Furthermore, due to recent high tef costs, even middle-class households have boosted their sorghum intake, combining sorghum and tef to make injera (Yetneberk et al., 2004; Salaam-Blyther, 2011; Fox et al., 2020). Sorghum is a near equivalent for tef, which is useful when tef prices fall, which also applies to sorghum consumption.

2.4 Sorghum grain quality

Sorghum quality is determined by aesthetic and cooking characteristics, processing attributes and consumer acceptability; all these factors contribute to the quality of sorghum. Furthermore, sorghum's nutritional features include protein and starch digestibility, nutrient bioavailability and anti-nutritional factors (tannin and phytic acid) (Mezgebe et al., 2018; Fox et al., 2020). Most cereal species, such as wheat, maize and sorghum, are dietary staples and are important commercial commodities, but they lack certain critical amino acids, such as lysine, threonine, tryptophan and methionine (Sudheesh et al., 2022).

Seed hardness, weight and size are all key factors in determining the quality of sorghum grain (Mezgebe et al., 2018; Tao et al., 2020). Because of its impact on processing, uniformity of these traits is important in sorghum grain quality assessment, in addition to average values for hardness, weight and size. The essential grain quality attributes that contribute to customer preferences and acceptance are grain size, shape, gloss and colour. Grain size correlates significantly with molecular mechanisms, offering a pathway to explore the underlying genetic composition (Tao et al., 2020). Moreover, grain size, grain shape and lustre are also important quality characters that determine the market price of sorghum grain. Sorghum varieties with round and large grain sizes have higher grain yields. Tao et al. (2020) observed the significant importance of genetic effects in breeding programmes and selection procedures to develop new sorghum cultivars with enhanced grain size and round grain shape.

2.5 Sorghum grain structure

Kernel physicochemical parameters determine the processing properties and product quality of cereal grain. Within cereal grain species, differences in grain morphological structures and chemical composition have been identified. These changes have a significant impact on the nutritional quality of cereal grain products (Koehler and Wieser, 2013). The pericarp (outer layer), endosperm (storage tissue) and germ (embryo) are the three primary elements of the sorghum grain (Figure 2.3). Serna-Saldivar and Espinosa-Ramírez (2019) found that, of the total grain, the endosperm makes up 84.2%, the germ 9.4% and the pericarp 6.5%. The shape, size, endosperm, germ, pericarp and pericarp colour of sorghum kernels are all genetically determined. Some

sorghum varieties (tannin sorghum varieties) contain a darkly pigmented layer beneath the pericarp termed the testa or sub-coat (Sen et al., 2023).

The outermost endosperm (aleurone layer), the peripheral endosperm (sub-aleurone layer), the corneous endosperm and the floury endosperm are the four types of endosperm that have been identified (Serna-Saldivar and Espinosa-Ramírez, 2019). The aleurone is made up of a single layer of rectangular cells close to the testa, or tube cells. A thick cell wall, many proteins (protein bodies, enzymes), ash (phytin bodies) and oil are all present in the cells (spherosomes). The peripheral endosperm (sub-aleurone layer) contains higher protein content but smaller starch granules than the other parts of the endosperm; it affects the processing and nutrient digestibility (Duodu et al., 2003; Serna-Saldivar and Espinosa-Ramírez, 2019). Therefore, the protein in the peripheral endosperm would provide information on the endosperm-specific genes. The aleurone layers have no starch, but contain minerals, protein, phytic acid, oil and vitamins (Rooney and Murty, 1982; Serna-Saldivar and Espinosa-Ramírez, 2019).

The corneous and floury endosperm cells are composed of starch granules. The corneous endosperm contains starch granules fixed in a dense protein matrix. The endosperm matrix protein is comprised mainly of glutelins and prolamins (kafirins) (Serna-Saldivar and Espinosa-Ramírez, 2019). The cells in the floury endosperm are loosely packed. The protein matrix does not hold the spherical starch granules together (Rooney and Murty, 1982). Endosperm texture refers to the floury endosperm and determine the functional qualities of sorghum (Ezeogu et al., 2005) as well as sorghum malting quality (Chiremba et al., 2011). The embryonic axis and the scutellum are the two major components of the germ. The new plant is contained within the embryonic axis, which is separated into radicals and plumules. The radical is responsible for the formation of basic roots, whereas the plumule is responsible for the formation of leaves and stems. The cotyledon of the sorghum caryopsis is the scutellum, the germ transport reserve tissue, containing a high concentration of oil, protein, enzymes, minerals and acts as a link between the endosperm and the germ (Rooney and Murty, 1982; Serna-Saldivar and Espinosa-Ramírez, 2019).

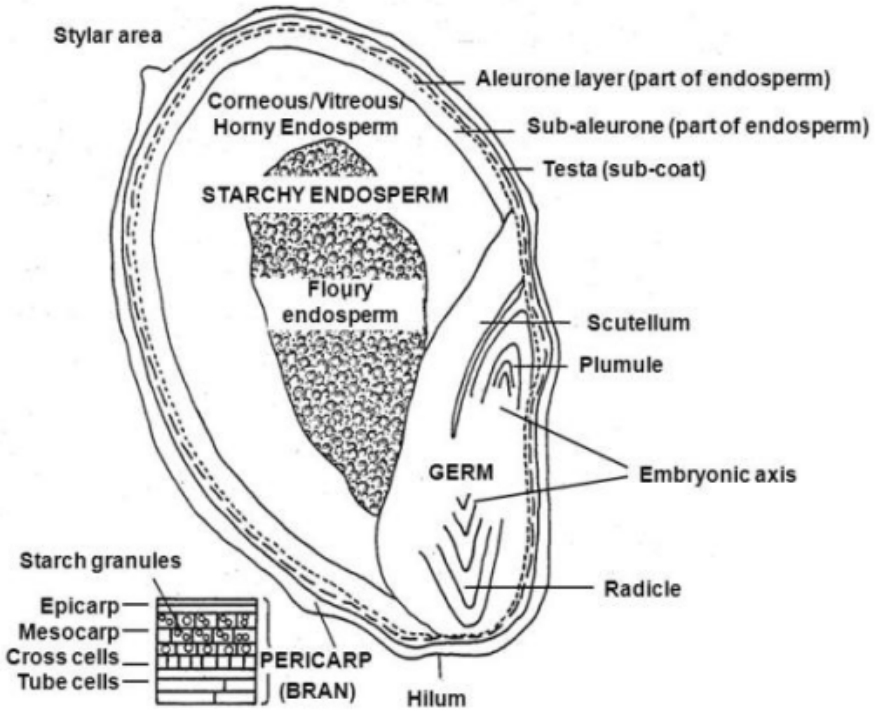


Figure 2.3 Longitudinal section of a sorghum kernel (Belton and Taylor, 2002).

2.6 Chemical composition of sorghum grain

Sorghum grain composition and functional qualities have been extensively researched. According to Rhodes et al. (2017), protein content in global sorghum germplasm range from 8.1 to 18.8%, fat content from 1.0 to 4.3% and starch content from 61.7 to 71.1%. When compared to other cereal grains, sorghum grain is known for its hardness. Sorghum grain with modified endosperm texture has improved nutritional quality and digestibility due to its increased lysine content. Breeding resulted in sorghum cultivars with modified hard endosperm with high lysine content and increased protein digestibility (Tesso et al., 2006). Sorghum contains complex carbohydrates (starch and dietary fibre) that are slowly digested and provide a feeling of fullness when consumed, while suppressing appetite. Sorghum is a great source of nutrients for people of all ages. Fe and Zn are essential minerals for normal human growth, yet there is a deficiency of these two micronutrients in underdeveloped nations (Are et al., 2019).

Starch

Starch is the dominant carbohydrate in sorghum grain and is stored as granules in the endosperm, which provides energy for humans and animals, and cellulose, which makes up many plant structures. The quality and quantity of starch present in sorghum are important quality parameters that can influence consumer acceptance of the end product (Pushpamma, 1993). Rhodes et al. (2017) reported that starch concentration of global sorghum germplasm ranged from 61.7 to 71.1%. Starch is found in the form of granules and makes up 56 to 76% of sorghum flour, depending on genetic and environmental factors (Bean et al., 2019). Starch granules are made up of a linear carbohydrate called amylose (20-30%) and a highly branched polymer called amylopectin (70-80%). Amylose (106-108 kDa) is a linear sugar made up of 1,6-linkage branching and 1,4-linked sugar units per 200 to 10 000 units. Amylopectin (108-109 kDa) has a higher branched structure than amylose-1, with 6-linkage branch points every 20 glucose units (Tetlow, 2011).

The ratio of amylose to amylopectin in sorghum grain starch is one of the most important factors determining its use as food, feed or biofuel. Amylose creates a compound with lipids after heating, increasing starch viscosity and decreasing fermentation and digestibility (Zhu, 2014). Wang and Copeland (2013) found that amylopectin does not form a lipid complex and has a lower viscosity, extending the shelf life of bread, cakes and pastes. Amylopectin improves grain digestibility and ethanol fermentation efficiency. Sorghum starch physicochemical properties are improving, making it more valuable for food and industrial applications. According to Bemiller (1997), molecular and plant breeding approaches such as selection, mutation, crossover genetics, physicochemical and enzymatic alterations can boost starch content. The application of genomic-based breeding methods including SNP and genotyping-by-sequencing (GBS) in sorghum has the potential to significantly accelerate selection gain. Bekele et al. (2013) showed that molecular breeding has advanced and improved the effectiveness of breeding techniques. Waxy genes produce starch that has high amylopectin with improved protein and starch digestibility in sorghum (Are et al., 2019). Starch structures and enzyme activities are involved in starch biosynthesis and many genes for these enzymes have been mapped. The enzymatic activities of starch biosynthesis require the coordinated activities of several enzymes like starch synthase (SSS), granule-bond

starch synthase (GBSS), starch branching enzymes (SBE) and starch de-branching enzymes (DBE) (Bing et al., 2014).

Protein

The protein quality of sorghum is related to the protein fraction distribution in grains and the attributed nutritional composition (Serna-Saldivar and Espinosa-Ramírez, 2019). Sorghum endosperm proteins can be found in two forms: a matrix and protein bodies that are encased in the matrix. Albumins, globulins, kafirins, cross-linked kafirins and glutelins are several types of proteins found in sorghum (De Mesa-Stonestreet et al., 2010). Kafirins are the most abundant storage proteins in sorghum endosperm, accounting for 77-82% of total protein content (Duodu et al., 2003). Kafirins are prolamin storage proteins with low quantities of certain amino acids, particularly lysine, a problem common to cereal grains (Taylor and Schüssler, 1986). In poor countries, sorghum is a reliable food crop. When kafirin, albumin, glutelin and globulin proteins are restricted in critical amino acids (lysine, tryptophan and threonine), this may result in an imbalanced diet. After starch, protein is the second most abundant component in sorghum grain (6-18%) (Duodu et al., 2003). Sorghum flour protein content can range from 6 to 20%, depending on genetic, environmental and management influences (Bean et al., 2019).

Sorghum proteins are found in the endosperm (80%), germ (16%) and pericarp (3%) of the grain (Taylor and Schüssler, 1986). The sub-aleurone layer of the peripheral endosperm has a higher protein concentration than the rest of the endosperm. As a result, the peripheral endosperm provides insight into the endosperm's unique gene function. Taylor and Schüssler (1986) reported that sorghum grain composition is affected by genotype and environmental factors.

Sorghum protein digestibility improvement

The use of sorghum in both food and feed applications has been limited, largely due to the low digestibility of its proteins. Improving sorghum protein digestibility is one of the major tasks and objectives of sorghum breeding programmes and some success has been achieved in the past (Tesso et al., 2006).

Food processing procedures affect nutrient bioavailability in cereal grains by modifying the effects of anti-nutritional agents or chemical alterations to proteins and their interactions with other molecules (Afify et al., 2012). The sorghum protein bodies and matrix are resistant to enzymatic digestion due to their physical and chemical properties (Wong et al., 2010; Duodu and Dowell, 2019). After wet cooking, sorghum protein digestibility ranged from 36.4 to 74% (Hussain et al., 2020), while Ezeogu et al. (2005) showed that in vitro protein digestibility significantly decreased in wet cooking, from 59.1 to 30.5%. Dry heating (e.g. popping) has a deleterious influence on protein digestion as well, but to a lesser extent (Opazo-Navarrete et al., 2019). The bioavailability of sorghum protein as well as in vitro protein digestibility are further affected by food processing procedures such as fermentation, dry roasting and malting or germination (Yetneberk et al., 2005; De Mesa-Stonestreet et al., 2010; Weerasooriya et al., 2018). Findings suggest that modifying food processing procedures and selecting appropriate genotypes for specific food products could help to improve in vitro protein digestibility and bioavailability.

Because sorghum products are gluten-free, they lack the elastic gluten network that is necessary for expanded texture development in staples like bread. Due to poor amino acid composition and low protein digestion, sorghum protein nutritional scores are often low. Both environmental and internal variables influence protein digestion. Physical inaccessibility due to trapping non-intact cell structures and the presence of anti-nutritional substances are examples of external influences. The amino acid sequence of proteins, as well as protein folding and cross-linking, are the most important internal factors. Food processing is usually done to preserve the nutritional value of the food. Linking of anti-nutritional elements such as tannins and phytic acid with sorghum components, particularly its protein, makes it less digestible for humans and monogastric animals than other cereals. To increase the nutritional quality of sorghum and fully utilise its potential as human food or animal feed, these unwanted components must be removed or reduced (Are et al., 2019). Proline-rich proteins bind sorghum tannins more effectively than other proteins. Proteins of good quality are easily digestible and contain essential amino acids in sufficient quantities to meet human needs. Digestibility is also strongly associated with the overall kafirin content (Wu et al., 2013; Elkonin et al., 2018).

Mineral elements

Micronutrient deficiency is one of the world's most serious problems, especially in Africa. Dietary diversity and food biofortification have been used to combat micronutrient deficiency (Ohanenye et al., 2021). Sorghum is high in minerals, which are found in the pericarp, aleurone layer and germ. Sorghum is a rich source of potassium (K) and a good supply of magnesium (Mg), Fe, Zn and copper (Cu) (Kumar et al., 2022). Sorghum has a wide adaptability range, is drought tolerant and has low variability in Fe and Zn content. As a result, developing varieties with higher micronutrient concentrations can help meet nutritional requirements. The grain Fe concentration in elite parents generally ranges from 20 to 40 ppm, while the grain Zn concentration ranges from 20 to 30 ppm (Are et al., 2019). The aleurone layer and embryo (particularly the scutellum) have the highest Fe and Zn concentrations, while the starchy endosperm has the lowest (Are et al., 2019).

The mineral content of sorghum grain varies greatly. Environmental variables have a greater impact on the mineral content of this cereal grain than genetic factors. Micronutrient deficits are a major public health issue with considerable health and nutritional implications (Ananda et al., 2020). Zn and Fe nutritional deficiencies have received a lot of attention in recent decades, especially in developing nations where many people rely on grains and legumes as their primary sources of these nutrients.

Biofortification is still being scaled up and more work is needed to achieve the overall goal of eradicating Fe and Zn deficiency in the world population and assuring nutritional security (Stangoulis and Knez, 2022). Plant susceptibility to biotic and abiotic stressors is also affected by micronutrient deficiency. However, the plant genotype plays a big role in mineral concentration and each mineral element has intricate interactions with a variety and different environmental conditions (Nakandalage and Seneweera, 2018).

Anti-nutritional factors

Sorghum contains phytochemicals such as phenolic acids, flavonoids and tannins of which tannins are considered as major anti-nutritional factor. Sorghum tannins are made up of oligomers or polymers of flavan-3-ol and flavan-3,4-diol that are mostly linked by B-type connections. In

addition, monomers (primarily catechin) and dimers (mostly procyanidin B1), which have a low molecular weight, are present in modest amounts in sorghum grain (Awika and Rooney, 2004; Xiong et al., 2019). The tannin concentration of sorghum grain was found to vary significantly between cultivars and values ranging from 2.3 to 8.8 mg/g, with an average of 5 mg/g were reported (Abdelhalim et al., 2019). Sorghum has a total polyphenol concentration of 1.7 to 102.6 g/kg, which includes tannins, proanthocyanidins, flavonoids and phenolic acids. In comparison, the polyphenol content of maize and wheat is low, while barley has a 13.5 g/kg polyphenol content (Bravo, 1998). One of the most important nutritional aspects influenced by tannins is protein digestibility. Tannins produce a protein-tannin complex, which decreases protein digestibility, delays growth and sometimes interferes with amino acid synthesis (Tapiwa, 2019). Tannins are divided into two groups: hydrolysable and condensed tannins (Okuda and Ito, 2011). Tannins can be easily treated if they are hydrolysable. Their effects can be reduced by soaking sorghum grains allowing it to absorb water. Condensed tannins can be removed through dehulling or heat treatment to remove the sorghum coat. If tannins in sorghum can be easily removed, sorghum will become one of the best sources of animal feed (Tapiwa, 2019). The use of high-tannin forages is lethal and even at low levels, their regular use reduces dairy, meat and egg productivity, reproduction and quality. Tannins have been found in nature as monomers or low-molecular-weight oligomers in plants such as grapes, tea and legumes (Terblanche, 2017; Watrelot and Norton, 2020). Sorghum tannins, on the other hand, are in a condensed state with a high molecular weight and a high degree of polymerisation, which is unusual among the major grains (Wu et al., 2012).

In sorghum, phenolic acids, 3-deoxyanthocyanidins and condensed tannins make up the majority of the phenolic components. Rhodes et al. (2017) reported on the phenolic contents of the sorghum races. The highest mean content of phenolic acids is seen in *bicolor* and *guinea-caudatum*. *Caudatum* has a lower level of phenolics than the other botanical races. *Bicolor-durra* accessions have high mean amounts of phenolics (Rhodes, 2014). Sorghum phenolic acids have been demonstrated to have high antioxidant action in vitro and consuming sorghum whole grains may enhance gut health and lower the risk of chronic diseases (Xiong et al., 2019). Sorghum with high tannin content provide agronomic benefits, such as protecting the plant against diseases and birds and have been widely produced in some underdeveloped areas with food security issues.

Phytic acid is another anti-nutritional compound found in seeds. Phytic acid can be found in most cereal grains, legumes, nuts, oilseeds, tubers, pollen, spores and organic soils. Phytic acid accounts for about 1-2% of the weight of cereal seeds and it can sometimes reach 3-6% (Febles et al., 2002). According to Syed (2019), the phytic acid concentration of grains ranged from 0.2 to 2.4 mg/g, with an average of 1.1 mg/g. Phytic acid is the primary phosphorus storage compound in plant seeds, accounting for up to 80% of total phosphorus in seeds (Lopez et al., 2002). In many plant species, 90% of phytic acid is found in the aleurone layer and only 10% in the embryo. Many factors can influence the phytic acid content and phosphorus availability in cereal grain, including genetics, environmental fluctuations, location, irrigation conditions, soil type, season and fertiliser application (Dost and Tokul, 2006). Phytic acid has long been regarded as an anti-nutrient due to its strong ability to form complex multi-charged metal ions, particularly Zn, Ca and Fe. As a result, consuming large amounts of phytic acid-rich foods may result in a deficit in the absorption of some dietary minerals (Reddy and Pierson, 1994).

2.7 Environmental factors affecting nutritional quality

Crop yields and nutritional security are heavily reliant on weather patterns and the changing climate is placing the bulk of food produced for human consumption at risk. As a result of climate change, soil quality is deteriorating, mostly owing to the disturbance of soil microbial communities, which harms the breakdown of organic pollutants and soil organic matter, as well as nitrogen fixation (Ai et al., 2021). Despite attempts to boost global food availability, which is critical for food and nutritional security, the worldwide burden of malnutrition and micronutrient deficiencies remains severe, especially among low-income groups and is intrinsically tied to climate change (FAO, 2018).

Developmental processes, morphological traits and fundamental physiological responses to varied environments all have distinct effects on plants. Increased carbon dioxide (CO_2) is well known to improve plant growth and production by speeding up photosynthesis while also boosting crop water uptake efficiency (Guo et al., 2013; Li et al., 2018). However, increased plant growth at increasing CO_2 levels is offset by decreased grain quality responses, which have been reported in many plant species (Morison and Lawlor, 1999). This indicates how rising CO_2 levels alter the balance of agricultural metabolic pathways, mineral intake and nutrient availability.

Environmental factors such as pH, air temperature, sunlight, rainfall, organic matter and soil texture have the potential to influence nutrient concentration in crops and must be considered when investigating micronutrient content variation in crops (Nakandalage and Seneweera, 2018).

Environmental conditions may also influence the concentration of provitamin A carotenoids in maize cultivars (Mengesha et al., 2019). Large variability for sorghum grain micronutrients (Fe and Zn), as well as the presence of genotype by environment interaction (GEI) for micronutrients, impacts the expression of these traits (Phuke et al., 2017). Grain physical characteristics (size and hardness) are controlled by genotype, location and their interactions. Tao et al. (2020) demonstrated that cross-environment correlations of these grain features are relatively high within the same population and heritability of the variables is moderate to high, indicating a strong genetic influence. To find and promote superior genotypes, stability assessments of potential genotypes must be reproducible across years. The use of factor-analysis in conjunction with mixed models for multi-environmental trial (MET) analysis of imbalanced data is crucial for determining the relative stability and GEI of varieties (Cullis et al., 2010; Smith et al., 2015). Sorghum protein level and composition vary by cultivar due to genetic and agronomic factors (water availability, soil fertility, temperatures and environmental variables during grain growth) (Taylor and Schüssler, 1986). Positively interacting varieties can be employed for specific agroecological conditions in the target areas and are thus best adapted to those conditions. For many cereal grains, including sorghum, geographical and/or environmental factors (such as soil characteristics) influence protein content (Kaufman et al., 2018).

The availability of minerals in the soil has a considerable influence on plant nutrition. Low nutrient availability typically has a limited influence on plant photosynthetic rates, perhaps resulting in less carbon available for secondary compound formation (Dong et al., 2018). There are considerable environmental effects on both yield and quality features. Because the crop is cultivated under a variety of agro-climatic conditions that impact the nutritional makeup of the grain, the essential amino acid composition of sorghum protein also varies greatly (Fano, 2017). Human selection for various food applications altered grain composition delivery patterns among genetic groupings. It is also likely that some of the changes in grain composition between genetic groups are caused by adaptation to environmental variables (Rhodes et al., 2017).

2.8 Drought effects on grain quality traits

Drought stress has been shown to reduce grain quality in cereals such as wheat, barley, maize and sorghum (Safian et al. 2022; Ullah et al., 2023). The impact of drought stress on sorghum productivity during each growing stage can vary depending on the severity, duration and timing of water deficit, as well as the sorghum variety's tolerance to drought (Assefa et al., 2010; Yahaya and Shimelis, 2022). Extended periods of water scarcity can limit sorghum's ability to complete its life cycle, resulting in reduced yield and quality (Bing et al 2014; Sehgal et al., 2018). Repeated exposure to drought stress can weaken sorghum plants, making them more susceptible to other stressors such as pests and diseases (Funnell-Harris et al., 2023). Drought stress during critical stages such as flowering, grain filling and ripening can lead to substantial yield losses (Sarshad et al., 2021). Water deficit during pre-flowering stage can lead to a decrease in the number of flowers produced per panicle and limit the potential for successful pollination and seed set in sorghum (Lehrer, 2022). Water scarcity during flowering and pollination stage can result in incomplete pollination, poor pollen viability and reduced fertilization rates, leading to a decrease in grain set and potential yield loss (Raymundo et al., 2023). Insufficient soil moisture during grain filling and ripening critical period can result in poor grain filling, smaller kernel size and reduced grain weight, ultimately leading to lower yields and inferior grain quality in sorghum (Raymundo et al., 2023).

Heat and drought have been found to alter protein production and cross-linking in wheat and maize. Drought stress lowers grain filling rate and shortens filling duration during flowering and grain filling, decreasing grain size and output (Sehgal et al., 2018). It reduces grain size by decreasing endosperm cell division and the amount of starch granules in the grain. According to Sehgal et al. (2018), drought stress reduce the availability of sucrose from leaves and hampered enzymatic activity in grains. Drought has a large impact on the physicochemical grain characteristics of cereals, particularly sorghum. Drought stress influences grain starch concentration, endosperm cell numbers and grain size and has a significant impact on starch production. Drought can impair protein and starch digestibility (Impa et al. (2019), resulting in insufficient nutritional absorption from drought-stressed sorghum. Drought decreases grain protein digestibility and micronutrient composition (Zn, Fe, Mn and Cu), while increasing kernel hardness and total protein in grain (Impa et al., 2019). Drought reduces sorghum feed digestion (in vitro organic matter digestibility)

(Somegowda et al., 2021). Drought stress at anthesis resulted in lower total starch, amylose and amylopectin accumulation, which is related to poor enzyme activity on sugar nucleotide precursors (Bing et al., 2014). SSS, GBSS, SBE and DBE enzymatic activity in sorghum grains were reduced to varying degrees due to drought stress, inhibiting starch synthesis during grain filling (Bing et al., 2014).

Drought and heat stress reduce grain filling rate and duration during flowering and grain filling, reducing seed size and quantity, influencing commercial attributes such as seed weight and seed quality. It reduces grain size by inhibiting cell proliferation in endosperm cells and the number of starch granules in grain (Dalling, 1985). The production and transfer of photo assimilates, as well as the import of precursors for biosynthesis of seed reserves, minerals and other functional elements, all have an impact on seed filling. These processes are especially susceptible to drought due to the participation of a varied array of enzymes and transporters present in leaves and seeds (Figure 2.4; Sehgal et al., 2018).

Figure 2.4 illustrates the movement of sucrose and other nutrients from the source (endosperm/cotyledon) to sink during seed growth (embryo). Carbon, nitrogen, phosphorus and other minerals obtained from the hydrolysis of endosperm/cotyledon stored nutrients are transferred to the embryo for development and mobilised toward the assimilate-transport pathway into growing seeds, as well as the starch and sucrose synthesis pathways. Both drought and heat can have an impact on seed filling by altering any of these or many processes, causing them to either accelerate (as in heat) or disrupt seed filling (as in heat and drought) (Sehgal et al., 2018). A summary of previously published results on the effect of drought stress on the quality of sorghum genotypes at different growth stages is presented in Table 2.1.

2.9 Genetics of kernel colour

Kernel colour has a key role in determining the nutritional content and consumer appeal of sorghum. Sorghum kernel colour varies greatly. Xiong et al. (2019) reported that sorghum kernel colour has been categorised into five types: white, yellow, red, black and brown. The kernel colour is related to phenolic compounds in the bran layer and is controlled by the *R* and *Y* genes (pericarp colour), the *Z* gene (pericarp thickness), the presence of *B1* and *B2* genes (pigmented testa), the *S*

gene (spreader), the *Tp* gene (testa colour) and the *I* gene (the colour intensifier) (Dykes et al., 2009; Dykes et al., 2013; Khoddami et al., 2023).

White sorghum possesses a *y* recessive gene as well as *RRyy* or *Rryy* alleles, which result in a white pericarp. White sorghum has low amounts of total phenolics, tannins and 3-deoxyanthocyanidins. Yellow sorghum contains a recessive *r* gene and a dominant *Y* gene, with *rrY* yielding a yellow pericarp. Yellow sorghum is high in flavanones and low in phenolics. Sorghum with a red or black pericarp possesses both dominant *R* and *Y* genes, with *RY* yielding a red or black pericarp. The pericarp of red and black sorghum contains a high concentration of phenolic compounds. Red sorghum has a somewhat higher phenolic content than white sorghum, but no tannins (Dykes et al., 2009; Dykes et al., 2013; Khoddami et al., 2023).

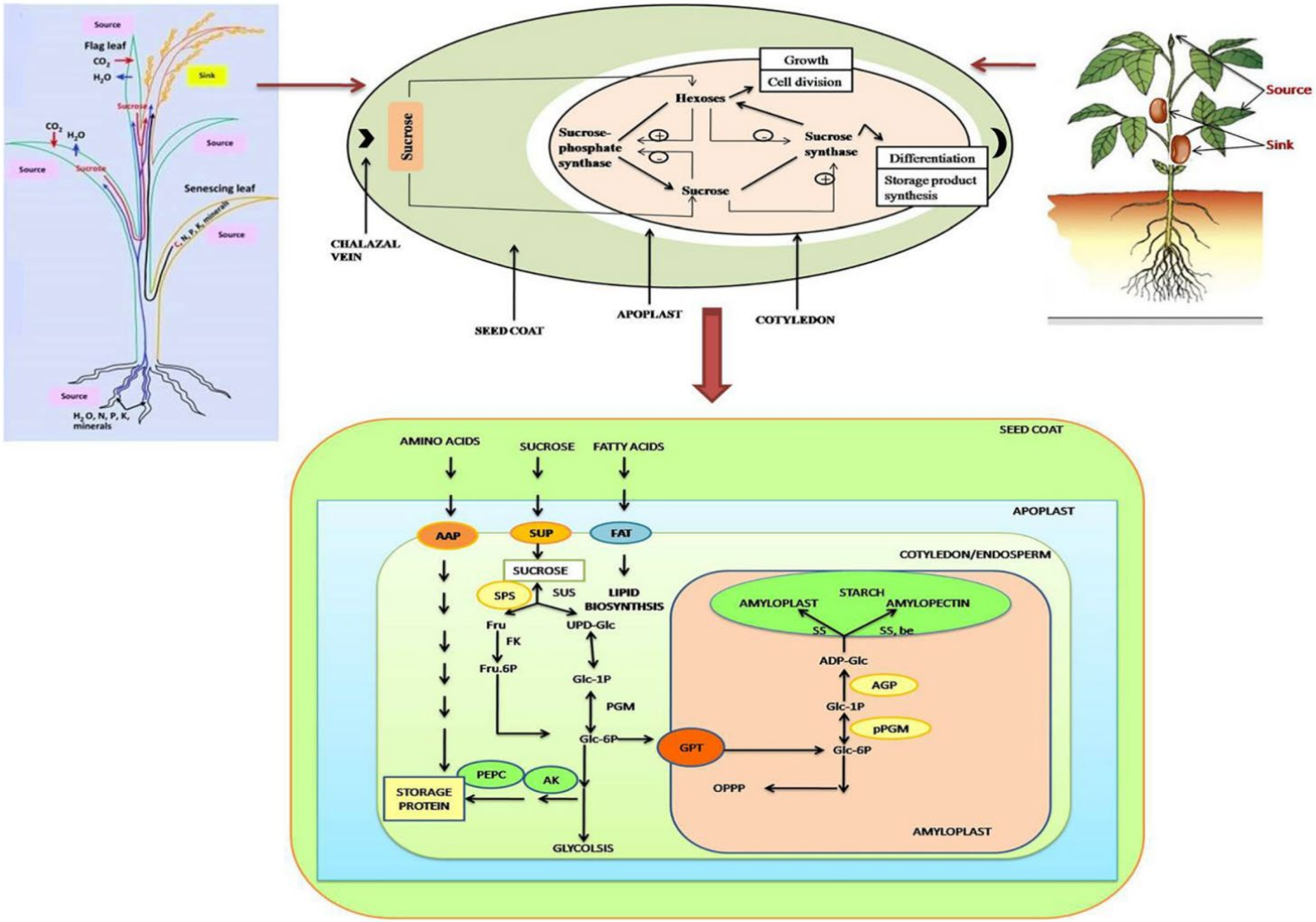


Figure 2.4 Schematic representation of various processes during the seed filling stage in both monocot (top left) as well as dicot (top right) plants (Sehgal et al., 2018).

SUT = sucrose transporter; AAP = amino acid protein; PEPC, AK, PEP = carboxylase and/or aspartate kinase; AGP = ADP glucose pyro-phosphorylase; GPT = plastidic glucose-6-P translocator; pPGM = plastidic phosphoglucomutase.

Table 2.1 Summary of previously published studies on the effect of drought stress on the quality of sorghum genotypes at different growth stages

Experiment	Genotype name	Stress level and conditions	Results	Genotypes that had high grain qualities under stress	Reference
Effect of drought stress during flowering stage on starch accumulation and starch synthesis enzymes in sorghum grains	2 sorghum hybrids (Tieza 17 than Liaoza 11)	7^{th}, 14^{th}, 21^{st} and 28^{th} day after flowering	Drought stress during the grain filling and flowing stages lowered starch synthesis enzyme activity, resulting in lower accumulations of starch, amylase and amylopectin in grains	Tieza 17	Bing et al., (2014)
Water deficit and heat stress-induced alterations in grain physicochemical characteristics and micronutrient composition in field-grown grain sorghum	24 sorghum genotypes	Two water stress conditions: flowering and grain filling stages	Drought stress decreased protein digestibility and micronutrient composition Drought stress increased kernel hardness and total proteins in grain	SC372 and RTx7000	Impa et al. (2019)
Nutritional composition of sorghum [*Sorghum bicolor* (L.) Moench] genotypes cultivated without and with water stress	100 sorghum genotypes	Post-flowering water stress	Post-flowering water stress reduced carbohydrate, protein and ash contents	**Carbohydrates:** BR007B, SC59 and SC1033 **Proteins:** SC325, SC320 and SC124 **Ash:** SC224, SC566, SC467 and SC1356	Queiroz et al. (2015)
The potential of some sweet sorghum (*Sorghum bicolor* L.) genotypes under two water regimes for sugar and bio-ethanol production	10 sweet sorghum varieties	Once every 15 days till the heading stage, plants exposed to water stress	Drought stress reduced protein digestibility and the content of micronutrients Drought-enhanced kernel hardness and overall protein content in grain Water stress after blooming lowered carbohydrate, protein and ash levels Water stress lowered yields of stalk, juice, sugar and ethanol	ICSSH 30 was superior in drought-tolerance-associated traits and high sugar and bio-ethanol potential yield under both water regimes.	Alhajturki et al. (2012)
Effect of moisture stress on morphological and yield attributes of four sorghum varieties	4 hybrid sorghum varieties	The stress was exposed in 2-day intervals starting from the tilling stage	Water stress decreased starch and protein contents	Varieties BD 740 and BD 731 had higher starch and protein contents	Hossain et al. (2016)
Morphological and biochemical responses of *Sorghum bicolor* (L.) Moench under drought stress	3 sorghum varieties (Sepideh, Kimia, and Payam)	The stress was at pollination, seed milky, and seed doughy stages	Drought stress negatively influenced morphological and yield-related traits	Payam variety is better to use in both drought stress and non-stress conditions	Sarshad et al. (2021)

Because of the presence of polyphenols, some sorghum has a red hue. Under adverse conditions such as humidity and rain, these polyphenolic chemicals move from the glumes to the grains. During this process, the grain becomes discoloured, compromising the end-use quality of the final products (Hikeezi, 2010). Because the red pericarp transforms to black under solar radiation during maturation, black sorghum is genetically red and a subspecies of red sorghum. The pericarp of black sorghum contains a high concentration of phenolic compounds, notably 3-deoxyanthocyanidins (Dykes et al., 2013). Brown sorghum, often known as tannin sorghum, contains coloured testa and a high concentration of condensed tannins (Awika and Rooney, 2004; Dykes et al., 2009). The existence of the testa in sorghum is controlled by the dominant testa genes *B1* and *B2*, with *B1* and *B2* generating brown pericarp and testa hues (Earp et al., 2004).

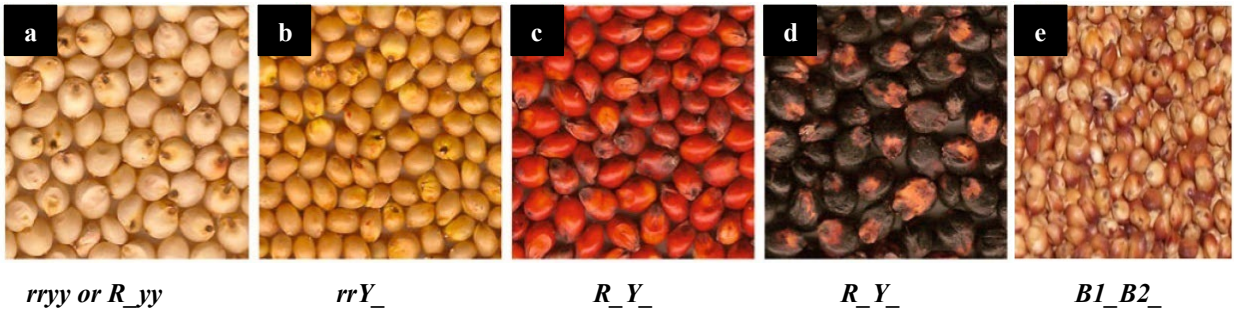


Figure 2.5 Genetics of sorghum pericarp colour (*R* and *Y and B1* and *B2* genes): (a) White (*rryy* or *R_yy*); (b) yellow (*rrY_*); (c) red (*R_Y_*); (d) black (*R_Y_*) and (e) brown (*B1_B2_*).

2.10 Association mapping for grain quality

GBS is a powerful technique that combines genotyping with next-generation sequencing. It offers a practical solution for genotyping diverse panels, addressing missing data and advancing association mapping studies (Fu and Yang, 2017; Wang et al., 2020). Association mapping is a technique for identifying genetic variants associated with phenotypic traits by analysing the patterns of genetic variation within a population (Mural et al., 2022). Advantages of association mapping over traditional QTL linkage studies is its ability to capture natural genetic variation and achieve higher mapping resolution. This will result in an extremely well-characterised genetic population (Poland and Rife, 2012). The use of genomics to investigate variability, inheritance and trait associations to enhance nutritional quality is growing. Sorghum has a small genome (730 Mb) with a whole genome sequence being available (Paterson, 2013). Sorghum is amenable to the use

of genomic breeding approaches. In recent years, there has been a rapid decrease in the cost of genotyping for GWAS for the identification of SNPs using high-density SNP array technologies (Batley and Edwards, 2007; Negro et al., 2019) or GBS (Elshire et al., 2011; Morris et al., 2013) or next-generation sequencing technologies (Metzker, 2010). This has paved the way for genomic selection or predictive breeding tactics (Jannink et al., 2010).

In plants, SNPs are used for a variety of molecular genetic marker applications. High-resolution genetic map creation, linkage disequilibrium-based association mapping, genetic diagnostics, genetic variability assessment, varietal identification, phylogenetic studies and characterisation of genetic resources are among these uses (Batley and Edwards, 2007). Recent advancements allow such techniques to significantly accelerate selection gain and increase breeding effectiveness (Bekele et al., 2013; Krishnappa et al., 2021). The genetic control of quality characteristics, as well as mutation research, have greatly increased our understanding of nutritional quality aspects in sorghum. Mutations in sorghum genes, such as the waxy gene, produce starch with high amylopectin content and low amylose content, as well as increased protein and higher starch digestion (Griebel et al., 2019).

Starch content has been genetically mapped in many cereal crops, such as maize (Cook et al., 2012), rice (Bing et al., 2014; Yang et al., 2019) and wheat (McCartney et al., 2006). Several SNPs linked to sorghum starch concentration have been found (Sukumaran et al., 2012; Boyles et al., 2017). There are 27 genes encoding kafirin proteins, with 22 gene clusters coding for kafirins on chromosome 5, two on chromosome 2, one on chromosome 9, one on chromosome 10 and one additional gene on chromosome 8 (Xu and Messing, 2008; Duressa et al., 2018). Even though these kafirin genes and gene sequences are well-known, research on gene variants and allelic combinations, as they relate to protein digestibility, is limited. Cremer et al. (2014) detected no relationship between allelic variants of β-, γ-, or δ-kafirins and grain protein digestibility.

In general, the candidate gene association mapping technique supplements genome-wide association research and classical linkage mapping. This strategy can establish the gene–trait association at the population level by employing a significant number of SNPs in conjunction with a thorough selection of candidate genes (Sukumaran et al., 2012). A SNP in the putative gene *SSIIa*

on chromosome 10 is linked to kernel hardness and accounts for 8% of the trait's variation (Sukumaran et al., 2012). Studies on kernel hardness in sorghum, wheat and rice suggests that starch concentration and the distribution of proteins and lipids on the surface of starch granules are essential determinants of grain hardness (Chen et al., 2012). Sorghum grain hardness is related to vitreousness and vitreousness is related to amylose content, which is a significant component of starch.

QTL discovered by GWAS for sorghum protein, fat and starch properties (Table 2.2) are used to identify likely genes inside a GWAS QTL region. A putative *alpha-amylase 3* gene, previously associated with grain composition characteristics, has been identified as a key possibility for protein and fat variation (Rhodes et al., 2017). Tao et al. (2020) detected 111 orthologues in the sorghum genotype for the suspected causative genes underlying the grain size QTL, 109 genes influencing grain size in rice and maize and the hypothesised causative genes underpinning the grain size QTL. According to Tao et al. (2020), 16 of the 111 candidate genes are identified within a 0.1 cM window of a grain size QTL. *SbGS3*, the ortholog of the rice grain size gene, *GS3* (Mao et al., 2010), was shown to be co-located with *qGS1.9*, a QTL for grain size (length, weight, volume, width and thickness) discovered in both the sorghum diversity and backcross-nested association mapping populations (Tao et al., 2020).

Table 2.2 Quantitative trait loci mapped for grain quality traits in sorghum

Trait	No. of sorghum population	No. of markers	Type of markers	Chr/LG	Method	Candidate genes	References
Kernel quality traits (protein, starch, fibre, fat and tannin)							
Protein	245 accessions	85 585	SNP	2, 5	MLM	*Sobic.002G217100*	Liu et al. (2018)
	390 accessions	404 627	SNP	2, 4	MLM	*Sb02g023790*	Xiong et al. (2019)
	40 accessions (MW467, ZMB6986 and TZ3966)	226	SSR	7	MLM	*SbAGB02*	Ng'uni et al. (2012)
	390 accessions	404 627	SNP	2, 4	MLM	*AMY3, Sb02g023790*	Rhodes et al. (2017)
	390 RILs		SNP	1		*Sobic.001G059100*	Boyles et al. (2017
Starch	40 accessions (ZMB3947 and ZMB4859)	200	SSR	5	MLM	*SbKAFGK1*	Ng'uni et al. (2012)
	196 RILs	841 038	SNP	4, 5, 8, 9	FarmCPU	*Sobic.005G089600*	Kimani et al. (2020)
	300 accessions	1 291	SNP	4, 3	MLM	*SSIIb, pSB1120*	Sukumaran et al. (2012)
	364 accessions	260 000	SNP	6, 9	MLM	10 candidate genes	Bing et al. (2014)
	390 RILs		SNP	1	MLM	*Sobic.001G029500*	Boyles et al. (2017
Waxy	20 RILs		SNP	10	MLM	*Wx*	McIntyre et al. (2008)
Amylose	390 RILs		SNP	10	MLM	*Wx (Sobic.010G022600)*	Boyles et al. (2017
Fibre	245 accessions	85 585	SNP	3, 10	MLM	*Sobic.003G272200, Sobic.010G172100*	Liu et al. (2018)
Fat	390 accessions	404 627	SNP	2, 4	MLM	*Sb02g023790*	Xiong et al. (2019)
	390 accessions	404 627	SNP	2, 4	MLM	*AMY3, Sb02g023790*	Rhodes et al. (2017)
Tannin	196 RILs	841 038	SNP	4	FarmCPU	*Tannin1, TT16*	Kimani et al. (2020)
Proanthocyanidin	373 accessions	404,628	SNP	1, 2, 4	MLM	*TT8, TTG1, Zm1, TT16*	Rhodes et al. (2014)
3-deoxyanthocyanidin	373 accessions	404,628	SNP	3, 4	MLM	*TT18, TT6, TT7*	Rhodes et al. (2014)

Table 2.3 Quantitative trait loci mapped for grain quality traits in sorghum (continued)

Trait	No. or sorghum population	No. of markers	Type of markers	Chr/LG	Method	Candidate genes	References
Minerals							
Fe	40 accessions (MW467, ZMB6986 and TZ3966)	226	SSR	7	MLM	*SbAGB02*	Ng'uni et al. (2012)
	407 accessions	270, 228	SNP	1	MLMM, CMLM	*Sobic.001G213400*	Shakoor et al. (2016)
Zn	407 accessions	270, 228	SNP	7	MLMM, CMLM	*Sobic.007G064900*	Shakoor et al. (2016)
	40 accessions (MW467, ZMB6986 and TZ3966)	226	SSR	7	MLM	*SbAGB02*	Ng'uni et al. (2012)
Ca	300 accessions	1 292	SNP	3, 6	MLM	*pSB0289, PRC1044*	Sukumaran et al. (2012)
P	407 accessions	270, 228	SNP	1	MLMM, CMLM	*Sobic.001G443900*	Shakoor et al. (2016)
	300 accessions	1 293	SNP	6	MLM	*pSB0140*	Sukumaran et al. (2012)
Physical kernel characteristics							
Kernel hardness	300 accessions	1 290	SNP	3, 10	MLM	*pSB1700, SSIIa*	Sukumaran et al. (2012)
Kernel/pericarp colour							
Red	373 accessions	404 628	SNP	3, 4	GLM, MLM	*ANS, F3H and TT19*	Rhodes et al. (2017)
Brown	373 accessions	404 628	SNP	2, 6	GLM, MLM	*TT12, TT16*	Rhodes et al. (2017)
	1386 accessions	459 304	SNP	4	MLM	*Tannin1 (Sobic.004G280800)*	Hu et al. (2019)
Yellow	1386 accessions	459 304	SNP	1	MLM	*Y1 (Sobic.001G397900)*	Hu et al. (2019)
Pericarp colour	1425 accessions	72, 190	SNP	1, 3, 4	CMLM	Eight candidate genes	Girma et al. (2019)

Chr = Chromosome; LG = Linkage group; RIL = Recombinant inbred line; SNP = Single nucleotide polymorphism; SSR = Simple sequence repeat; MLM = Mixed linear model; FarmCPU = Fixed and random model circulating probability unification; GLM = General linear model; CMLM = Compressed mixed linear model; MLMM = Multi-locus mixed model.

1189

2.11 Conclusions

Ethiopia is the centre of origin and genetic diversity of sorghum. However, the level of diversity and its genetics are poorly understood. Yield and kernel properties that determine traditional food recipes are not understood, which has affected improvement. Kernel composition is a key trait that determines end-use and functional properties. The quality of grain sorghum is determined by nutritional quality traits such as protein and starch digestibility, nutrient bioavailability and anti-nutritional factors such as tannins, processing characteristics, cooking quality and consumer acceptability. Sorghum is used for fermented flatbreads and injera, which is the staple food of Ethiopia.

Environments have varying effects on plant responses at the level of molecular function, developmental processes, morphological traits, basic physiological responses and nutritional grain quality of sorghum. Drought stress conditions decrease starch, grain protein digestibility, micronutrient compositions and increased kernel hardness and total protein in sorghum grain. The molecular mechanisms of drought tolerance and nutritional qualities using genetic, genomic, proteomic and metabolomic studies would lead to the identification of molecular markers that can be targeted for improving the drought tolerance and nutritional grain quality of desirable sorghum cultivars using molecular breeding techniques.

2.12 References

Abah, C., Ishiwu, C., Obiegbuna, J. and Oladejo, A. (2020). Sorghum grains: nutritional composition, functional properties and its food applications. *European Journal of Nutrition and Food Safety,* 12: 101-111.

Abdelhalim, T.S., Kamal, N.M. and Hassan, A.B. (2019). Nutritional potential of wild sorghum: Grain quality of Sudanese wild sorghum genotypes (*Sorghum bicolor* L. Moench). *Food Science and Nutrition,* 7: 1529-1539.

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

Acquaah, G. (2012). Principles of Plant Genetics and Breeding, 2nd ed. John Wiley & Sons Ltd.

Adugna, A. and Bekele, E. (2013). Geographical distribution and phenotypic diversity of wild/weedy sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia: implications for germplasm conservation and crop–wild gene flow. *Plant Genetic Resources,* 11: 68-76.

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

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*, 60:1313-1328.

Afify, A. E.M.M., El-Beltagi, H.S., Abd El-Salam, S.M. and Omran, A.A. (2012). Protein solubility, digestibility and fractionation after germination of sorghum varieties. *PLoS One,* 7: e31154.

Ai, X., Wang, L., Xu, D., Rong, J., Ai, S., Liu, S., Li, C. and Ai, Y. (2021). Stability of artificial soil aggregates for cut slope restoration: A case study from the subalpine zone of southwest China. *Soil and Tillage Research,* 209: 104934.

Aldrich, P., Doebley, J., Schertz, K. and Stec, A. (1992). Patterns of allozyme variation in cultivated and wild *Sorghum bicolor*. *Theoretical and Applied Genetics*, 85: 451-460.

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

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: 1108.

Are, A.K., Srivastava, R.K., Mahalingam, G., Gorthy, S., Gaddameedi, A., Kunapareddy, A., Kotla, A. and Jaganathan, J. (2019). Application of plant breeding and genomics for improved sorghum and pearl millet grain nutritional quality. In: Taylor, J.R.N., Duodu, K.G. (eds). Sorghum and Millets: Chemistry, Technology and Nutritional Attributes, 2nd Ed. Elsevier, pp. 51-68.

Assefa, Y., Staggenborg, S.A. and Prasad, V.P. (2010). Grain sorghum water requirement and responses to drought stress: A review. *Crop Management*, 9: 1-11.

Awika, J.M. and Rooney, L.W. (2004). Sorghum phytochemicals and their potential impact on human health. *Phytochemistry,* 65: 1199-1221.

Ayana, A. and Bekele, E. (1999). Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resources and Crop Evolution,* 46*:* 273-284.

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

Batley, J. and Edwards, D. (2007). SNP Applications in Plants. In: Oraguzie, N.C., Rikkerink, E.H.A., Gardiner, S.E., De Silva, H.N. (eds). Association Mapping in Plants. Springer, New York, NY, pp. 95-102. https://doi.org/10.1007/978-0-387-36011-9_6

Bean, S.R., Zhu, L., Smith, B.M., Wilson, J.D., Ioerger, B.P. and Tilley, M. (2019). Starch and protein chemistry and functional properties. In: Taylor, J.R.N., Duodu, K.G. (eds). Sorghum and Millets: Chemistry, Technology and Nutritional Attributes, 2nd Ed., pp. 131-170.

Bekele, W.A., Wieckhorst, S., Friedt, W. and Snowdon, R.J. (2013). High-throughput genomics in sorghum: from whole-genome resequencing to a SNP screening array. *Plant Biotechnology Journal,* 11: 1112-1125.

Belton, P.S. and Taylor, J.R. (2002). Sorghum. In: Belton, P.S., Taylor, J.R. (ed.). Pseudocereals and less common cereals: grain properties and utilization potential, Springer Science and Business Media, pp. 93-122.

Bemiller, J.N. (1997). Starch modification: challenges and prospects. *Starch-Stärke* 49, 127-131.

Bing, Y., Zhou, Y.F., Gao, M.Y., Zhang, Z., Yi, H., Yang, G.D., Wenjuan, X. and Huang, R.D. (2014). Effect of drought stress during flowering stage on starch accumulation and starch synthesis enzymes in sorghum grains. *Journal of Integrative Agriculture,* 13: 2399-2406.

Birhan, T., Dong, H., Abajebel, N., Wakjira, M., Lemke, C., Vadez, V. and Bantte, K. (2022). Exploiting genetic variation from unadapted germplasm - An example from improvement of sorghum in Ethiopia. *Plants, People, Planet*, 4: 523-536.

Boyles, R.E., Pfeiffer, B.K., Cooper, E.A., Rauh, B.L., Zielinski, K.J., Myers, M.T., Brenton, Z., Rooney, W.L. and Kresovich, S. (2017). Genetic dissection of sorghum grain quality traits using diverse and segregating populations. *Theoretical and Applied Genetics,* 130: 697-716.

Brachi, B., Morris, G.P. and Borevitz, J.O. (2011). Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology*, 12: 1-8.

Bravo, L. (1998). Polyphenols: chemistry, dietary sources, metabolism and nutritional significance. *Nutrition Review,* 56: 317-333.

Brown, P.J., Myles, S. and Kresovich, S. (2011). Genetic support for phenotype-based racial classification in sorghum. *Crop Science,* 51: 224-230.

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

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

Chen, Y., Wang, M. and Ouwerkerk, P.B. (2012). Molecular and environmental factors determining grain quality in rice. *Food and Energy Security,* 1: 111-132.

Chiremba, C., Rooney, L.W. and Taylor, J.R. (2011). Relationships between simple grain quality parameters for the estimation of sorghum and maize hardness in commercial hybrid cultivars. *Cereal Chemistry,* 88: 570-575.

Cook, J.P., McMullen, M.D., Holland, J.B., Tian, F., Bradbury, P., Ross-Ibarra, J., Buckler, E.S. and Flint-Garcia, S.A. (2012). Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiology,* 158: 824-834.

Cremer, J.E., Liu, L., Bean, S.R., Ohm, J.B., Tilley, M., Wilson, J.D., Kaufman, R.C., Vu, T.H., Gilding, E.K. and Godwin, I.D. (2014). Impacts of kafirin allelic diversity, starch content and protein digestibility on ethanol conversion efficiency in grain sorghum. *Cereal Chemistry*, 91: 218-227.

Cui, Y., Xu, G., Magill, C., Schertz, K. and Hart, G. (1995). RFLP-based assay of *Sorghum bicolor* (L.) Moench genetic diversity. *Theoretical and Applied Genetics,* 90: 787-796.

Cullis, B.R., Smith, A., Beeck, C. and Cowling, W. (2010). Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. *Genome,* 53: 1002-1016.

Dalling, M. (1985). The physiological basis of nitrogen redistribution during grain filling in cereals. In: Harper, J.E., Schrader, L.E., Howell, R.W. (eds). Exploitation of Physiological and Genetic Variability to Enhance Crop Productivity. Waverly Press, Baltimore, MD, pp. 55-71.

De Mesa-Stonestreet, N.J., Alavi, S. and Bean, S.R. (2010). Sorghum proteins: the concentration, isolation, modification and food applications of kafirins. *Journal of Food Science,* 75: 90-104.

Dong, J., Gruda, N., Lam, S.K., Li, X. and Duan, Z. (2018). Effects of elevated CO_2 on nutritional quality of vegetables: a review. *Frontiers in Plant Science,* 9: 924.

Dost, K. and Tokul, O. (2006). Determination of phytic acid in wheat and wheat products by reverse phase high-performance liquid chromatography. *Analytica Chimica Acta*, 558: 22-27.

Doggett, H. (1991). Sorghum history in relation to Ethiopia. In: Engels, J.M.M., Hawkes, J.G., Worede, M. (eds). Plant Genetic Resources of Ethiopia, Cambridge University Press, Cambridge, pp. 140-159.

Duodu, K.G. and Dowell, F.E. (2019). Sorghum and millets: quality management systems. In: Taylor, J.R.N., Duodu, K.G. (eds). Sorghum and Millets: Chemistry, Technology and Nutritional Attributes, 2nd Ed. Elsevier, pp. 421-442.

Duodu, K., Taylor, J., Belton, P. and Hamaker, B. (2003). Factors affecting sorghum protein digestibility. *Journal of Cereal Science,* 38: 117-131.

Duressa, D., Weerasoriya, D., Bean, S.R., Tilley, M. and Tesso, T. (2018). Genetic basis of protein digestibility in grain sorghum. *Crop Science,* 58: 2183-2199.

Dykes, L., Rooney, W.L. and Rooney, L.W. (2013). Evaluation of phenolics and antioxidant activity of black sorghum hybrids. *Journal of Cereal Science*, 58: 278-283.

Dykes, L., Seitz, L.M., Rooney, W.L. and Rooney, L.W. (2009). Flavonoid composition of red sorghum genotypes. *Food Chemistry,* 116: 313-317.

Earp, C., McDonough, C. and Rooney, L. (2004). Microscopy of pericarp development in the caryopsis of *Sorghum bicolor* (L.) Moench. *Journal of Cereal Science,* 39: 21-27.

Elkonin, L., Italyanskaya, J. and Panin, V. (2018). Genetic modification of sorghum for improved nutritional value: state of the problem and current approaches. *Journal of Investigating Genomics,* 5: 39-48.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. and Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high-diversity species. *PloS One,* 6: e19379.

Enyew, M., Feyissa, T., Carlsson, A.S., Tesfaye, K., Hammenhag, C., Seyoum, A. and Geleta, M. (2022). Genome-wide analyses using multi-locus models revealed marker-trait associations for major agronomic traits in *Sorghum bicolor*. *Frontiers in Plant Science*, 13: 999692.

Ezeogu, L., Duodu, K. and Taylor, J. (2005). Effects of endosperm texture and cooking conditions on the in vitro starch digestibility of sorghum and maize flours. *Journal of Cereal Science,* 42: 33-44.

Fano, D. (2017). Breeding of sorghum for high lysine in the seed. *International Journal of Agriculture Sciences,* 9: 4702-4707.

FAO (2018). Food and Agriculture Organization of the United Nations. Rome. http://faostat.fao.org (accessed 01/12/2024).

Febles, C., Arias, A., Hardisson, A., Rodrıguez-Alvarez, C. and Sierra, A. (2002). The phytic acid level in wheat flour. *Journal of Cereal Science*, 36: 19-23.

Fox, G., Nugusu, Y., Nida, H., Tedessa, T., McLean, G. and Jordan, D. (2020). Evaluation of variation in Ethiopian sorghum injera quality with new imaging techniques. *Cereal Chemistry,* 97: 362-372.

Fu, Y.B. and Yang, M.H. (2017). Genotyping-by-sequencing and its application to oat genomic research. *Methods in Molecular Biology,* 1536:169-187

Funnell-Harris, D.L., Sattler, S.E., O'Neill, P.M., Gries, T., Ge, Z. and Nersesian, N. (2023). Effects of altering three steps of monolignol biosynthesis on sorghum responses to stalk pathogens and water deficit. *Plant Disease*, 107: 3984-3995.

Geleta, N. and Labuschagne, M. (2005). Qualitative traits variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from, eastern highlands of Ethiopia. *Biodiversity and Conservation,* 14: 3055-3064.

Girma, G., Nida, H., Seyoum, A., Mekonen, M., Nega, A., Lule, D., Dessalegn, K., Bekele, A., Gebreyohannes, A. and Adeyanju, A. (2019). A large-scale genome-wide association analysis of Ethiopian sorghum landrace collection reveals loci associated with important traits. *Frontiers in Plant Science*, 10: 691.

Girma, G., Nida, H., Tirfessa, A., Lule, D., Bejiga, T., Seyoum, A. and Mengiste, T. (2020). A comprehensive phenotypic and genomic characterization of Ethiopian sorghum germplasm defines core collection and reveals rich genetic potential in adaptive traits. *The Plant Genome*, 13: e20055.

Griebel, S., Westerman, R.P., Adeyanju, A., Addo-Quaye, C., Craig, B.A., Weil, C.F. and Tuinstra, M.R. (2019). Mutations in sorghum SBEIIb and SSIIa affect alkali spreading value, starch composition, thermal properties and flour viscosity. *Theoretical and Applied Genetics*, 132: 3357-3374.

Guo, Y., Samuels, D.C., Li, J., Clark, T., Li, C.I. and Shyr, Y. (2013). Evaluation of allele frequency estimation using pooled sequencing data simulation. *The Scientific World Journal,* 2013: 895496.

Harlan, J.R. (1969). Ethiopia: a center of diversity. *Economic Botany*, 23: 309-314.

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

Hikeezi, D.M. (2010). The importance of sorghum grain colour and hardness and their causes and measurement. https://digitalcommons.unl.edu/intsormilpresent/18. Accessed 29 February 2024.

Hossain, A., Mahmud, M., Islam, M., Tajkia, J., Sagar, A. and Khaton, M. (2016). Effect of moisture stress on morphological and yield attributes of four sorghum varieties. *Progressive Agriculture*, 27: 265-271.

Hu, Z., Olatoye, M.O., Marla, S. and Morris, G.P. (2019). An integrated genotyping-by-sequencing polymorphism map for over 10,000 sorghum genotypes. *The Plant Genome,* 12:180044.

Hussain, S., Maqsood, M., Ijaz, M., Ul-Allah, S., Sattar, A., Sher, A. and Nawaz, A. (2020). The combined application of potassium and zinc improves water relations, stays green, irrigation water use efficiency and grain quality of maize under drought stress. *Journal of Plant Nutrition,* 43: 2214-2225.

Impa, S., Perumal, R., Bean, S.R., Sunoj, V.J. and Jagadish, S.K. (2019). Water deficit and heat stress-induced alterations in grain physicochemical characteristics and micronutrient composition in field-grown grain sorghum. *Journal of Cereal Science,* 86: 124-131.

Jannink, J.L., Lorenz, A.J. and Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics,* 9: 166-177.

Kaufman, R., Wilson, J., Bean, S., Galant, A., Perumal, R., Tesso, T., Herald, T. and Shi, Y.C. (2018). Influence of genotype× location interaction on grain sorghum grain chemistry and digestibility. *Agronomy Journal,* 110: 1681-1688.

Khalifa, M. and Eltahir, E.A. (2023). Assessment of global sorghum production, tolerance, and climate risk. *Frontiers in Sustainable Food Systems*, 7: 1184373.

Khoddami, A., Messina, V., Vadabalija Venkata, K., Farahnaky, A., Blanchard, C.L. and Roberts, T. H. (2023). Sorghum in foods: Functionality and potential in innovative products. *Critical Reviews in Food Science and Nutrition,* 63: 1170-1186.

Kimani, W., Zhang, L.M., Wu, X.Y., Hao, H.Q. and Jing, H.C. (2020). Genome-wide association study reveals that different pathways contribute to grain quality variation in sorghum (*Sorghum bicolor*). *BMC Genomics,* 21: 1-19.

Koehler, P. and Wieser, H. (2013). Chemistry of Cereal Grains. In: Gobbetti, M., Gänzle, M. (eds). Handbook on Sourdough Biotechnology. Springer, New York, NY, pp. 11-45. https://doi.org/10.1007/978-1-4614-5425-0_2

Krishnappa, G., Savadi, S., Tyagi, B.S., Singh, S.K., Mamrutha, H.M., Kumar, S. and Singh, G.P. (2021). Integrated genomic selection for rapid improvement of crops. *Genomics*, 113: 1070-1086.

Kumar, A.A. (2016). Botany, Taxonomy and Breeding. In: Rakshit, S., Wang, YH. (eds). The Sorghum Genome. Compendium of Plant Genomes. Springer, Cham, pp. 27-45. https://doi.org/10.1007/978-3-319-47789-3_2

Kumar, R., Dhiman, M., Sharma, L., Dadhich, A., Kaushik, P. and Sharma, M.M. (2022). Nanofertilizers: The targeted nutrient supplier and enhanced nutrient uptake by pearl millets (*Pennisetum glaucum*). *Biocatalysis and Agricultural Biotechnology,* 45: 102524.

Lehrer, M.A. (2022). Phenomic and Genetic Controls of the Drought Stress Response in Sorghum. PhD thesis, West Virginia University, USA.

Li, Y., He, N., Hou, J., Xu, L., Liu, C., Zhang, J., Wang, Q., Zhang, X. and Wu, X. (2018). Factors influencing leaf chlorophyll content in natural forests at the biome scale. *Frontiers in Ecology and Evolution,* 6: 64.

Liu, R., Gong, J., Xiao, X., Zhang, Z., Li, J., Liu, A., Lu, Q., Shang, H., Shi, Y. and Ge, Q. (2018). GWAS analysis and QTL identification of fiber quality traits and yield components in upland cotton using enriched high-density SNP markers. *Frontiers in Plant Science,* 9: 1067.

Lopez, H.W., Leenhardt, F., Coudray, C. and Remesy, C. (2002). Minerals and phytic acid interactions: is it a real problem for human nutrition? *International Journal of Food Science and Technology*, 37: 727-739.

Mao, H., Sun, S., Yao, J., Wang, C., Yu, S., Xu, C., Li, X. and Zhang, Q. (2010). Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proceedings of the National Academy of Sciences,* 107: 19579-19584.

Marshall, D.R. and Brown, A.H.D. (1975) Optimum Sampling Strategies in Genetic Conservation. In: Frankel, O.H., Hawkes, J.G. (eds). Crop Genetic Resources for Today and Tomorrow, Cambridge University Press, London, pp. 53-80.

McCartney, C., Somers, D., Lukow, O., Ames, N., Noll, J., Cloutier, S., Humphreys, D. and McCallum, B. (2006). QTL analysis of quality traits in the spring wheat cross RL4452×'AC Domain'. *Plant Breeding,* 125: 565-575.

McIntyre, C., Drenth, J., Gonzalez, N., Henzell, R. and Jordan, D. (2008). Molecular characterization of the waxy locus in sorghum. *Genome,* 51: 524-533.

Menamo, T., Kassahun, B., Borrell, A.K., Jordan, D.R., Tao, Y., Hunt, C. and Mace, E. (2021). Genetic diversity of Ethiopian sorghum reveals signatures of climatic adaptation. *Theoretical and Applied Genetics*, 134: 731-742.

Mengesha, W., Menkir, A., Meseka, S., Bossey, B., Afolabi, A., Burgueno, J. and Crossa, J. (2019). Factor analysis to investigate genotype and genotype × environment interaction effects on pro-vitamin A content and yield in maize synthetics. *Euphytica*, 215: 1-15.

Mengistu, G., Shimelis, H., Laing, M., Lule, D. and Mathew, I. (2020). Genetic variability among Ethiopian sorghum landrace accessions for major agro-morphological traits and anthracnose resistance. *Euphytica*, 216: 1-15.

Metzker, M.L. (2010). Sequencing technologies - the next generation. *Nature Reviews Genetics* 11: 31-46.

Mezgebe, A.G. (2018). Sorghum waxy and high protein digestibility traits and their relationship with malting and dough-based product-making quality. PhD thesis, University of Pretoria, South Africa.

Mezgebe, A.G., Abegaz, K. and Taylor, J.R. (2018). Relationship between waxy (high amylopectin) and high protein digestibility traits in sorghum and malting quality. *Journal of Cereal Science,* 79: 319-327.

Mishra, A., Singh, A., Sharma, M., Kumar, P. and Roy, J. (2016). Development of EMS-induced mutation population for amylose and resistant starch variation in bread wheat (*Triticum aestivum*) and identification of candidate genes responsible for amylose variation. *BMC Plant Biology*, 16: 1-15.

Mohan Jain, S. and Suprasanna, P. (2011). Induced mutations for enhancing nutrition and food production. *Gene Conserve*, 10: 201-215.

Morison, J. and Lawlor, D. (1999). Interactions between increasing CO_2 concentration and temperature on plant growth. *Plant, Cell and Environment*, 22: 659-682.

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

Mural, R.V., Sun, G., Grzybowski, M., Tross, M.C., Jin, H., Smith, C. and Schnable, J.C. (2022). Association mapping across a multitude of traits collected in diverse environments in maize. *GigaScience*, 11: giac080.

Muraya, M.M., Geiger, H.H., Mutegi, E., Kanyenji, B.M., Sagnard, F., de Villiers, S.M. and Parzies, H.K. (2010). Geographical patterns of phenotypic diversity and structure of Kenyan wild sorghum populations (*Sorghum* spp.) as an aid to germplasm collection and conservation strategy. *Plant Genetic Resources*, 8: 217-224.

Muraya, M.M., Geiger, H.H., De Villiers, S., Sagnard, F., Kanyenji, B.M., Kiambi, D. and Parzies, H.K. (2011). Investigation of pollen competition between wild and cultivated sorghums (*Sorghum bicolor* (L.) Moench) using simple sequence repeats markers. *Euphytica,* 178: 393-401.

Mutegi, E., Sagnard, F., Muraya, M., Kanyenji, B., Rono, B., Mwongera, C., Marangu, C., Kamau, J., Parzies, H., de Villiers, S. and Semagn, K. (2010). Ecogeographical distribution of wild, weedy and cultivated *Sorghum bicolor* (L.) Moench in Kenya: implications for conservation and crop-to-wild gene flow. *Genetic Resources and Crop Evolution*, 57: 243-253.

Nakandalage, N. and Seneweera, S. (2018). Micronutrients use the efficiency of crop-plants under changing climate. In: Hossain, M.A., Kamiya, T., Burritt, D.J., Tran, L.P., Fujiwara, T. (eds). Plant micronutrient use efficiency. Elsevier, pp. 209-224.

Negro, S.S., Millet, E.J., Madur, D., Bauland, C., Combes, V., Welcker, C. and Nicolas, S.D. (2019). Genotyping-by-sequencing and SNP-arrays are complementary for detecting quantitative trait loci by tagging different haplotypes in association studies. *BMC Plant Biology*, 19, 1-22.

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

Ohanenye, I.C., Emenike, C.U., Mensi, A., Medina-Godoy, S., Jin, J., Ahmed, T. and Udenigwe, C.C. (2021). Food fortification technologies: Influence on iron, zinc and vitamin A bioavailability and potential implications on micronutrient deficiency in sub-Saharan Africa. *Scientific African*, 11: e00667.

Okuda, T. and Ito, H. (2011). Tannins of constant structure in medicinal and food plants are hydrolyzable tannins and polyphenols related to tannins. *Molecules,* 16: 2191-2217.

Opazo-Navarrete, M., Tagle Freire, D., Boom, R.M. and Janssen, A.E. (2019). The Influence of starch and fiber on in vitro protein digestibility of dry fractionated quinoa seed (Riobamba Variety). *Food Biophysics*, 14: 49-59.

Paterson, A.H. (2013). The Sorghum Genome Sequence: A Core Resource for Saccharinae Genomics. In: Paterson, A. (eds) Genomics of the Saccharinae. Plant Genetics and Genomics: Crops and Models, vol 11. Springer, New York, NY, pp. 105-119. https://doi.org/10.1007/978-1-4419-5947-8_5.

Phuke, R.M., Anuradha, K., Radhika, K., Jabeen, F., Anuradha, G., Ramesh, T., Hariprasanna, K., Mehtre, S.P., Deshpande, S.P. and Anil, G. (2017). Genetic variability, genotype× environment interaction, correlation and GGE biplot analysis for grain iron and zinc concentration and other agronomic traits in RIL population of sorghum (*Sorghum bicolor* L. Moench). *Frontiers in Plant Science,* 8: 712.

Piguet, F. (2003). Ethiopia: Hararghe Food Security is hampered by long-term drought conditions and economic constraints. UNDP-EUE. Reliefweb, 29 June - 5 July, 2003.

Poland, J.A. and Rife, T.W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome*, 5: 92-102.

Pushpamma, P. (1993). Sorghum as food in the semi-arid tropics: studies in the dryland communities of Andhra Pradesh, India. IDRC, India, 59-63

Queiroz, V.A.V., da Silva, C.S., de Menezes, C.B., Schaffert, R.E., Guimarães, F.F.M., Guimarães, L.J.M., de Oliveira Guimarães, P.E. and Tardin, F.D. (2015). Nutritional composition of sorghum [*Sorghum bicolor* (L.) Moench] genotypes cultivated without and with water stress. *Journal of Cereal Science,* 65: 103-111.

Raymundo, R., Mclean, G., Sexton-Bowser, S. and Morris, G.P. (2023). Crop modeling suggests limited transpiration would increase yield of sorghum across drought-prone regions of the United States. Frontiers in Plant Science, 14:1283339

Raymundo, R., Wang, X., Felderhoff, T., Sexton-Bowser, S., Poland, J., Lipka, A.E. and Morris, G.P. (2023). Field phenomics reveals genetic variation for transpiration response to vapor pressure deficit in sorghum. *bioRxiv*, 2023-06.

Reddy, N.R. and Pierson, M.D. (1994). Reduction in antinutritional and toxic components in plant foods by fermentation. *Food Research International,* 27: 281-290.

Rhodes, D. (2014). Diversity, genetics and health benefits of sorghum grain. PhD thesis, University of South Carolina, USA.

Rhodes, D.H., Hoffmann, L., Rooney, W.L., Ramu, P., Morris, G.P. and Kresovich, S. (2014). Genome-wide association study of grain polyphenol concentrations in global sorghum [*Sorghum bicolor* (L.) Moench] Germplasm. *Journal of Agricultural and Food Chemistry*, 62: 10916-10927.

Rhodes, D. H., Hoffmann, L., Rooney, W.L., Herald, T.J., Bean, S., Boyles, R., Brenton, Z.W. and Kresovich, S. (2017). Genetic architecture of kernel composition in global sorghum germplasm. *BMC Genomics,* 18: 1-8.

Rooney, L.W. and Murty, D.S. (1982) Evaluation of Sorghum Food Quality. In: House, L.R., Mughogho, L.K., Peacock, J.M. (eds). Sorghum in the eighties. Proceedings of the International Symposium on Sorghum, Patancheru, India, 2-7 November 1981. ICRISAT: Andhra Pradesh, India, pp. 571-588.

Safian, N., Naderi, M.R., Torabi, M., Soleymani, A. and Salemi, H.R. (2022). Corn (*Zea mays* L.) and sorghum (*Sorghum bicolor* (L.) Moench) yield and nutritional quality affected by drought stress. *Biocatalysis and Agricultural Biotechnology*, 45: 102486.

Salaam-Blyther, T. (2011). USAID Global Health Programs: FY2001-FY2012 Request. Washington DC: *Congressional Research Service*, 1-16.

Sarshad, A., Talei, D., Torabi, M., Rafiei, F. and Nejatkhah, P. (2021). Morphological and biochemical responses of *Sorghum bicolor* (L.) Moench under drought stress. *SN Applied Sciences*, 3: 1-12.

Sehgal, A., Sita, K., Siddique, K.H., Kumar, R., Bhogireddy, S., Varshney, R.K., HanumanthaRao, B., Nair, R.M., Prasad, P.V. and Nayyar, H. (2018). Drought and/or heat-stress effects on seed filling in food crops: impacts on functional biochemistry, seed yields and nutritional quality. *Frontiers in Plant Science,* 9: 1705.

Sen, D., Fernández, A., Crozier, D., Henrich, B., Sokolov, A.V., Scully, M.O., Rooney, W.L. and Verhoef, A.J. (2023). Non-Destructive direct pericarp thickness measurement of sorghum kernels using extended-focus optical coherence microscopy. *Sensors*, 23: 707.

Serna-Saldivar, S.O. and Espinosa-Ramírez, J. (2019). Grain structure and grain chemical composition. In: Taylor, J.R.N., Duodu, K.G. (eds). Sorghum and Millets: Chemistry, Technology and Nutritional Attributes, 2nd Ed. Elsevier, pp. 85-129.

Sirany, T., Tadele, E., Aregahegn, H. and Wale, D., 2022. Economic Potentials and Use Dynamics of Sorghum Food System in Ethiopia: Its Implications to Resolve Food Deficit. *Advances in Agriculture*, 2022: 1-11.

Smith, A. B., Ganesalingam, A., Kuchel, H. and Cullis, B.R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theoretical and Applied Genetics,* 128: 55-72.

Somegowda, V.K., Vemula, A., Naravula, J., Prasad, G., Rayaprolu, L., Rathore, A., Blümmel, M. and Deshpande, S.P. (2021). Evaluation of fodder yield and fodder quality in sorghum and its interaction with grain yield under different water availability regimes. *Current Plant Biology,* 25: 100191.

Stangoulis, J.C. and Knez, M. (2022). Biofortification of major crop plants with iron and zinc-achievements and future directions. *Plant Soil,* 474: 57-76.

Sudheesh, C., Bhat, Z.R., Aaliya, B. and Sunooj, K.V. (2022). Cereal proteins. In: Kour, J., Nayik, G.A. (eds). Nutraceuticals and Health Care. Academic Press, pp. 29-60.

Sukumaran, S., Xiang, W., Bean, S.R., Pedersen, J.F., Kresovich, S., Tuinstra, M.R., Tesso, T.T., Hamblin, M.T. and Yu, J. (2012). Association mapping for grain quality in a diverse sorghum collection. *The Plant Genome,* 5: 126-135.

Syed, R. (2019). A study of health-promoting components in pigeon pea and its application in food models, PhD thesis, Tennessee State University, USA.

Tao, Y., Zhao, X., Wang, X., Hathorn, A., Hunt, C., Cruickshank, A.W., van Oosterom, E.J., Godwin, I.D., Mace, E.S. and Jordan, D.R. (2020). Large-scale GWAS in sorghum reveals common genetic control of grain size among cereals. *Plant Biotechnology Journal,* 18: 1093-1105.

Tapiwa, K.A. (2019). Polyphenols in sorghum, their effects on broilers and methods of reducing their effects: A Review. *Biomedical Journal of Scientific and Technical Research ,* 19: 14058-14061.

Taylor, J. and Schüssler, L. (1986). The protein compositions of the different anatomical parts of sorghum grain. *Journal of Cereal Science,* 4: 361-369.

Taylor, J.R. and Duodu, K.G. (2019). Traditional sorghum and millet food and beverage products and their technologies. In: Taylor, J.R.N., Duodu, K.G. (eds). Sorghum and Millets: Chemistry, Technology and Nutritional Attributes, 2nd Ed. Elsevier, pp. 259-292.

Terblanche, E. (2017). Development of novel methods for tannin quantification in grapes and wine. MSc dissertation, Stellenbosch University, South Africa.

Tesso, T., Ejeta, G., Chandrashekar, A., Huang, C.P., Tandjung, A., Lewamy, M., Axtell, J.D. and Hamaker, B.R. (2006). A novel modified endosperm texture in a mutant high-protein digestibility/high-lysine grain sorghum (*Sorghum bicolor* (L.) Moench). *Cereal Chemistry,* 83: 194-201.

Teshome, M. and Feyissa, T. (2013). Genetic diversity of wild sorghum (*Sorghum bicolor* ssp. *verticilliflorum* (L.) Moench) germplasm from Ethiopia as revealed by ISSR markers. *Journal of Plant Sciences*, 12:137-144.

Tetlow, I.J. (2011). Starch biosynthesis in developing seeds. *Seed Science Research*, 21: 5-32.

Ullah, A., Zhao, C., Zhang, M., Sun, C., Liu, X., Hu, J., Zeeshan, M., Zaid, A., Dai, T. and Tian, Z. (2023). Nitrogen enhances the effect of pre-drought priming against post-anthesis drought stress by regulating starch and protein formation in wheat. *Physiologia Plantarum*, 175: e13907.

Wang, N., Yuan, Y., Wang, H., Yu, D., Liu, Y., Zhang, A. and Zhang, X. (2020). Applications of genotyping-by-sequencing (GBS) in maize genetics and breeding. *Scientific Reports*, 10: 16308.

Wang, S. and Copeland, L. (2013). Molecular disassembly of starch granules during gelatinization and its effect on starch digestibility: a review. *Food and Function*, 4: 1564-1580.

Watrelot, A.A. and Norton, E.L. (2020). Chemistry and reactivity of tannins in Vitis spp.: A review. *Molecules,* 25: 2110.

Weerasooriya, D.K., Bean, S.R., Nugusu, Y., Ioerger, B.P. and Tesso, T.T. (2018). The effect of genotype and traditional food processing methods on in-vitro protein digestibility and micronutrient profile of sorghum cooked products. *PLoS One,* 13: e0203005.

Wong, J.H., Marx, D.B., Wilson, J.D., Buchanan, B.B., Lemaux, P.G. and Pedersen, J.F. (2010). Principal component analysis and biochemical characterization of protein and starch reveal primary targets for improving sorghum grain. *Plant Sci*ence, 179: 598-611.

Wu, Y., Li, X., Xiang, W., Zhu, C., Lin, Z., Wu, Y., Li, J., Pandravada, S., Ridder, D.D. and Bai, G. (2012). The presence of tannins in sorghum grains is conditioned by different natural alleles of Tannin1. *Proceedings of the National Academy of Science,* 109: 10281-10286.

Wu, Y., Yuan, L., Guo, X., Holding, D.R. and Messing, J. (2013). Mutation in the seed storage protein kafirin creates a high-value food trait in sorghum. *Nature Communications*, 4: 2217.

Xiong, Y., Zhang, P., Warner, R.D. and Fang, Z. (2019). Sorghum grain: From genotype, nutrition and phenolic profile to its health benefits and food applications. *Comprehensive Reviews in Food Science and Food Safety,* 18: 2025-2046.

Xu, J.H. and Messing, J. (2008). Diverged copies of the seed regulatory *Opaque-2* gene by a segmental duplication in the progenitor genome of rice, sorghum and maize. *Molecular Plant,* 1: 760-769.

Yahaya, M.A. and Shimelis, H. (2022). Drought stress in sorghum: Mitigation strategies, breeding methods and technologies - A review. *Journal of Agronomy and Crop Science,* 208: 127-142.

Yang, G., Chen, S., Chen, L., Sun, K., Huang, C., Zhou, D., Huang, Y., Wang, J., Liu, Y. and Wang, H. (2019). Development of a core SNP array based on the KASP method for molecular breeding of rice. *Rice,* 12: 1-18.

Yetneberk, S., de Kock, H.L., Rooney, L.W. and Taylor, J.R. (2004). Effects of sorghum cultivar on injera quality. *Cereal Chemistry,* 81: 314-321.

Yetneberk, S., Rooney, L.W. and Taylor, J.R. (2005). Improving the quality of sorghum injera by decortication and compositing with tef. *Journal of the Science of Food and Agriculture*, 85:1252-1258.

Zhu, F. (2014). Structure, physicochemical properties and uses of millet starch. *Food Research International,* 64: 200-211.

CHAPTER 3

PHENOTYPIC DIVERSITY AND DROUGHT ADAPTATION IN AN ETHIOPIAN SORGHUM LANDRACE COLLECTION

3.1 Abstract

Climate-resilient and high-yielding sorghum cultivars must be developed for sustainable food systems and improved livelihoods in sub-Saharan Africa. This study aimed to assess the genetic diversity and drought tolerance of 365 Ethiopian sorghum landrace accessions using agronomic traits, to select promising lines for production and/or breeding. In stress environments, yield and yield components had low means, genetic variance and heritability values, except for days to flowering. Irrigated conditions lead to a 44.6% higher mean yield and 14% higher thousand seed weight. Ten genotypes were identified, that had the highest stress tolerance index, geometric mean productivity, mean productivity and yield, in all trials. Cluster analysis separated the sorghum landraces into seven distinct genetic groups, largely based on geographic origin. Grain yield correlated significantly with head weight, thousand seed weight and head width. The current study showed significant genetic variation among Ethiopian sorghum landraces under irrigated and drought-stress conditions for strategic conservation and breeding in similar agro-ecologies.

3.2 Introduction

Sorghum was shown to have significant genetic and geographic diversity for both phenotypic and grain quality parameters. Genetic variation in plant genetic resources enables plant breeders to develop new improved varieties with desirable traits, including farmer-preferred traits (Govindaraj et al., 2015). Genetic variation is required not only for the identification of superior genotypes that could be recommended for direct release for commercial production, but also for the identification of desirable parents to use as source material in genetic improvement. High levels of genetic diversity are essential for climate adaptation development of crop varieties to mitigate the effects of climate change. Ethiopia is one of the centres of origin of sorghum and a major source of sorghum diversity (Harlan, 1969; Harlan and Wet, 1972; Mekbib, 2006). The genetic diversity of sorghum has been studied using morphological, biochemical and molecular markers (Appa Rao et al., 1996; Geleta et al., 2005; Girma et al., 2019).

Sorghum is generally well adapted to abiotic stress conditions, especially drought stress, which is the primary abiotic stress in sorghum production. The impact of drought stress on grain yield and its components will vary depending on the stage of plant development at which the water deficit occurs. It was reported that drought stress slowed panicle development before anthesis but accelerated it after anthesis (Manjarrez-Sandoval, 1989). Grain sorghum's adaptation to various environments has resulted in the evolution and occurrence of comprehensive genetic variation for drought tolerance (Amelework et al., 2015). Drought and heat stress have a greater impact on crop growth and yield if they occur during the reproductive stages of crop development (Sehgal et al., 2018). Tolerant genotypes can be used to increase yield and crop performance among small-holder farmers, reducing food insecurity, poverty and famine. This study aimed to assess the level and pattern of genetic variation and drought adaptation in a core collection of Ethiopian sorghum landraces, using phenotypic traits.

3.3 Materials and methods

3.3.1 Plant material and design

In this study, 361 Ethiopian sorghum landraces and four improved varieties were evaluated for genetic diversity during two consecutive seasons (2020 and 2021). Landraces included the recently defined representative core subset (n=387) selected from an initial set of over 2 000 accessions (Girma et al., 2019). A few accessions from the core were not included in the current study due to lack of seed. A germination test was done on seeds in 2020 before planting in 2021. Seven sorghum landraces that did not germinate were replaced by seven different landraces in 2021, so this was the only difference between trials of the two seasons.

In these extensive field experiments, covering diverse environments, a combination of row-column design, spatial and MET analysis and linear mixed models (LMM) was used. Employing LMMs is crucial for managing correlated and uncorrelated observations across different environments, leading to more precise estimations of treatment effects and their variability. LMMs effectively capture variability among different experimental units like plots, blocks, and locations, accounting for unmeasured sources of variation that may impact on traits across diverse environments, by

integrating covariates related to environmental conditions (such as temperature, precipitation and soil type) as fixed effects (Smith et al., 2001; Smith and Cullis, 2018).

Due to the large number of landrace accessions included in this study, partial replication was used in all trial designs (Cullis et al., 2006). This customised design was used to minimise spatial error effects within each trial and had a row-column arrangement. A multi-environment trial design with 354 landraces and four commercial checks, totalling 358 sorghum entries across two sites (Melkassa and Miesso) was used. Of the total of 354 sorghum landrace accessions, a third were replicated at each site and the replicated set was different at each site. The four checks were replicated at all sites. Each plot consisted of a single row, 4 m in length. Experiments consisted of a total of 10 columns and 48 rows. A total of 48 sorghum landraces were organised per single column.

Soon after ploughing, seeds were drilled in the row at 75 cm between rows and thinned to maintain 15 cm intra-row spacing. Fields were ploughed in the main summer cropping season. Basal fertiliser was applied at the recommended rate and normal recommended agronomic practices for sorghum production in Ethiopia were applied.

3.3.2 Locations

Trials were planted at two dry lowland sorghum production environments, which are both found in the central Rift Valley of Ethiopia. Miesso and Melkassa are located 225 and 100 km east of the capital, Addis Ababa, respectively. Miesso, where sorghum is traditionally grown as a rainfed crop, is a drought stress sorghum testing site in Ethiopia (MoA, 1998) while Melkassa serves as the national sorghum research coordinating station and has facilities for supplemental irrigation to create non-stressed growing conditions for comparison with stress conditions. Information for the two sites including planting and harvesting dates, rainfall patterns, relative humidity and maximum/minimum temperatures are presented in Table 3.1. Data on monthly temperatures and rainfall are listed in Appendix 1 for the two consecutive seasons (2020 and 2021).

The sorghum landrace collections display differences in flowering and maturity periods, as well as in morphological characteristics. This research highlights that the landraces have an obvious

extended flowering duration of two months during particular seasons and locations, as indicated by the range of flowering periods detailed in Table 3.7.

In Miesso, the sorghum landraces were planted on July 20, 2020, which corresponded to a severe drought period. Throughout the flowering stage, only 20.5 mm of rainfall was recorded, with 6 mm occurring in October and 14.5 mm in November. In the subsequent year, the sorghum landraces were planted on May 11, 2021 and received a considerably higher total rainfall of 166.6 mm throughout the flowering stage, with 95.6 mm in August and 69.9 mm in September, exceeding the precipitation of 2020 (Table 3.1 and Appendix 1). In MET analysis, both LMM and integrating random effects are critical. LMMs allow for the incorporation of both fixed effects (treatment effects or covariates) and random effects (environmental effects or experimental errors) into the model, providing a flexible framework for analysing MET data (Smith et al., 2001).

Table 3.1 Long-term average temperature and rainfall properties and planting and harvesting dates for each experimental site

Location	Longitude	Latitude	Altitude (m.a.s.l.)	Soil type	Rainfall (mm)	Minimum T˚C	Maximum T˚C
Melkassa	39˚9'E	14˚6'N	1 550	Clay loam	763	20.4	34
Miesso	39˚21'E	8˚30'N	1 470	Clay loam	570	16	31

	Planting date		Harvesting date	
Location	**Year 1**	**Year 2**	**Year 1**	**Year 2**
Melkassa	Jul 15, 2020	Aug 3, 2021	Jan 1-20, 2021	Jan 10-12, 2022
Miesso	Jul 20, 2020	May 11, 2021	Nov 27-28, 2020	Aug 30-Dec 24, 2021

M.a.s.l. = meter about sea level; T˚C = temperature in degrees centigrade.

Both locations have a brief rainy season, with most rain falling in July and August. During these months, there is a good quantity of rain, but a part of it is squandered as runoff. After September, rain becomes considerably less. As a result, sorghum accessions produced under rainfed conditions at both sites often face moisture stress during the anthesis and post-anthesis development stages. From planting to grain filling, supplemental irrigation was provided every week at Melkassa. Seven supplemental irrigations (three before and four after anthesis) were given, which was sufficient for normal crop growth and development in this non-stress environment.

3.3.3 Phenotypic data collection and sampling techniques

Data were collected from the single rows for plot-based data and five randomly sampled plants for plant-based data, following the descriptors for sorghum (IBPGR/ICRISAT, 1993). The data were reported as an average of five plants per entry. Details of the data collection are listed in Table 3.2.

Table 3.2 List of quantitative characteristics recorded at Melkassa and Miesso

Quantitative character	Code	Description
Days to 50% flowering (days)	DTF	Number of days from emergence to when 50% of plants have started flowering in a plot
Days to 90% maturity (days)	DTM	Number of days from emergence to when 90% of plants have started maturity in a plot
Head length (cm)	HLP	Length of the panicle from its base to tip
Head width (cm)	HWD	Width of panicle in a natural position at the widest part
Head weight (g)	HWT	The average weight of the harvested whole panicles
Internode length per plant (cm)	ILP	Average internode length of the internode from its base to tip
Number of internodes per plant (count)	NIP	Count of total internodes per plant
Plant height (cm)	PHT	Height of the main stalk from the ground to the tip of the main panicle
Grain yield per plant (g)	GY	Weight of grain per panicle (average of harvested plants or panicles in the plot)
Thousand seed weight (g)	TSW	The weight of a thousand seeds at 12% moisture content

3.3.4 Data analysis

The concurrence of genotypes across the two seasons allowed trials to be analysed as four METs for each of the measured quantitative parameters. Each MET was analysed by fitting a linear mixed model for random effects, which can obtain the best linear unbiased predictor (BLUP) of the variety effect, that is, the best estimate of the performance of that variety in future trials using the package ASReml (Butler et al., 2017) and R statistical software (Matloff, 2011; Brown, 2021). The model consisted of the targeted trait at each trial, random effects (BLUPs) for genotypes within the trial and spatial error for each trial (Smith and Cullis, 2018). Spatial effects were fitted to each

trial and then a variance structure was created to produce a correlation of environments in a factor analytic framework (Smith et al., 2001). Heritability was calculated for the different environments according to the method proposed by Cullis et al. (2006).

The model used for analysis is the spatial mixed model for MET and can be written as:

$$\begin{aligned} y &= X_\tau + Z_\mu + e \\ &= X_\tau + Z_0\mu_0 + Z_g\,\mu_g + e \end{aligned}$$

Where the fixed effect τ includes environmental main effects and trial-specific effects for extraneous field variation (Butler et al., 2017), y is the matrix of the response variable and X is the design matrix of the fixed variable. Z_μ is the design matrix of random variables, μ is the vector of random effect parameters and e is the error, u_g is the variety effect at each environment with the associated design matrix $Z_g^{(nxmp)}$ and u_0 comprises an additional random effect with design matrix Z_0.

3.3.5 Cluster analysis

The use of multiple measurements subjected to multivariate statistical analysis can provide a measure based on generalised distance as indicated by D^2 statistics (Mahalanobis, 2018; Yang and Delpha, 2022). Mahalanobis D^2 statistics is a powerful tool for describing divergence among lines based on multiple characters (Deepashree et al., 2023). This statistical tool is especially valuable in fields such as pattern recognition and cluster analysis and classification, where the goal is to determine the degree of dissimilarity or similarity among different samples based on multiple variables (Madhusri et al., 2022).

Before clustering, standardisation of the scoring system across different traits enables data analysis and interpretation by converting trait scores to a 1 to 10 scale for clustering purposes. This improves the clustering process by ensuring that all traits are measured on a consistent scale, making it easier to identify patterns and relationships in the data (Manly, 1986; Deutsch et al., 2015). The optimum number of clusters in the data set is determined by using gap statistics (Tibshirani et al., 2001). R statistical software (Matloff, 2011; Brown, 2021) was used for the UPGMA clustering analysis method.

3.3.6 Drought stress tolerance indices

The following relationships were used to calculate drought stress tolerance indices:

$$\text{Stress susceptibility index (SSI)} = \frac{1-(Ys/Yp)}{1-(\bar{Y}s\ /\bar{Y}p\ )}$$

$$\text{Mean productivity (MP)} = \frac{Yp+Ys}{2}$$

$$\text{Tolerance index (TOL)} = Yp\text{-}Ys$$

$$\text{Stress tolerance index (STI)} = \frac{Yp+Ys}{\bar{Y}p^2}$$

$$\text{Geometric mean productivity (GMP)} = (YpxYs)^{0.5}$$

$$\text{Yield stability index (YSI)} = \frac{Ys}{Yp}$$

Where Yp is yield in non-stress conditions; Ys is yield in stress conditions; $\bar{Y}s$ and $\bar{Y}p$ represent the mean yield of all cultivars under stress and non-stress conditions, respectively and 1- ($\bar{Y}s$ /$\bar{Y}p$) represents the stress intensity (Fischer and Maurer, 1978; Husain et al., 1990; Fernandez, 1992; Gavuzzi et al., 1997).

3.4 Results

3.4.1 Genotypic variation, genotype by environment interaction and assessment of heritability

Grain yield (GY) and yield-related components, except head width per plant (HWD) and thousand seed weight (TSW) for Miesso in 2020 (MS20), showed highly significant genotypic variances (σ^2g) in each environment. Low mean, range, σ^2g and heritability (H^2) values were observed for GY and yield-related components in stress environments MS20 and Miesso 2021 (MS21), except for days to 50% flowering (DTF) (Table 3.3). The study also examined σ^2g for different traits under non-stress conditions in Melkassa 2020 (MK20) and Melkassa 2021 (MK21). MK20 had the highest σ^2g for all traits, while MS20 had the lowest values in stress environments. Highly significant genotype (σ^2g), environment (σ^2en) and genotype x environment (σ^2gen) interaction variances were present for all traits, except for σ^2gen for days to 90% maturity (DTM) (Table 3.4). H^2 was consistently higher in non-stress environments than in stress environments (Table 3.3). DTF and head length per plant (HLP) had the highest H^2 across all environments while GY and TSW at MS20 had the lowest H^2.

Table 3.3 Heritability and variance components for each trait in each environment

Trait	MK20			MK21			MS20			MS21		
	σ^2g	σ^2e	H^2	σ^2g	σ^2e	H^2	σ^2g	σ^2e	H^2	σ^2g	σ^2e	H^2
DTF	163.30**	40.89	87.15	125.80**	20.74	89.28	117.40**	9.17	87.24	472.20**	86.19	84.73
DTM	78.10**	55.98	73.38	18.90**	17.12	72.04	3.10*	1.73	63.01	55.70**	26.47	66.76
HLP	44.20**	9.51	88.72	36.10**	10.33	90.26	25.00**	10.37	80.05	34.00**	8.66	86.91
HWD	0.90**	0.99	56.78	0.60**	1.34.00	51.37	0.1	2.25	50.65	0.40*	1.71	45.25
HWT	795.20**	1501.07	41.57	162.90**	699.56	26.53	271.90**	1474.30	62.47	112.00**	379.00	20.93
ILP	4.80**	1.21	79.02	-	-	-	11.30**	4.28	54.99	9.13**	12.04	51.88
INP	19.50**	8.12	70.62	2.90**	3.10	69.55	1.20**	0.47	68.57	0.80**	2.75	60.71
PHT	2091.10**	250.56	78.34	1507.40**	1342.38	69.75	1405.40**	1276.71	50.32	1699.10**	1133.15	69.27
GY	447.40**	593.60	46.30	77.60**	279.00	39.50	140.20**	456.70	19.90	100.70**	222.40	39.70
TSW	44.60**	28.80	80.90	28.50**	14.50	77.30	4.6	17.80	14.40	17.50**	85.60	810.00

MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021; DTF = Days to 50% flowering (days); DTM = Days to 90% maturity (days); HLP = Head length per plant (cm); HWD = Head width per plant (cm); HWT = Head weight per plant (g); ILP = Internode length per plant (cm); INP = Internode number per plant (count); PHT = Plant height per plant (cm); GY = Grain yield per plant (g); TSW = Thousand seed weight (g); σ^2g = genotype variance; σ^2en = environment variance; σ^2gen = genotype x environment interaction variances; H^2 = heritability.

Table 3.4 Genotype (σ^2g), environment (σ^2en) and genotype x environment (σ^2gen) interaction variances and standard errors for agronomic traits for combined environments

Trait	σ^2g	σ^2en	σ^2gen	SE
DTF	121.10**	893.00**	21.12**	10.33
DTM	6.30**	7884.00**	0.00	0.86
HLP	34.43**	27.20**	2.20**	2.90
HWD	0.41*	247.50**	0.19**	0.07
HWT	184.50**	637.60**	266.62**	29.83
ILP	0.81**	1576.90**	5.44**	0.40
INP	0.98**	769.40**	0.28**	0.12
PHT	1295.30**	188.80**	203.25**	114.70
GY	37.94**	734.50**	136.69**	8.75
TSW	3.94*	934.00**	3.57**	0.47

** P ≤ 0.01; * P ≤ 0.05; DTF = Days to 50% flowering (days); DTM = Days to 90% maturity (days); HLP = Head length per plant (cm); HWD = Head width per plant (cm); HWT = Head weigh per plant (cm); ILP = Internode length per plant (cm); INP = Internode number per plant (count); PHT = Plant height per plant (cm); GY = Grain yield per plant (g); TSW = Thousand seed weight (g); SE = standard error.

When environments are correlated (similar landrace responses in different environments), the best genotypes selected in one environment will also be the best in the correlated environment. In this case, MET analysis can also assist in recognising genotype broad and specific adaptation across a wide range of genetic materials. Correlation values of -1 indicated that results of these environments were opposite, implying that the highest-performing genotypes in one environment were the lowest-performing genotypes in the other. A correlation of +1 indicated that the two environments were perfectly similar, so selecting superior genotypes in one environment will also apply to the correlated environment (Figure 3.1 and Appendix 6).

3.4.2 Influence of drought on measured characteristics

Low GY, TSW and head weight (HWT) mean values were observed under drought-stress conditions compared to the non-stress conditions (Table 3.5). The GY, TSW and HWT values for the four environments ranged from 32.66 to 79.59 g/plant, 28.26 to 34.86 g and 54.36 to 121.16 g/plant, respectively, with irrigated environments (Melkassa) typically producing the highest GY, TSW and HWT. The MK20 trial (irrigated) resulted in the highest yield, most likely due to the supplemental and rain throughout the growing season.

Melkassa (irrigated) had a 54.5% higher mean GY than Miesso (stress) in 2020 and Melkassa (irrigated) had a 26.9% higher mean GY than Miesso (stress) in 2021. Overall, the irrigated conditions had a 44.6% higher GY and 14% higher TSW than the stress conditions. Because of

the impact of rainfall patterns and other agro-climatic factors, the average yield in 2020 was 31.5% higher than in 2021. HLP, HWD, plant height (PHT) and DTF were higher under stress than non-stress conditions (Table 3.5).

Table 3.5 Mean values for grain yield and yield components of sorghum landraces in non-stress and stressed environments for two growing seasons

Environment	GY	TSW	HWT	HLP	HWD	PHT	ILP	INP	DTF	DTM
2020										
Melkassa	79.59	34.86	121.16	25.84	6.90	269.16	12.15	19.79	104.49	165.94
Miesso	36.18	30.64	86.07	27.93	7.87	261.82	18.58	10.92	92.94	129.41
Mean	57.89	32.75	103.62	26.89	7.38	265.49	15.36	15.35	98.72	147.68
2021										
Melkassa	44.66	33.92	70.68	26.19	5.94	232.86	-	10.83	99.65	155.30
Miesso	32.66	28.26	54.36	24.96	6.99	302.07	22.81	11.88	122.14	152.69
Mean	39.66	31.09	62.52	25.58	6.46	267.47	22.81	11.35	110.90	154.00
Mean per location										
Melkassa	62.13	34.39	95.92	26.02	6.42	251.01	12.15	15.40	102.07	160.62
Miesso	34.42	29.45	70.22	26.45	7.43	281.95	20.69	11.40	107.54	141.05
LSD (0.05)	1.55	0.31	1.98	0.22	0.07	1.92	0.22	0.25	0.89	0.53

Melkassa = dry lowland (irrigated); Miesso = dry lowland (stress); GY = Grain yield per plant (g); TSW = Thousand seed weight (g); HWT = Head weight (g); HLP = Head length per plant (cm); HWD = Head width per plant (cm); PHT = Plant height per plant (cm); ILP = Internode length per plant (cm); INP = Internode number per plant (count); DTF = Days to 50% flowering (days); DTM = Days to 90% maturity (days).

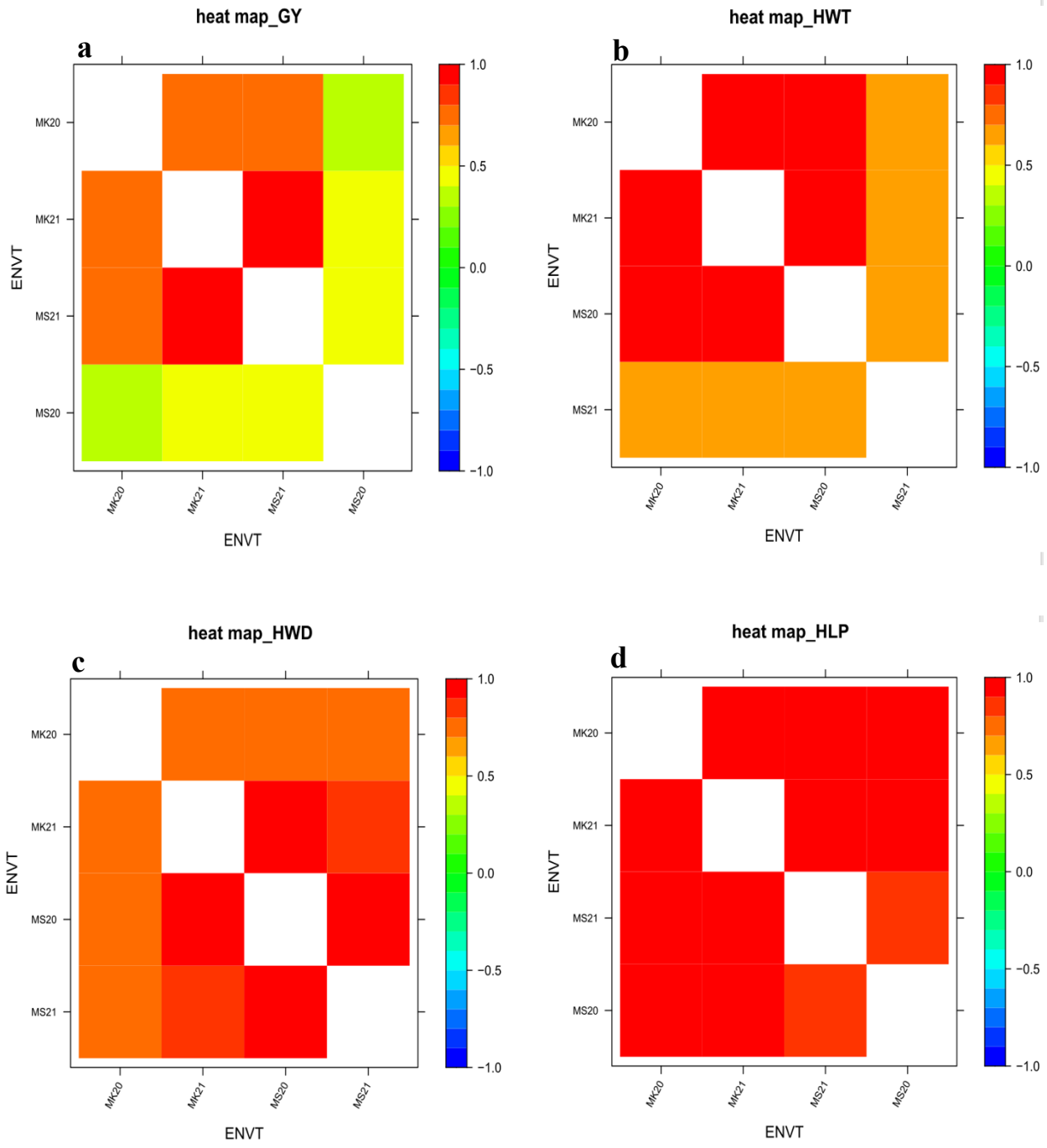


Figure 3.1 Correlations of testing environments for (a) grain yield (GY), (b) head weight (HWT), (c) head width (HWD) and (d) head length (HLP)

ENVT = Environment; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021; 1.0 indicates a good positive correlation, - 1.0 negative correlation and 0.0 no correlation.

3.4.3 Stress tolerance indices

The GY of landraces across environments was used to compute TOL. A high TOL shows stress sensitivity and as a result, tolerant cultivars are selected based on lower TOL values. There were positive and significant correlations between yield in non-stress conditions (Yp) and yield in stress conditions (Ys), GMP, MP, STI and TOL (Table 3.6). TOL revealed G18 as the most tolerant genotype, with a TOL index of 10.41, followed by G296 (11.44), G141 (11.57) and G132 (11.93). According to TOL, the least stress-tolerant sorghum landraces were G351 (47.95), G324 (44.61), G71 (43.81) and G97 (43.58) (Appendix 2).

The YSI had a significant positive correlation with Ys (r = 0.40), which suggests that a higher value of Ys is associated with a higher YSI value. On the other hand, YSI had a significant negative correlation with Yp (r = -0.26) and TOL (r = -0.76), indicating that higher Yp and TOL values were associated with lower YSI values. The SSI showed a significant negative correlation with YSI (r = -1.00) and Ys (r = -0.40). SSI had a significant positive correlation with Yp (r = 0.26) and TOL (r = 0.76), indicating that higher Yp and TOL values are associated with higher SSI values. Overall, YSI and SSI were inversely related, with YSI reflecting yield stability under stress conditions, while SSI reflected stress susceptibility (Table 3.6). G132, G10 and G134 had the lowest SSI (0.48, 0.57 and 0.59, respectively), but the highest YSI (0.79, 0.75 and 0.74, respectively). G136 had the highest SSI (1.52), but the lowest YSI (0.32), with non-stress and stress yields of 56.6 and 18.3 g/plant, respectively (Appendix 2).

Table 3.6 Correlation of stress tolerance indices in sorghum landraces based on grain yield

Indices	Yp	Ys	TOL	MP	GMP	YSI	SSI
Ys	0.77**						
TOL	0.82**	0.26**					
MP	0.97**	0.91**	0.64**				
GMP	0.94**	0.95**	0.56**	0.99**			
YSI	-0.26**	0.40**	-0.76**	-0.01	0.09*		
SSI	0.26**	-0.40**	0.76**	0.01	-0.09*	-1.00**	
STI	0.93**	0.94**	0.56**	0.99**	0.99**	0.09*	-0.09*

** P ≤ 0.01; * P ≤ 0.05; Yield in non-stress conditions (Yp); Yield in stress conditions (Ys); Tolerance index (TOL); Mean productivity (MP); Geometric mean productivity (GMP); Yield stability index (YSI); Stress susceptibility index (SSI); Stress tolerance index (STI).

The STI, as well as GMP and MP are the averages of genotype yield under non-stress and stress conditions, with higher values indicating greater yield potential under such conditions. As a result, G189, G210, G125, G353, G324, G319, G219, G187, G229 and G269 had the highest STI, GMP and MP values (Appendix 2). Based on genotype yield averages under non-stressed and stress conditions, there were strong correlations between Yp, Ys, STI, GMP and MP (Table 3.6).

3.4.4 Performance of evaluated sorghum landraces for yield and related traits across environments

The sorghum landraces were analysed and ranked individually in different environments and genotypes G210, G125, G189 and G219 exhibited the highest GY across all tested environments: MK20, MK21, MS20 and MS21. Meanwhile, genotypes G324, G319 and G187 displayed higher GY specifically in the environments MK20, MK21 and MS21, whereas G353 consistently performed well in both locations during 2021. In TSW, genotypes G208, G333, G282, G126, G358 and G188 demonstrated superior performance across MK20, MK21 and MS21, while G274, G19 and G292 excelled in MK20 and MS21. Genotypes G8, G190 and G189 showed notable performance in MS20 in TSW. For HWT, genotypes G269, G219, G95, G210, G73, G68, G333, G332, G219 and G188 exhibited the highest performance across all environments. Similarly, in HLP, genotypes G296, G311, G157, G192, G152, G17, G28, G278, G298 and G279 displayed the highest performance across all environments. In HWD, genotypes G321, G322, G211, G219, G353, G282 and G136 demonstrated the highest performance across all environments. Regarding DTF, genotypes G78, G75, G134, G3 and G296 were the earliest across all environments, whereas genotypes G294, G111, G121, G132, G365 and G75 displayed early DTM (Table 3.7).

In terms of GY performance, 30 landraces had a 23% higher GY than the check cultivar (Bonsa) (Table 3.8). Across environments, a mean yield of 48.3 g/plant for these 30 landraces was obtained. ETSL100956 (G189) had the highest grain yield of 74.2 g/plant, which was 51.8% higher than the check cultivar (Bonsa). The highest grain yield and ten top ranked landraces were ETSL100956, ETSL101061, ETSL100639, ETSL101676, IS38303, ETSL101640, ETSL101127, ETSL100951, ETSL101190 and ETSL101353, ranging from 67.15 to 74.19 g/plant. The top 30 yielding landraces had higher TSW and HWT than the commercial check cultivars. When compared to

check cultivars like Bonsa and Dagim, most of the high-yielding genotypes were late to medium-maturing, with tall to medium PHT (Table 3.8).

3.4.5 Cluster analysis

Cluster analysis divided the 361 landraces (the 354 entries of season 1 and the additional seven landraces included in season 2), as well as the four checks, into seven main groups based on quantitative traits (Table 3.9, Figure 3.2 and Table 3.10). The core sorghum landrace collection represents various sorghum growing areas and agro-climatic zones in Ethiopia (dry highland, dry lowland, dry intermediate, wet lowland, wet highland and wet intermediate altitude). However, data on the precise growing regions and agro-climatic zones were not captured. In this study, different clusters were identified based on certain traits. The length of branches in the dendrogram reflects the genetic distance or dissimilarity between clusters. Longer branches indicate greater genetic differences, while shorter branches indicate closer genetic relationships (Figure 3.2). These clusters were characterised by their values for the traits, which could be categorised as high, moderate or low when compared to values of those traits in other clusters.

Table 3.7 Range and mean values and 10 top-ranking genotypes for each trait in each environment

Trait	Envt	Range	Mean	1	2	3	4	5	6	7	8	9	10
GY	MK20	42.18-126.33	79.59	G189	G324	G269	G229	G319	G187	G339	G210	G125	G293
	MK21	29.83-62.49	44.67	G324	G125	G189	G319	G353	G210	G219	G192	G187	G143
	MS20	23.18-60.24	36.18	G210	G125	G189	G190	G66	G228	G188	G121	G219	G181
	MS21	12.55-54.14	32.66	G353	G324	G187	G189	G319	G210	G219	G229	G125	G342
TSW	MK20	21.05-49.77	34.86	G208	G333	G282	G126	G358	G274	G19	G280	G292	G188
	MK21	22.19-64.04	33.92	G208	G214	G188	G282	G358	G333	G38	G221	G192	G126
	MS20	29.73-31.66	30.64	G8	G190	G189	G112	G342	G300	G205	G221	G23	G200
	MS21	19.66-37.53	28.26	G208	G282	G333	G274	G126	G292	G358	G188	G19	G38
HWT	MK20	71.18-202.34	121.10	G269	G291	G95	G210	G73	G68	G333	G332	G219	G188
	MK21	50.61-102.88	70.50	G269	G291	G95	G210	G73	G68	G333	G332	G219	G188
	MS20	55.2-128.70	83.18	G269	G291	G95	G210	G73	G68	G333	G332	G219	G188
	MS21	39.65-77.16	54.76	G209	G332	G269	G219	G291	G95	G68	G73	G364	G252
HLP	MK20	11.55-47.54	25.81	G296	G311	G157	G298	G278	G152	G17	G28	G192	G279
	MK21	12.63-45.47	26.11	G296	G311	G157	G17	G192	G28	G278	G298	G152	G279
	MS20	16.16-43.56	27.78	G311	G296	G157	G298	G152	G17	G28	G192	G278	G279
	MS21	44.03-11.44	24.76	G296	G311	G157	G192	G152	G17	G28	G278	G298	G279
HWD	MK20	5.32-9.63	6.89	G321	G322	G211	G282	G136	G275	G219	G205	G186	G97
	MK21	4.59-7.93	5.93	G322	G321	G211	G219	G282	G353	G136	G168	G100	G205
	MS20	7.50-8.40	7.86	G322	G321	G211	G282	G219	G353	G136	G168	G100	G275
	MS21	5.87-8.66	6.97	G321	G322	G211	G219	G353	G282	G136	G275	G100	G168
PHT	MK20	153.56-351.25	270.96	G279	G74	G120	G73	G97	G192	G251	G137	G276	G322
	MK21	126.58-301.41	232.53	G250	G137	G266	G139	G267	G70	G255	G311	G297	G138
	MS20	117.99-344.94	256.00	G290	G154	G238	G160	G298	G222	G186	G296	G128	G252
	MS21	174.61-373.73	302.47	G32	G328	G28	G267	G117	G264	G187	G104	G250	G326
ILP	MK20	5.77-18.30	12.15	G286	G74	G81	G284	G266	G283	G274	G60	G291	G269
	MS20	5.99-27.54	18.47	G276	G160	G131	G242	G296	G78	G250	G166	G311	G180
	MS21	16.39-31.36	22.81	G18	G32	G8	G297	G237	G329	G244	G245	G20	G16
INP	MK20	12.66-35.87	19.80	G296	G311	G157	G38	G78	G166	G131	G310	G178	G84
	MK21	5.68-14.55	10.83	G274	G209	G10	G20	G286	G196	G186	G74	G83	G189
	MS20	7.16-12.99	10.80	G189	G20	G284	G186	G155	G187	G46	G274	G252	G271
	MS21	9.23-13.61	11.87	G20	G189	G155	G284	G186	G187	G274	G252	G83	G65
DTF	MK20	74.83-138.09	104.37	G75	G296	G134	G3	G78	G131	G254	G130	G84	G344
	MK21	73.59-128.46	99.52	G78	G75	G134	G3	G296	G132	G130	G131	G12	G272
	MS20	69.67-117.91	92.91	G78	G134	G75	G111	G132	G254	G131	G3	G362	G296
	MS21	80.34-167.40	122.10	G3	G362	G361	G75	G134	G296	G132	G213	G87	G78
DTM	MK20	141.65-179.27	165.77	G294	G111	G121	G132	G63	G365	G75	G254	G275	G296
	MK21	143.16-161.62	155.18	G294	G111	G121	G63	G365	G272	G75	G230	G132	G316
	MS20	124.44-131.61	129.29	G294	G316	G111	G254	G132	G75	G121	G3	G78	G272
	MS21	134.93-168.67	152.44	G254	G294	G272	G121	G365	G111	G360	G75	G132	G275

Envt = Environment; GY = Grain yield per plant (g); TSW = Thousand seed weight (g); HWT = Head weight per plant (g); HLP = Head length per plant (cm); HWD = Head width per plant (cm); PHT = Plant height per plant (cm); ILP = Internode length per plant (cm); INP = Internode number per plant (count); DTF = Days to 50% flowering (days); DTM = Days to 90% maturity (days).

Table 3.8 Yield and agronomic traits mean values of the top 30 yielding landraces (≥23% higher yield than the best commercial check) compared to four commercial checks across four environments

Code	Genotype	GY	Rank	% Above best check	TSW	HWT	HLP	HWD	PHT	ILP	INP	DTF	DTM
G189	ETSL100956	74.20	1	51.80	38.80	79.90	13.00	7.10	270.40	19.30	14.49	102.30	150.50
G210	ETSL101061	71.90	2	47.00	32.50	119.40	14.50	7.20	276.40	16.90	14.31	105.60	151.20
G125	ETSL100639	71.20	3	45.60	34.30	97.60	21.70	7.40	215.00	15.60	12.59	86.70	142.90
G324	ETSL101676	70.20	4	43.60	36.90	94.40	16.30	6.50	266.50	18.30	13.62	108.60	152.60
G353	IS38303	69.30	5	41.90	36.00	98.50	26.00	8.00	291.40	17.40	13.98	95.70	150.50
G319	ETSL101640	69.10	6	41.30	34.60	103.70	19.70	7.70	261.10	17.10	13.05	95.70	151.40
G219	ETSL101127	68.50	7	40.20	35.50	109.80	23.40	8.10	288.10	17.50	12.29	105.80	150.00
G187	ETSL100951	68.40	8	40.00	34.70	87.10	21.60	7.40	294.50	17.90	14.22	103.10	153.00
G229	ETSL101190	68.20	9	39.50	32.20	100.60	16.50	7.10	264.30	16.80	14.43	106.50	146.80
G269	ETSL101353	67.20	10	37.40	35.50	132.90	18.30	7.40	275.90	16.40	13.05	108.60	153.90
G192	ETSL100977	66.40	11	35.80	38.60	85.10	36.90	6.90	294.00	18.90	14.69	96.70	149.10
G190	ETSL100964	65.60	12	34.20	36.40	106.10	15.20	7.30	276.60	18.40	13.54	103.40	151.50
G201	ETSL101019	65.60	13	34.20	33.20	100.60	28.90	7.20	298.20	18.90	14.07	99.10	148.40
G293	ETSL101477	65.00	14	33.00	35.90	107.80	17.10	7.50	256.40	16.10	13.56	97.10	150.20
G143	ETSL100739	64.40	15	31.70	28.20	94.20	33.20	6.70	293.60	18.60	13.56	116.00	152.60
G339	ETSL101760	64.10	16	31.10	33.20	99.10	30.00	7.20	246.40	17.70	13.33	103.20	155.40
G332	ETSL101719	63.80	17	30.50	36.90	114.30	22.80	6.90	276.20	18.30	12.06	108.40	152.60
G121	ETSL100619	63.70	18	30.20	33.60	89.80	19.70	7.50	215.70	16.80	13.12	83.10	138.20
G188	ETSL100953	63.60	19	30.00	39.60	106.10	29.10	7.40	289.80	18.10	14.08	98.90	149.70
G6	DS10	63.30	20	29.60	34.40	96.10	23.50	7.20	199.10	16.10	13.05	91.50	143.30
G342	ETSL101851	63.30	21	29.40	37.30	92.50	34.70	7.30	237.40	15.20	12.89	93.00	149.10
G344	ETSL101858	62.20	22	27.30	35.00	94.40	28.30	6.90	261.80	16.00	14.36	86.80	144.80
G84	ETSL100406	62.20	23	27.20	34.00	84.90	23.50	7.00	278.70	17.70	14.40	86.30	143.60
G71	ETSL100315	61.90	24	26.70	34.80	105.50	15.50	6.80	266.10	18.90	14.28	100.30	151.00
G291	ETSL101464	61.80	25	26.40	34.70	123.40	14.40	7.80	279.20	17.70	13.82	107.30	156.00
G69	ETSL100307	61.20	26	25.30	37.70	94.30	35.80	6.60	249.10	16.70	12.93	90.80	149.70
G364	PML981488	61.10	27	24.90	32.60	99.10	34.60	6.60	190.30	16.00	12.89	91.30	141.30
G318	ETSL101629	61.00	28	24.80	34.60	101.70	35.80	7.00	292.40	17.30	14.36	98.10	150.10
G203	ETSL101023	61.00	29	24.80	34.30	86.30	34.10	7.30	261.40	18.40	13.38	86.90	143.80
G109	ETSL100529	60.30	30	23.30	36.90	90.30	32.10	7.50	227.80	17.40	12.38	90.30	150.10
Commercial checks													
G4	Bonsa	48.90	152	0.00	27.20	81.30	27.90	6.90	168.30	11.90	10.73	98.70	148.40
G5	Dagim	45.10	231	-7.70	30.20	77.70	33.30	5.90	193.60	14.70	10.71	91.00	145.10
G2	Assosa-1	45.10	232	-7.70	31.10	74.10	25.20	6.80	216.00	17.50	11.38	126.00	158.50
G1	Adukar	41.70	294	-14.70	30.30	66.30	21.00	6.60	190.30	18.00	11.68	128.70	156.40
Mean (N-365)		**48.30**			**31.90**	**83.10**	**26.20**	**6.90**	**266.50**	**17.80**	**13.30**	**104.80**	**150.80**

GY = Grain yield per plant (g); TSW = Thousand seed weight (g); HWT = Head weight per plant (g); HLP = Head length per plant (cm); HWD = Head width per plant (cm); PHT = Plant height per plant (cm); ILP = Internode length per plant (cm); INP = Internode number per plant (count); DTF = Days to 50% flowering (days); DTM = Days to 90% maturity (days).

Table 3.9 Grouping of Ethiopian sorghum landraces into different clusters based on climate and adaptation regions

Sorghum landraces	Ethiopian sorghum climate classifications	No of entries	Region	Cluster
ETSL101000	Dry lowland	1	Afar	
ETSL101246, ETSL101715	Dry highland	2	Amhara	
ETSL100587, ETSL101058, ETSL101059	Dry intermediate	3		
ETSL100141, ETSL101021, ETSL101051, ETSL101056, ETSL101128, ETSL101362, ETSL101389, IS38306, IS38326	Wet highland	9		
ETSL100584	Wet intermediate	1		
ETSL100007, ETSL100016, ETSL100020, ETSL100030, ETSL100060, ETSL100063, ETSL100064, ETSL100079, ETSL100090, ETSL100091, ETSL100105, ETSL100107, ETSL100167, ETSL100169, ETSL100178, ETSL100198, ETSL100200, ETSL100238, ETSL101461, IS25531	Wet lowland	20	Benishangul Gumuz	
ETSL101445, ETSL101586	Dry lowland	2	Dire Dawa	
ETSL100634, ETSL100876, ETSL100877, ETSL100889, ETSL101408	Wet lowland	5	Gambela	
ETSL100801, ETSL101687	Dry highland	2	Oromia	
ETSL100328	Dry intermediate	1		
ETSL101446	Dry lowland	1		
ETSL100409, ETSL100436, ETSL100443, ETSL100491, ETSL100492, ETSL100502, ETSL100565, ETSL100782, ETSL100798, ETSL101218, ETSL101224, ETSL101231, ETSL101255, ETSL101256, ETSL101312, ETSL101321, ETSL101347, ETSL101443, ETSL101523, ETSL101524, ETSL101555	Wet highland	21		I
ETSL100204, ETSL100392, ETSL100445, ETSL100476, ETSL100763, ETSL100851, ETSL101335, ETSL101550	Wet intermediate	8		
ETSL100192, ETSL100474	Wet lowland	2		
ETSL101712	Dry lowland	1	SNNP	
ETSL100568	Wet highland	1		
ETSL100739	Dry lowland	1	Somalia	
ETSL100431	Dry intermediate	1	Tigray	
ETSL101160, ETSL101161	Dry lowland	2		
ETSL101101	Wet highland	1		
Adukar, Assosa-1, ETSL100005, ETSL100022, ETSL100025, ETSL100034, ETSL100040, ETSL100046, ETSL100048, ETSL100066, ETSL100068, ETSL100072, ETSL100100, ETSL100112, ETSL100149, ETSL100170, ETSL100181, ETSL100366, ETSL100367, ETSL100388, ETSL100465, ETSL100575, ETSL100651, ETSL100735, ETSL100771, ETSL100867, ETSL100898, ETSL100970, ETSL101206, ETSL101540, ETSL101643, ETSL101732, ETSL101868, IS38275, IS38290, ETSL100560, ETSL101249, ETSL101310, ETSL101848, ETSL100669	Unknown	40	Unknown	

Table 3.9 Grouping of Ethiopian sorghum landraces into different clusters based on climate and adaptation regions (continued)

Sorghum landraces	Ethiopian sorghum climate classifications	No of entries	Region	Cluster
ETSL100978	Dry lowland	1	Afar	
ETSL100234	Wet intermediate	1	Benishangul Gumuz	
ETSL100246, ETSL100249, ETSL100267	Wet lowland	3		
ETSL101451	Dry lowland	1	Dire Dawa	
ETSL101316	Wet lowland	1	Gambela	
ETSL100808, ETSL101262, ETSL101297	Dry highland	3	Oromia	
ETSL100485, ETSL100494, ETSL100707, ETSL101217, ETSL101225, ETSL101232, ETSL101254, ETSL101258, ETSL101277, ETSL101313, ETSL101410, ETSL101506, ETSL101527, ETSL101556, ETSL101559, ETSL101730, ETSL101766	Wet highland	17		II
ETSL100450, ETSL100699, ETSL100702, ETSL101416, ETSL101453, ETSL101459	Wet intermediate	6		
ETSL100444	Wet lowland	1		
ETSL101081, ETSL101142	Dry highland	2	Tigray	
ETSL101114	Wet highland	1		
ETSL100039, ETSL100044, ETSL100078, ETSL100093, ETSL100096, ETSL100214, ETSL100250, ETSL100488, ETSL100825, ETSL100834, ETSL101325, ETSL101323	Unknown	12	Unknown	
B35	Unknown	1	Unknown	III
ETSL100985	Dry lowland	1	Afar	
ETSL100121, ETSL101369	Dry lowland	2	Amhara	
ETSL100152, ETSL101292	Wet highland	2		
IS38412	Wet lowland	1		
ETSL100023	Wet lowland	1	Benishangul Gumuz	
ETSL100638, ETSL100644, ETSL100661, ETSL100683, ETSL100722, ETSL100890	Wet lowland	6	Gambela	
ETSL100365	Dry highland	1	Oromia	
IS38331	Dry lowland	1		
ETSL100874	Wet highland	1		
Dagim	Wet intermediate	1		
ETSL100297	Dry lowland	1	SNNP	IV
ETSL100689, ETSL101181	Wet highland	2		
ETSL101486, ETSL101491	Dry intermediate	2	Somalia	
ETSL100349, ETSL100350, ETSL100353, ETSL100419, ETSL100430, ETSL100842	Dry highland	6	Tigray	
ETSL101512	Dry intermediate	1		
ETSL100417, ETSL100544, ETSL100550, ETSL101163, ETSL101535, ETSL101571, ETSL101578	Dry lowland	7		
ETSL101118, ETSL101536	Wet highland	2		
Bonsa, ETSL100127, ETSL100184, ETSL100667, ETSL100872, ETSL101749, IS38353, P9830, PML981446, PML981474, TAM428, ETSL100275, ETSL100284	Unknown	13	Unknown	

Table 3.9 Grouping of Ethiopian sorghum landraces into different clusters based on climate and adaptation regions (continued)

Sorghum landraces	Ethiopian sorghum climate classifications	No of entries	Region	Cluster
ETSL100983, ETSL101074	Dry lowland	2	Afar	
ETSL100964	Dry intermediate	1	Amhara	
ETSL100310, ETSL100813	Dry lowland	2		
ETSL100951, ETSL100953, ETSL100956, ETSL101001, ETSL101015, ETSL101029, ETSL101038, ETSL101299, ETSL101654, ETSL101722, IS38279	Wet highland	11		
ETSL101640	Wet intermediate	1		
ETSL101464	Wet lowland	1	Benishangul Gumuz	
ETSL101673, ETSL101686	Dry highland	2	Oromia	
ETSL101444	Dry lowland	1		V
ETSL100332, ETSL100340, ETSL100400, ETSL100438, ETSL100768, ETSL101671	Wet highland	6		
ETSL100593, ETSL101337, ETSL101346, ETSL101353, ETSL101418, ETSL101436	Wet intermediate	6		
ETSL101470, ETSL101477, ETSL101676	Dry intermediate	3	Somalia	
ETSL101719	Dry lowland	1	Tigray	
ETSL101127, ETSL101404	Wet highland	2		
ETSL100062, ETSL101061, ETSL101190, ETSL101358, ETSL101388, IS38303, PML981476, ETSL100303, ETSL100315	Unknown	9	Unknown	

Table 3.9 Grouping of Ethiopian sorghum landraces into different clusters based on climate and adaptation regions (continued)

Sorghum landraces	Ethiopian sorghum climate classifications	No of entries	Region	Cluster
ETSL100977	Dry lowland	1	Afar	
ETSL101247, ETSL101248	Dry highland	2	Amhara	
ETSL101629	Dry intermediate	1		
ETSL100120, IS38259	Dry lowland	2		
ETSL100150, ETSL100919, ETSL100945, ETSL101006, ETSL101014, ETSL101023, ETSL101289, ETSL101431, IS38319	Wet highland	9		
ETSL100286, ETSL100619, ETSL100620, ETSL100639, ETSL100642, ETSL100648, ETSL100666, ETSL100888, ETSL100900, ETSL100902, ETSL101293, ETSL101396	Wet lowland	12	Gambela	
ETSL100751, ETSL100786	Dry highland	2	Oromia	
ETSL101858	Dry lowland	1		
ETSL100720, ETSL100785, ETSL100797, ETSL100861, ETSL100863, ETSL101199, ETSL101210, ETSL101215, ETSL101270, ETSL101760	Wet highland	10		
ETSL100406, ETSL100408, ETSL100761, ETSL100792, ETSL100928	Wet intermediate	5		VI
ETSL100716, ETSL100745, ETSL101305, ETSL101698, ETSL101704	Dry lowland	5	SNNP	
ETSL101173, ETSL101180	Wet highland	2		
ETSL101681	Dry intermediate	1	Somalia	
ETSL100428, ETSL100432, ETSL100529, ETSL100841, ETSL101508, ETSL101515, ETSL101563	Dry highland	7	Tigray	
ETSL100427, ETSL100539, ETSL101568	Dry intermediate	3		
ETSL100307, ETSL101138, ETSL101851	Dry lowland	3		
ETSL101097, ETSL101110	Wet highland	2		
DS10, ETSL100233, ETSL100561, ETSL100837, ETSL100862, ETSL100893, ETSL100895, ETSL101019, ETSL101084, ETSL101274, ETSL101564, ETSL101723, ETSL101725, IS38254, IS38255, PML981488, ETSL100292, ETSL100916, ETSL100921, ETSL101856	Unknown	20	Unknown	
ETSL100805	Wet intermediate	1	Oromia	
ETSL101492	Dry intermediate	1	Tigray	VII
ETSL101561	Dry highland	1		

Unknown = No available data for landraces, SNNP = Southern Nations, Nationalities and People; lowland (< 1600 m above sea level), intermediate elevation (1600-1900 m.a.s.l.), highland (> 1900 m.a.s.l.), wet (high rainfall and long growing period), dry (low rainfall and short growing period).

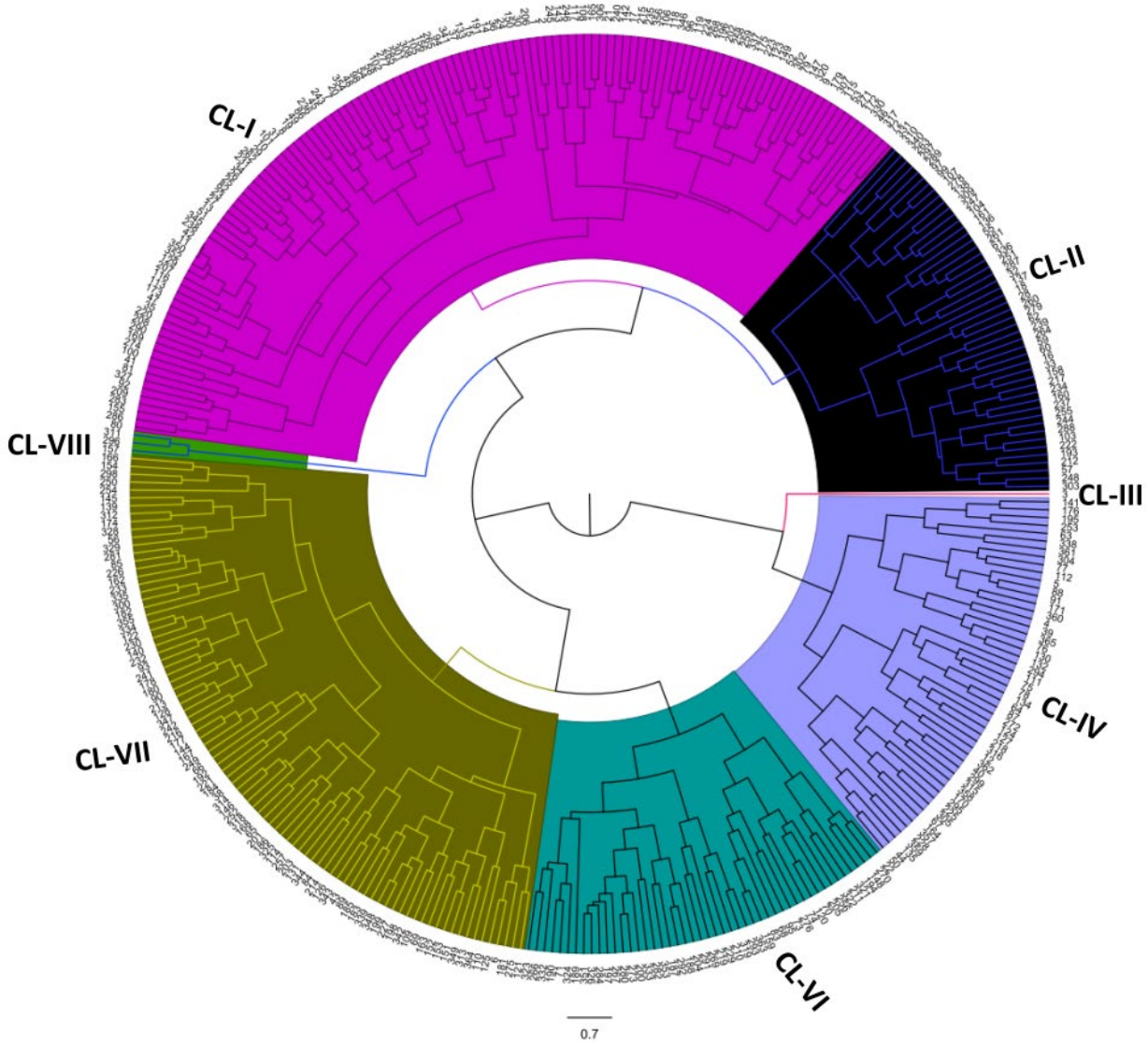


Figure 3.2 Dendrogram of cluster analysis displaying the relationship among landraces

Longer branches indicate greater genetic differences, while shorter branches indicate closer genetic relationships

Table 3.10 Summary of the cluster analysis displaying the origin of the 361 sorghum landraces and four checks, the number of genotypes and the mean values for 10 phenotypic traits

	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V	Cluster VI	Cluster VII
				Number of landraces			
Source of landraces							
Afar	1	1		1	2	1	
Amhara	15			5	15	14	
Benishangul Gumuz	20	4		1	1		
Dire Dawa	2	1					
Gambela	5	1		6		12	
Oromia	35	27		4	15	18	1
SNNP	2			3		7	
Somalia	1			2	3	1	
Tigray	4	3		16	3	15	2
Unknown	40	12	1	13	9	20	
Total	125	49	1	51	48	88	3
Traits							
GY	45.79	42.01	40.25	44.60	56.94	53.19	35.99
TSW	31.57	27.63	31.32	32.27	34.26	33.46	28.20
HWT	81.91	81.76	62.40	70.67	99.15	84.82	61.25
HLP	25.28	29.83	22.61	26.21	21.11	27.87	42.89
HWD	6.84	7.03	6.30	6.50	7.40	7.00	6.02
PHT	266.64	309.38	114.56	217.40	281.17	263.78	287.47
ILP	17.98	19.87	8.77	15.55	18.12	17.75	19.42
INP	13.14	14.06	8.65	12.18	13.74	13.70	15.97
DTF	112.48	118.20	76.78	91.94	105.20	94.65	85.94
DTM	153.60	154.64	141.77	145.59	152.35	147.42	141.67

SNNP = Southern Nations, Nationalities and People; GY = Grain yield per plant (g); TSW = Thousand seed weight (g); HWT = Head weight per plant (g); HLP = Head length per plant (cm); HWD = Head width per plant (cm); PHT = Plant height per plant (cm); ILP = Internode length per plant (cm); INP = Internode number per plant (count); DTF = Days to 50% flowering (days); DTM = Days to 90% maturity (days).

Most of the assessed landraces (125 or 34.25%) grouped in cluster I, which contained landraces from Afar (1), Amhara (15), Benishangul Gumuz (20), Dire Dawa (2), Gambela (5), Oromia (35), Southern Nations, Nationalities and Peoples *(*SNNP) (2), Somalia (1), Tigray (4) and unknown (data not available; 40). All six the sorghum agro-climatic zones contributed landraces to this group. Landraces in cluster I had a moderate mean GY and HWT (45.79 and 81.91 g/plant, respectively), medium to late flowering and maturity (112 and 154 days, respectively) and medium to tall PHT. Other characteristics were also moderate compared to other clusters (Table 3.10).

Cluster II contained 49 landraces [Afar (1), Benishangul Gumuz (4), Oromia (27), Dire Dawa (1), Gambela (1), Tigray (3) and unknown (12)] showing low mean GY and TSW (42.01 g/plant and 27.63 g, respectively), the tallest PHT (309.36 cm), the longest ILP (19.87 cm) and latest flowering and maturity (118 and 155 days, respectively) (Table 3.10).

Cluster III contained one genotype with a low mean GY (40.25 g/plant) and HWT (62.40 g/plant), the shortest PHT and ILP of 114.56 and 8.77 cm, respectively and the earliest DTF and DTM of 77 and 142 days, respectively; a moderate TSW of 31.32 g and the lowest INP number (9). Fourteen percent of the landraces were found in cluster IV and were from Afar (1), Amhara (5), Benishangul Gumuz (1), Oromia (4), Gambela (6), SNNP (3), Somalia (2), Tigray (16) and unknown (13). This group of landraces had medium to late mean flowering and maturity (92 and 146 days, respectively), moderate GY and HWT (44.6 and 70.67 g/plant) and medium to short PHT and ILP (217.4 and 15.55 cm). Furthermore, in comparison to other groups, other characteristics were moderate.

Forty-eight of the assessed landraces (13%) were found in cluster V, which originated from Afar (2), Amhara (15), Benishangul Gumuz (1), Oromia (15), Somalia (3), Tigray (3) and unknown (9). Cluster V's landraces had the highest mean GY and HWT (56.94 and 99.15 g/plant, respectively), the highest TSW (34.26 g), highest HWD (7.40 cm), medium to late flowering and maturity (105 and 152 days, respectively) and medium to high PHT and ILP (281.17 and 18.12 cm).

Cluster VI contained 88 genotypes, with a medium mean GY and HWT of 53.19 and 84.82 g/plant respectively. A medium TSW of 33.46 g, a moderate HWD of 7.00 cm and a moderate PHT of 263.78 cm, 95 and 147 DTF and DTM and a moderate ILP of 17.75 cm.

Cluster VII had a low number of landraces (3). Landraces in this cluster exhibited the lowest mean GY and HWT (35.99 and 61.25 g/plant, respectively), the smallest HWD (6.02 cm), the earliest flowering and maturity (85 and 142 days, respectively) and medium to high PHT and ILP (287.47 and 19.42 cm, respectively).

3.4.6 Inter- and intra-clustering distances

Based on Mahalanobis' D^2 statistics, the pairwise generalised squared distance (D^2) between the seven clusters showed that the highest inter-cluster genetic distance was recorded between clusters II and VI ($D^2 = 13.28$), followed by clusters II and IV ($D^2 = 12.54$). Cluster VII had the highest intra-cluster distance ($D^2 = 9.67$), indicating that genotypes in this cluster had comparatively high genetic divergence (Table 3.11).

Table 3.11 Inter- (below diagonal) and intra- (diagonal) cluster Mahalanobis' D^2 statistics based genetic distances

Cluster	I	II	III	IV	V	VI	VII
I	**5.63**						
II	10.61	**9.67**					
III	7.44	11.76	**6.07**				
IV	7.05	12.54	8.70	**6.88**			
V	6.38	11.05	6.70	6.51	**5.49**		
VI	9.35	13.28	9.94	8.94	8.74	**7.82**	
VII	8.07	11.45	7.51	9.10	7.29	10.87	**5.42**

3.4.7 Trait correlation

Even small correlations were statistically significant because the large dataset enabled the detection of even subtle patterns. GY was positively and significantly correlated with TSW (r = 0.44), HWT (r = 0.62) and HWD (r = 0.38) (Table 3.11). DTF, DTM, HLP and ILP correlated significantly negatively with GY. These negative correlations may aid in the selection of early maturing genotypes with high GY for areas prone to terminal drought. PHT, INP, ILP, DTF and DTM were significantly positively correlated. TSW was significantly negatively correlated with DTF, DTM, HLP, INP, ILP and PHT.

Table 3.12 Phenotypic correlation coefficients of 10 characteristics from 365 sorghum landraces for four trials

Trait	GY	TSW	HWT	HLP	HWD	PHT	ILP	INP	DTF
TSW	0.44**								
HWT	0.62**	0.21**							
HLP	-0.22**	-0.25**	-0.23**						
HWD	0.38**	0.077	0.56**	-0.19**					
PHT	-0.04	-0.22**	0.27**	0.13*	0.27**				
ILP	-0.12*	-0.20**	0.14**	0.10*	0.20**	0.79**			
INP	-0.069	-0.24**	0.23**	0.27**	0.24**	0.99*	0.80**		
DTF	-0.29**	-0.28**	0.16**	-0.092	0.10*	0.43**	0.43**	0.41**	
DTM	-0.18**	-0.11*	0.25**	-0.096	0.15**	0.46**	0.41**	0.44**	0.83**

** $P \leq 0.01$; * $P \leq 0.05$; GY = Grain yield per plant (g); TSW = Thousand seed weight (g); HWT = Head weight per plant (g); HLP = Head length per plant (cm); HWD = Head width per plant (cm); PHT = Plant height per plant (cm); ILP = Internode length per plant (cm); INP = Internode number per plant (count); DTF = Days to 50% flowering (days); DTM = Days to 90% maturity (days).

3.5 Discussion

The genetic variation present in the collection of sorghum landraces evaluated for GY and other agronomic characteristics across the test environments indicated that there is potential for selecting high-yielding landraces that could be used in variety development. Low mean, range, heritability and genetic variances were observed for agronomic traits in stress environments compared to non-stress environments, except for DTF. Heritability assessment provides information about trait architecture and high heritability improves selection efficiency, optimising resource use. GY generally has a low heritability because it is controlled by many genes and selection progress is generally slow (Ogunniyan and Olakojo, 2014). Somegowda et al. (2021) reported that genotypic variances were higher in non-stress conditions than in stress conditions for agronomic traits, except for DTF, which correlates with results obtained in the current study. Rainfall alone can be the most significant cause of differences in sorghum production for a given location. The soil moisture level is recharged after rain and when drought occurs, the surface soil moisture gradually decreases toward the end of the crop cycle, affecting yield and yield-related components (Rossato et al., 2017).

Climate change will have an impact on yield, so maintaining selection for high yield, either through direct selection or selection of correlated secondary traits, is important. This study showed that

drought increased physiological growth stages, shortened regular growth periods and reduced final yield. Drought stress affects the entire panicle during microsporogenesis and reduces yield at all stages of panicle development. Kilic and Yağbasanlar (2010) reported that drought stress reduced mean grain yield by 44.6% and grain weight by 14%. Manjarrez-Sandoval et al. (1989) reported that drought caused panicle abortion and a 25-55% reduction in the number of grains per mature panicle; later drought stress periods reduced individual grain weight by up to 50%. Under stress conditions PHT and DTF and DTM were found to be shorter than under non-stress conditions except under MS21. In the current study, a small number of genotypes did perform better than commercial varieties in both drought-stress and irrigated environments. For all environments, genotypes G189, G210, G125, G353, G324, G319, G219, G187, G229 and G269 had the highest Yp, Ys, MP, GMP and STI and had high GY in both environments, indicating that these genotypes could perform well in METs, which will be the next stage of evaluation. Wasae (2021) and Mwamahonje et al. (2021) found high TOL values in some sorghum genotypes and concluded that such genotypes are susceptible to drought stress. In the current study, four genotypes (G18, G132, G141 and G296) had low TOL values, indicating good drought tolerance.

Landraces were divided into seven clusters using ordinal variable cluster analysis. Landraces from similar altitude classes and eco-sites largely grouped together. As a result, the clustering pattern of origin also reflects the distribution patterns of different cultivated sorghum races in Ethiopia (Stemler et al., 1977; Ayana and Bekele, 1999). Thus, genotype adaptation is identified and genotype evaluation for potential adaptation is improved when classification is based on multiple agronomic characteristics (Abdi et al., 2002). Furthermore, as Desmae et al. (2016) hypothesised, clustering of accessions from geographically similar areas indicates the evolution of co-adaptive quantitative character associations. Clustering analysis was conducted to identify different types of sorghum landraces based on their trait values and growing conditions (lowland, intermediate and highland) in specific agro-ecologies within different regions. Results of the clustering analysis support the existence of distinct sorghum landraces corresponding to these different ecological regions.

GY had positive and significant correlations with HWT, TSW and HWD. Akatwijuka et al. (2019) reported that the positive correlation between GY and panicle weight suggests that selection for

panicle weight will also increase yield. In this study, GY was inversely related to HLP, DTF and DTM, supporting previously reported results (Enyew et al., 2021). These negative correlations may be useful in selecting early maturing genotypes with high GY in stress areas.

3.6 Conclusions

This study revealed the existence of significant genetic variation in Ethiopian sorghum landraces for GY and related traits. Drought stress reduced yield, but promising landraces with drought tolerance have been identified. The geographic clustering of landraces emphasises the importance of regional diversity in breeding programmes. Additionally, the positive correlations found between GY and certain traits provide valuable insights for selecting desirable traits in sorghum breeding. Overall, these conclusions highlight the potential for harnessing the genetic diversity present in Ethiopian sorghum landraces to develop improved varieties that are better adapted to drought conditions and possess high GY and related traits.

3.7 References

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

Akatwijuka, R., Rubaihayo, P.R. and Odong, T.L. (2019). Correlations and path analysis of yield traits in sorghum grown in southwestern highlands of Uganda. *African Crop Science Journal*, 27: 437-444.

Amelework, B., Shimelis, H., Tongoona, P. and Laing, M. (2015). Physiological mechanisms of drought tolerance in sorghum, genetic basis and breeding methods: a review. *African Journal of Agricultural Research*, 10: 3029-3040.

Appa Rao, S., Prasada Rao, K.E., Mengesha, M.H. and Reddy, V.G. (1996). Morphological diversity in sorghum germplasm from India. *Genetic Resources and Crop Evolution,* 43: 559-567.

Ayana, A. and Bekele, E. (1999). Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resources and Crop Evolution*, 46: 273-284.

Brown, V.A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, 4: 1-19.

Butler, D.G., Cullis, B.R., Gilmour, A.R., Gogel, B.J. and Thompson, R. (2017). ASReml-R /reference manual version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK, pp. 30-49.

Cullis, B.R.C., Mith, A.B.S. and Oombes, N.E.C. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological and Environmental Statistics*, 11: 381-393.

Deepashree, G., Raut, N., Gasti, V.D., Manjunathagowda, D. C., Chittapur, R., Gunnaiah, R. and Chinthan, K.N. (2023). Molecular and morphological diversity among the cluster bean [*Cyamopsis tetragonoloba* (L.) Taub.] genotypes. *Genetic Resources and Crop Evolution*, 70: 159-168.

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: 3111-3122.

Deutsch, E.W., Mendoza, L., Shteynberg, D., Slagel, J., Sun, Z. and Moritz, R.L. (2015). Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *Proteomics-Clinical Applications*, 9: 745-754.

Enyew, M., Feyissa, T., Geleta, M., Tesfaye, K., Hammenhag, C. and Carlsson, A.S. (2021). Genotype by environment interaction, correlation, AMMI, GGE biplot and cluster analysis for grain yield and other agronomic traits in sorghum (*Sorghum bicolor* L. Moench). *Plos One*, 16: 258211.

Fernandez, G.C. (1992). Effective selection criteria for assessing plant stress tolerance. In: Proceedings of the International Symposium on Adaptation of Vegetables and other Food Crops in Temperature and Water Stress, Aug. 13-16, Shanhua, Taiwan, pp. 257-270.

Fischer, R.A. and Maurer, R. (1978). Drought resistance in spring wheat cultivars. I. Grain yield responses. *Australian Journal of Agricultural Research*, 29: 897-912.

Gavuzzi, P., Rizza, F., Palumbo, M., Campanile, R.G., Ricciardi, G.L. and Borghi, B. (1997). Evaluation of field and laboratory predictors of drought and heat tolerance in winter cereals. *Canadian Journal of Plant Science*, 77: 523-531.

Geleta, N., Labuschagne, M.T., Osthoff, G., Hugo, A. and Bothma, C. (2005). Physical and chemical properties associated with food quality in sorghum. *South African Journal of Plant and Soil*, 22: 175-179.

Girma, G., Nida, H., Seyoum, A., Mekonen, M., Nega, A., Lule, D., Dessalegn, K., Bekele, A., Gebreyohannes, A., Adeyanju, A. and Tirfessa, A. (2019). A large-scale genome-wide association analysis of Ethiopian sorghum landrace collection reveals loci associated with important traits. *Frontiers in Plant Science*, 10: 691.

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: 431487.

Harlan, J.R. (1969). Ethiopia: a center of diversity. *Economic Botany*, 23, 309-314.

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

Husain, M.M., Reid, J.B., Othman, H. and Gallagher, J.N. (1990). Growth and water use of faba beans (*Vicia faba)* in a sub-humid climate I. Root and shoot adaptations to drought stress. *Field Crops Research*, 23: 1-17.

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

Kilic, H. and Yağbasanlar, T. (2010). The effect of drought stress on grain yield, yield components and some quality traits of durum wheat (*Triticum turgidum* ssp. *durum*) cultivars. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 38: 164-170.

Madhusri, D., Lavanya, G.R. and Elangovan, M. (2022). Phenotypic diversity of finger millet (*Eleusine coracana* (L.) Gaertn.) genotypes for grain yield characters. *International Journal of Plant and Soil Science*, 34: 139-148.

Mahalanobis, P.C. (2018). On the generalized distance in statistics. *Sankhyā: The Indian Journal of Statistics, Series A,* 80: S1-S7.

Manjarrez-Sandoval, P, González-Hernández, V.A., Mendoza-Onofre, L.E. and Engleman, E.M. (1989). Drought stress affects the grain yield and panicle development of sorghum. *Canadian Journal of Plant Science*, 69: 631-641.

Manly, B.F. (1986). Randomization and regression methods for testing for associations with geographical, environmental and biological distances between populations. *Researches on Population Ecology*, 28: 201-218.

Matloff, N. (2011). The art of R programming: A tour of statistical software design. No Starch Press, pp.152.

Mekbib, F. (2006). Farmer and formal breeding of sorghum (*Sorghum bicolor* (L.) Moench) and the implications for integrated plant breeding. *Euphytica*, 152: 163-176.

MoA (Ministry of Agriculture) (1998). Agro-ecological zones of Ethiopia: natural resource management and regulatory department. Ethiopian Ministry of Agriculture, Addis Ababa.

Mwamahonje, A., Eleblu, J.S.Y., Ofori, K., Feyissa, T., Deshpande, S. and Tongoona, P. (2021). Evaluation of traits' performance contributing to drought tolerance in sorghum. *Agronomy*, 11: 1698.

Ogunniyan, D.J. and Olakojo, S.A. (2014). Genetic variation, heritability, genetic advance and agronomic character association of yellow elite inbred lines of maize (*Zea mays* L.). *Nigerian Journal of Genetics*, 28: 24-28.

Rossato, L., Alvala, R.C.D.S., Marengo, J.A., Zeri, M., Cunha, A.P.D.A., Pires, L.B. and Barbosa, H.A. (2017). Impact of soil moisture on crop yields over Brazilian semiarid. *Frontiers in Environmental Science,* 5: 73.

Sehgal, A., Sita, K., Siddique, K.H., Kumar, R., Bhogireddy, S., Varshney, R.K., Hanumantha Rao, B., Nair, R.M., Prasad, P.V. and Nayyar, H. (2018). Drought and/or heat-stress effects on seed filling in food crops: impacts on functional biochemistry, seed yields and nutritional quality. *Frontiers in Plant Science,* 9: 1705.

Smith, A.B. and Cullis, B.R. (2018). Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. *Euphytica*, 214: 1-19.

Smith, A., Cullis, B. and Thompson, R. (2001). Analysing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics,* 57: 1138-1147.

Somegowda, V.K., Vemula, A., Naravula, J., Prasad, G., Rayaprolu, L., Rathore, A. and Deshpande, S.P. (2021). Evaluation of fodder yield and fodder quality in sorghum and its interaction with grain yield under different water availability regimes. *Current Plant Biology,* 25: 100191.

Stemler, A.B.L., Harlan, J.R. and De Wet, J.M.J. (1977). The sorghums of Ethiopia. *Economic Botany*, 31: 446-460.

Tibshirani, R., Walther, G. and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 63: 411-423.

Wasae, A. (2021). Evaluation of drought stress tolerance based on selection indices in Haricot Bean varieties exposed to stress at different growth stages. *International Journal of Agronomy*, 2021:1-9.

Yang, J. and Delpha, C. (2022). An incipient fault diagnosis methodology using local Mahalanobis distance: Detection process based on empirical probability density estimation. *Signal Processing*, 190: 108308.

CHAPTER 4

ENVIRONMENTAL INFLUENCES ON, AND GENOTYPE BY ENVIRONMENT INTERACTIONS FOR KERNEL COMPOSITION OF AN ETHIOPIAN SORGHUM LANDRACE COLLECTION

4.1 Abstract

Sorghum has a distinct advantage over other cereals in terms of beneficial bioactive compounds and climate resilience. Low grain protein, Fe and Zn concentrations are a frequent problem in food crops and affect the crop's nutritional value, particularly in cereals. Many people in Africa rely heavily on sorghum as a source of energy and micronutrients. For the creation of improved cultivars with high nutritional value, it is crucial to determine genetic variation, GEI, and the relationship between nutritional traits. In the current study, GEI for starch, protein, amylose, amylopectin, ash, moisture, Fe and Zn was highly significant. A total of 361 sorghum landraces and four commercial checks were grown under well-watered, wet intermediate and natural drought conditions for two consecutive seasons. In stress environments, starch content decreased by 5%, while protein content increased by 21% when compared to the wet intermediate conditions. The Zn content was 60% and 55% higher in wet intermediate environments than in irrigated and stress environments, respectively. The Fe content was 21% and 13% higher in wet intermediate environments than in irrigated and stress environments, respectively. Wet intermediate environments had higher total organic carbon, organic matter, nitrogen and available phosphorus soil levels than soils in irrigated and stress environments.

The most reliable and high-ranking genotypes can be identified using the genotype main effects with genotype by environment interaction (GGE) ranking biplots. Genotypes G246, G84, G4, and G175 had the highest starch contents, whereas G218, G43, and G161 had the highest protein contents. The highest Fe content was found in the G137 and G142 genotypes, with the G264 having the highest Zn content.

High grain moisture content and biotic and abiotic factors likely contributed to decrease starch and mineral contents. Starch showed negative correlations with protein, TSW, Zn, ash and moisture,

while it had positive correlations with GY, amylose and Fe. Protein positively correlated with TSW and moisture, and negatively with GY, amylose, Fe, Zn and ash. Fe, Zn and ash were positively correlated, indicating their interrelation. Grain moisture content had negative correlations with starch, amylose, Zn, Fe and ash, but a positive correlation with protein content and TSW.

4.2 Introduction

Sorghum is a cereal grain native to the arid regions of sub-Saharan Africa. The genetic and geographical diversity of sorghum cause significant variation in agronomic and grain quality attributes. Genetic improvement has played an important role in improving sorghum quality, while environmental factors such as soil quality, temperature, rainfall and pests can also have a significant impact on sorghum growth and quality. Thus, genotype and location variation will allow for the selection of the most desirable traits at a given location (Kaufman et al., 2018). Sorghum is grown in arid and semi-arid areas of Ethiopia due to its tolerance to heat and drought stress. Drought and heat stress have been shown to have a significant impact on grain quality in other cereals such as wheat, barley and maize. Drought and high-temperature stress have been shown to alter protein formation and crosslinking in wheat and maize grains (Thitisaksakul et al., 2012; Ashraf, 2014). It was reported that grain physical characteristics as well as chemical composition and properties were significantly influenced by genotype, location and their interactions (Kaufman et al., 2018). Gemechu et al. (2016) assessed the impact of climate change on agricultural production in the Miesso district of the West Hararghe zone. The study found that climate change related factors reduced agricultural productivity. According to Piguet (2003), low rainfall and poor harvests in Hararghe prompted the United Nations Emergency Unit for Ethiopia to organise a follow-up assessment mission during the small rainy season (*Belg*).

Adaptability and yield stability are important measures for crop cultivation in different agro-climatic regions. The stability and adaptability of genotypes across different environments can be assessed using various statistical tools such as joint regression (Finlay and Wilkinson, 1963) and genotype main effects in addition to GGE biplots (Yan, 2001). GGE biplots serve as the most efficient and commonly used multivariate models for determining genotype stability, adaptability and ranking, as well as for identifying suitable mega-environments (Kang, 2003; Yan et al., 2007; Yan and Holland, 2010; Yan and Tinker, 2005, 2006).

Most sorghum landrace studies have focused on the evaluation of GY, with less emphasis placed on grain quality traits using MET data. Although GEI has been used to assess the stability of improved varieties of sorghum using MET data (Rao et al., 2011; Admas and Tesfaye, 2017; Worede *et al.*, 2020), information on the stability of sorghum landraces for quality traits using GGE biplot models is lacking. As a result, the objectives of this research were to assess chemical composition GE interaction, drought effects, environmental influence, the performance and stability of sorghum landraces, the correlation of grain quality traits and the representativeness and discriminating ability of different environments where sorghum is grown.

4.3 Materials and methods

4.3.1 Study areas and crop management

This study quantified GEI for the wet intermediate (Jimma) and dry lowlands, specifically Melkassa (irrigation) and Miesso (stress) agroecology from 2020 to 2021. The two dry lowland experimental sites are found in the central rift valley of Ethiopia. Jimma is in the wet intermediate agroecology with good rainfall in the cropping season. Both Miesso and Melkassa are lowland sorghum production areas and moisture-stress sorghum testing sites in Ethiopia (Figure 4.1) (MoA 1998). Tables 4.1 to 4.3 and Appendix 1 show data for the three sites on planting and harvesting dates, rainfall patterns, relative humidity, maximum/minimum temperatures and soil analytical data for macronutrients and organic matter. The supplemental irrigation was provided at Melkassa as described in section 3.3.2. Seeds were drilled in the row at 75 cm between rows soon after ploughing and then thinned to maintain 15 cm intra-row spacing at all sites. Basal fertiliser was applied at the recommended rate and other agronomic practices were applied for normal sorghum production in Ethiopia.

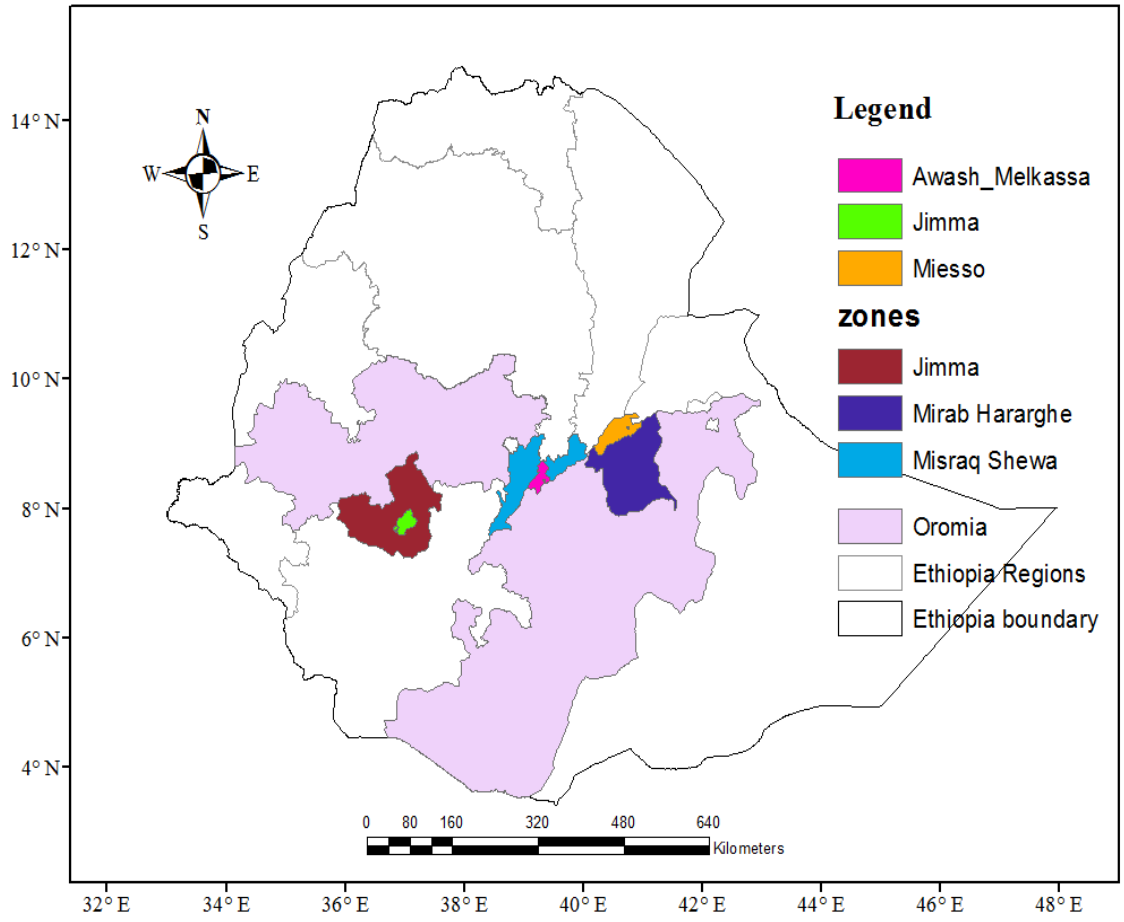


Figure 4.1 The three test locations in Ethiopia, Jimma, Miesso and Melkassa, depicted on a map using an ArcGIS geographic information system (Redlands, 2011).

Jimma, Miesso and Melkassa are situated in the Oromia regional state's Jimma, west Hararghe (Mirab Haraghe) and east Hararghe (Misrak Hararghe) zones, respectively.

Table 4.1 Long-term average temperature and rainfall properties for each trial site

Location	Longitude	Latitude	Altitude (m.a.s.l)	Soil type	Rainfall (mm)	Minimum T°C	Maximum T°C
Miesso	39°21'E	8˚30'N	1 470	Clay loam	570	16	31
Melkassa	39°9'E	14˚6'N	1 550	Clay loam	763	20.4	34
Jimma	36°47′ E	7˚40′ N	1 753	Clay	1 500	20	29

m.a.s.l. = meter above sea level; T°C = temperature in degrees centigrade.

Table 4.2 The planting and harvesting dates of trials in two consecutive seasons

	Planting date		Harvesting date	
Location	**Year 1**	**Year 2**	**Year 1**	**Year 2**
Melkassa	Jul 15, 2020	Aug 3, 2021	Jan 1-20, 2021	Jan 10-12, 2022
Miesso	Jul 20, 2020	May 11, 2021	Nov 27- Dec 5, 2020	Aug 30-Dec 24, 2021
Jimma	May 26, 2020	May 6, 2021	Dec 8, 2020 to Jan 8, 2021	Dec 4-30, 2021

4.3.2 Design and data analysis

The 361 Ethiopian sorghum landraces and four improved varieties were evaluated for genetic diversity and GEI in two consecutive seasons (2020 to 2021). Trial design and layout was done as described in section 3.3.1 for Melkassa and Miesso, and Jimma was added as a third site with a trial design and layout similar to Melkassa and Miesso. This experimental design allowed for the testing of effects of both genotype and environment, as well as GEI. The sorghum grain quality and GY data collected from the six environments were subjected to a combined analysis of variance (ANOVA) using the mixed linear model in R software (Butler et al., 2017). METs were analysed by fitting a linear mixed model using the spatial analysis with package ASReml (Smith et al., 2019). Each model consisted of a fixed effect for the targeted trait at each trial, random effects (BLUPs) for genotypes within the trial and spatial error for each trial (Smith and Cullis, 2018). The GGE biplot analyses were done using GENSTAT software (Horn et al., 2018).

Table 4.3 Soil analytical data for some macronutrients and organic matter for the three trial sites

Environment	Soil depth (cm)	pH	EC (µS/cm)	Textural class	%TOC	%OM	%TN	Av. P (ppm)
Jimma	0-30	6.67	96.50	Clay	1.57	2.71	0.16	8.45
	30-60	6.88	72.70	Clay	1.21	2.09	0.13	8.33
Melkassa	0-30	6.79	190.80	Clay loam	1.29	2.23	0.08	7.95
	30-60	7.28	230.00	Clay	1.25	2.16	0.08	7.73
Miesso	0-30	7.25	139.70	Clay loam	1.06	1.83	0.09	8.73
	30-60	7.28	231.50	Clay	1.03	1.78	0.09	8.61

EC = Electrical conductivity; TOC = Total organic carbon; OM = Organic matter; TN = Total nitrogen; Av. P = available phosphorus.

4.3.3 Analysis of sorghum grain quality

The near-infrared spectroscopy (NIRS) calibration model was created at Melkassa Agricultural Research Center to analyse sorghum grain quality, offering a cost-effective and time-efficient alternative to the more expensive and time-consuming wet chemistry analysis. The development of NIRS models for sorghum grain quality analysis involved the use of over 108 Ethiopian sorghum accession samples (Zerihun et al., 2020). The NIRS calibration equations exhibited high coefficients of determination for calibrations ($R^2c = 0.815$) and marginally lower coefficients of determination for cross-validations ($R^2cv = 0.589$) regarding tannin content. Similarly, the total starch content demonstrated high $R^2c = 0.983$ with slightly lower $R^2cv = 0.910$, while amylose content displayed high $R^2c = 0.762$ and slightly lower $R^2cv = 0.697$. The standard error of calibration (SEC) and standard error of cross-validation (SECV) were minimal (SECV/mean = 3.8%, 1.2%, and 4.3% for tannin, total starch, and amylose, respectively).

Ethiopian sorghum accessions grain samples were analysed for moisture, ash, protein, Fe and Zn content using both NIRS and chemical analysis at Melkassa Agricultural Research Centre. A Inframatic 9500 NIR System Grain Analyzer was used to scan 200 g of cleaned whole grain for every environment. Calibration models were established using NIRS spectra from 168 Ethiopian sorghum accession samples for calibration purposes and 46 samples for external validation, as reported by Zerihun et al. (2023).

Moisture, ash and protein were expressed as percentages, while Fe and Zn were measured in mg/kg, with respective ranges of 3.33%, 3.85%, 7.40%, 86.35 mg/kg, and 58.66 mg/kg.

Predictions for moisture, ash, protein, Fe and Zn demonstrated high accuracy, with R^2c and R^2cv as follows: moisture (0.999, 0.99), ash (0.997, 0.996), protein (0.942, 0.890), Fe (0.849, 0.599) and Zn (0.953, 0.883), respectively. The SEC of prediction values for moisture, ash, protein, Fe and Zn were 0.210, 0.197, 0.049, 0.058 and 0.03, respectively.

4.3.4 Broad-sense heritability

Broad sense H^2 was estimated from individual and combined ANOVA using the formula of Johnson et al. (1955). H^2 is classified as follows: 0-30% is considered low, 30-60% is considered moderate and >60% is considered high (Abate and Mekbib, 2015; Kumar et al., 2015).

$$\text{Heritability } (H^2) = \frac{\sigma^2 g}{\sigma^2 p} \times 100$$

Where: H^2 = heritability in a broad sense, $\sigma^2 p$ = Phenotypic variance and $\sigma^2 g$ = Genotypic variance.

4.4 Results

4.4.1 Genotypic variation, genotype by environment interaction and assessment of heritability

Starch and protein concentration showed highly significant $\sigma^2 g$ in each environment and for pooled analysis. Highly significant $\sigma^2 g$, $\sigma^2 en$ and $\sigma^2 gen$ effects were present for all traits (Table 4.4). In stress environments, starch and protein concentration had low mean $\sigma^2 g$ and H^2 values, however, protein concentrations had high mean values (Table 4.5).

Table 4.4 Genotype ($\sigma^2 g$), environment ($\sigma^2 en$) and genotype x environment ($\sigma^2 gen$) interaction variance, and standard errors (SE) for quality traits

Trait	$\sigma^2 g$	SE	$\sigma^2 gen$	SE	$\sigma^2 en$
Starch	0.30**	0.06	0.29**	0.11	2182**
Protein	0.45**	0.05	0.227**	0.06	1321**
Amylose	0.01**	0.00	0.01**	0.00	1967**
Amylopectin	0.01**	0.00	0.01**	0.00	1967**
Zn	4.38**	0.67	4.32*	1.79	250.70*
Fe	11984.76**	126.11	0.11**	0.32	748.21**
Ash	0.09**	0.01	0.05**	0.01	537.60**
Moisture	0.12**	0.03	0.16**	0.05	2933**

* $P \leq 0.05$; ** $P \leq 0.01$.

In the non-stress conditions, starch had the highest heritability (44.3%) at MK20, while protein had a high heritability (68.4%) at MK21. However, under stress environments, starch had the lowest heritability (23.2%) at MS20, while MS21 had the lowest protein heritability (54.7%). In non-stress environments (irrigated and intermediate-Jimma 2021 (JM21), Zn and Fe concentrations were highly heritable. In the stress environment, Zn and Fe had low heritability values (MS20). Zn and Fe heritability values in the intermediate environment Jimma 2020 (JM20) were average. There was a significant correlation for values across environments for some of the measured traits (Figure 4.2). When environments are correlated (have similar landrace responses), selecting the best materials in one environment would also reflect the best materials in another (Figure 4.2 and Table 4.6). In this case, MET analysis can help identify genotype broad and specific adaptations across a wide range of genetic materials. In MET analysis, using both the LMM and including random effects is essential. LMMs enable the integration of fixed effects (treatment effects or covariates) and random effects (environmental impacts or experimental errors) into the model, offering flexibility for MET data analysis.

Table 4.5 Mean, heritability and variance components for each trait in each environment

Trait	JM20				JM21				MK20			
	Mean	**σ^2g**	**σ^2e**	**H^2**	**Mean**	**σ^2g**	**σ^2e**	**H^2**	**Mean**	**σ^2g**	**σ^2e**	**H^2**
Starch	70.80	0.10	1.80	37.80	69.50	0.40	0.70	39.10	67.50	0.60	0.80	44.30
Protein	8.10	0.50	1.40	38.70	9.00	0.70	0.70	60.50	10.80	0.40	1.00	66.20
Amylose	19.90	0.10	0.00	44.40	19.70	0.10	0.00	83.70	19.50	0.00	0.00	85.60
Amylopectin	80.10	0.10	0.00	44.40	80.30	0.10	0.00	83.70	80.50	0.00	0.00	85.60
Zn	38.50	126.00	49.30	51.40	25.50	25.60	20.10	67.90	24.50	25.30	20.50	73.30
Fe	44.20	186.00	64.30	51.00	28.80	112.00	0.80	76.10	30.20	252.00	112.00	83.70
Ash	2.00	0.30	0.10	44.60	1.80	0.10	0.00	74.80	1.50	0.10	0.10	63.80
Moisture	14.80	0.98	0.95	36.30	14.90	0.01	1.62	36.30	17.10	0.22	0.48	38.20
	MK21				**MS20**				**MS21**			
Trait	**Mean**	**σ^2g**	**σ^2e**	**H^2**	**Mean**	**σ^2g**	**σ^2e**	**H^2**	**Mean**	**σ^2g**	**σ^2e**	**H^2**
Starch	71.60	0.50	2.10	38.40	66.80	0.70	1.40	23.20	66.50	0.40	3.20	25.10
Protein	7.60	1.10	0.60	68.40	11.60	0.80	0.50	57.10	10.30	1.10	0.80	54.70
Amylose	19.80	0.00	0.00	87.30	19.40	0.00	0.00	**76.00**	19.40	0.10	0.00	76.50
Amylopectin	80.20	0.00	0.00	87.30	80.70	0.00	0.00	**76.00**	80.60	0.10	0.00	76.50
Zn	24.70	18.10	17.00	73.40	27.20	79.30	17.80	24.40	27.90	31.60	10.40	58.80
Fe	27.10	175.00	54.00	82.00	29.80	229.00	19.60	60.40	33.60	192.00	52.00	81.30
Ash	1.80	0.10	0.10	75.20	1.50	0.20	0.00	32.30	2.20	0.20	0.00	67.30
Moisture	13.80	0.35	0.00	37.70	17.00	0.92	0.26	31.20	15.50	0.28	1.26	14.00

σ^2g = Genotypic variance; σ^2en = Environment variance; σ^2gen = Genotype x environment interaction variance; H^2 = heritability; JM20 and JM2 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021.

Table 4.6 The genetic correlation matrix of the six environments for starch (below diagonal) and protein (above diagonal) content

Environments	JM20	JM21	MK20	MK21	MS20	MS21
JM20		0.62**	0.73**	0.62**	0.65*	0.5**
JM21	0.45*		0.85**	0.71**	0.75**	0.58*
MK20	0.69**	0.31		0.84**	0.89**	0.68**
MK21	1.00**	0.45*	0.69**		0.75**	0.58**
MS20	0.65**	0.29	0.45*	0.65**		0.61**
MS21	0.74**	0.33	0.51**	0.74**	0.48*	

* $P \leq 0.05$; ** $P \leq 0.01$; JM20 and JM2 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021.

High starch contents were found in commercial cultivar G4 (Bonsa) and landraces G246, G175, G84, G63 and G3, which were ranked in the same order in MK21 and JM20. Test environment JM21 was non-significantly correlated with other environments for starch content as shown in Figure 4.2 and Table 4.6. The range and mean values of protein content were higher in stress environments than in the non-stress and intermediate environments. Landrace G358 (ETSL101310), had a high protein content in all environments. The JM20 environment had higher Zn and Fe ranges (22.3 to 95.7 ppm and 19.0 to 77.6 ppm, respectively) and mean values (38.5 and 44.2 ppm, respectively) than the other test environments. The top ten and bottom five performing landraces, as well as commercial check cultivars G1 (Adukar) and G4 (Bonsa), with their mean values for grain quality traits, are listed in Table 4.8. Sorghum landraces G264, G278, G142, G175 and G148 had the highest mean values for amylose, Zn and ash content. The eight grain quality traits were investigated and evaluated using METs with spatial analysis, and the combined mean of the BLUP was determined (Appendix 3).

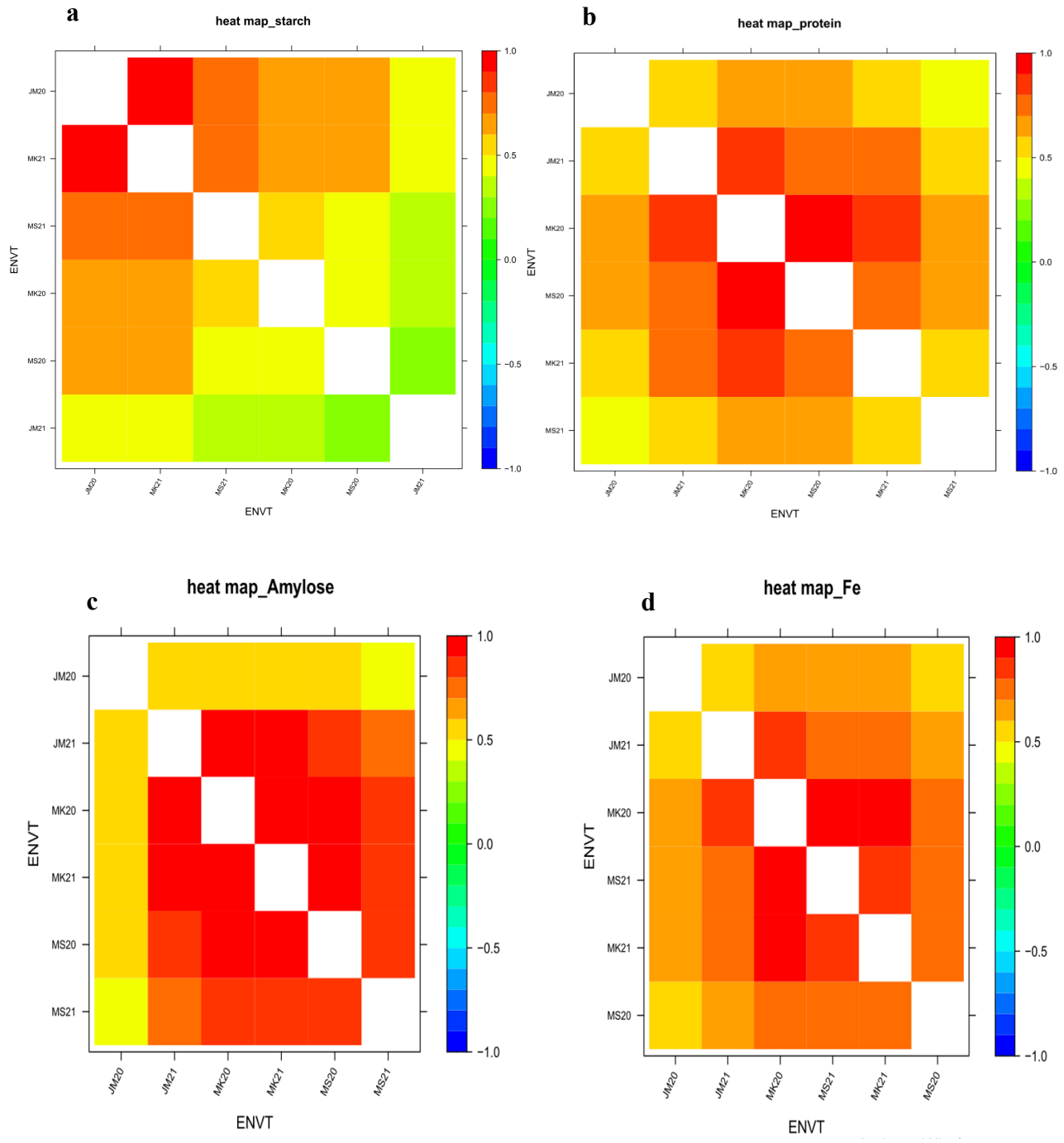


Figure 4.2 The heat map of the restricted maximum likelihood (REML) estimates of (a) starch, (b) protein (c) amylose and (d) Fe content of the environments' genetic correlation matrix. The key on the right side of each map depicts the correlation colour scale.

ENVT = Environment; JM20 and JM 21 = Jimms 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021; 1.0 indicates a good positive correlation; - 1.0 negative correlation and 0.0 no correlation.

4.4.3 Influence of drought on sorghum grain quality traits

Protein and starch concentrations

Melkassa and Miesso had higher protein concentrations in 2020 than in 2021, but lower starch concentrations. However, at Jimma, protein concentration was higher in 2021 than in 2020, while starch concentration was higher in 2020 (Table 4.9), confirming the expected inverse relationships between protein and starch concentrations. Protein concentrations at Miesso were higher in both years than in other environments, but starch concentrations were lower. MK21 produced grain with higher starch (71.56%) and lower protein (7.68%) concentrations than other environments (Table 4.10). Starch content was 5% higher in wet intermediate environments than in stress environments. Protein content was 21% higher in stress environments than in wet intermediate environments. Figure 4.3 and Table 4.10 show that the sorghum landraces had higher protein concentration trends in stress environments (MS20 and MS21) than in wet intermediate environments (JM20 and JM21), but sorghum landraces had higher starch concentration trends in wet intermediate environments (JM20 and JM21) than in stress environments (MS20 and MS21).

Stress environments had higher mean protein concentrations than irrigated and wet intermediate environments, while wet intermediate environments had higher mean starch concentrations compared to stress environments (Table 4.10 and Figure 4.3). In general, the mean protein content was higher in 2020 than in 2021 across all environments, but the mean starch content was higher in 2021. Miesso had the highest ash and moisture content in 2021, while Jimma had the highest ash content in 2020, but Melkassa and Miesso had higher moisture content in 2020.

Table 4.7 Range and mean values, and six top-ranking genotypes for each trait in each environment

Trait	Envt	Range	Mean	1	2	3	4	5	6
Starch (%)	JM20	69.90 - 71.30	70.80	G4	G246	G175	G84	G63	G3
	JM21	67.90 - 70.40	69.50	G265	G246	G257	G286	G261	G115
	MS20	64.60 - 67.90	66.80	G162	G84	G3	G175	G178	G4
	MS21	65.20 - 67.20	66.50	G175	G4	G246	G194	G63	G84
	MK20	64.94 - 68.94	67.50	G4	G175	G41	G246	G40	G87
	MK21	69.62 - 72.71	71.60	G4	G246	G175	G84	G63	G3
Protein (%)	JM20	7.00 - 9.20	8.10	G358	G161	G171	G218	G221	G293
	JM21	7.10 - 10.60	9.00	G358	G221	G293	G143	G161	G171
	MS20	9.40 - 13.20	11.60	G358	G161	G218	G225	G171	G184
	MS21	7.30 - 12.50	10.30	G358	G88	G184	G231	G360	G151
	MK20	9.36 - 11.97	10.80	G358	G161	G171	G218	G221	G293
	MK21	5.36 - 9.97	7.70	G171	G161	G218	G221	G293	G143
Amylose (%)	JM20	19.40 - 21.10	19.90	G278	G264	G163	G142	G230	G175
	JM21	19.00 - 20.90	19.70	G264	G278	G142	G175	G148	G122
	MS20	18.80 - 20.40	19.40	G264	G142	G278	G175	G148	G122
	MS21	18.70 - 20.40	19.40	G264	G142	G175	G278	G148	G122
	MK20	18.89 - 20.59	19.50	G264	G142	G278	G175	G148	G319
	MK21	19.19 - 20.98	19.80	G264	G142	G278	G175	G148	G319
Amylopectin	JM20	78.90 - 80.60	80.10	G184	G20	G48	G341	G356	G46
(%)	JM21	79.10 - 81.00	80.30	G341	G46	G48	G2	G1	G12
	MS20	79.60 - 81.20	80.70	G48	G2	G341	G46	G1	G12
	MS21	79.60 - 81.30	80.60	G294	G48	G12	G2	G341	G46
	MK20	79.41 - 81.11	80.50	G48	G2	G341	G46	G1	G12
	MK21	79.02 - 80.81	80.20	G48	G2	G341	G46	G1	G12
Zn (ppm)	JM20	22.30 - 95.70	38.50	G264	G142	G278	G175	G148	G230
	JM21	20.80 - 60.60	25.50	G264	G278	G142	G175	G148	G1
	MS20	19.70 - 63.60	27.20	G135	G245	G190	G119	G149	G207
	MS21	19.30 - 50.8	27.90	G264	G230	G116	G142	G355	G175
	MK20	19.46 - 59.53	24.80	G264	G278	G142	G175	G148	G137
	MK21	20-24 - 53.45	24.70	G264	G142	G278	G175	G148	G137
Fe (ppm)	JM20	19.00 - 77.60	44.20	G142	G137	G322	G194	G258	G276
	JM21	10.50 - 71.00	28.60	G137	G142	G278	G166	G180	G248
	MS20	8.00 - 78.0	29.80	G142	G119	G137	G87	G245	G166
	MS21	12.00 - 87.30	33.60	G142	G137	G23	G166	G248	G260
	MK20	3.75 - 93.13	30.80	G142	G137	G166	G248	G260	G172
	MK21	7.15 - 81.10	27.50	G142	G137	G279	G260	G248	G28
Ash (%)	JM20	0.90 - 3.80	2.00	G148	G142	G278	G264	G1	G163
	JM21	1.20 - 3.30	1.80	G1	G264	G278	G142	G148	G46
	MS20	0.80 - 3.00	1.50	G135	G245	G119	G149	G190	G87
	MS21	1.50 - 3.30	2.20	G202	G142	G230	G264	G148	G1
	MK20	0.79 - 2.80	1.51	G264	G278	G142	G148	G1	G137
	MK21	1.13 - 3.19	1.80	G142	G148	G264	G1	G278	G137
Moisture (%)	JM20	11.40 - 16.20	14.80	G184	G34	G358	G49	G208	G27
	JM21	14.80 - 15.20	14.90	G278	G163	G230	G245	G135	G207
	MS20	13.20 - 18.50	17.00	G184	G225	G125	G318	G314	G73
	MS21	15.10 - 16.80	15.50	G346	G164	G30	G182	G169	G9
	MK20	15.69 - 17.85	17.10	G34	G358	G49	G184	G27	G208
	MK21	13.21 - 16.54	13.80	G90	G355	G212	G290	G173	G106

Envt = Environment; JM20 and JM2 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021.

Table 4.8 The mean (BLUP) performance of the top ten and bottom five sorghum genotypes across the six environments for each of the grain quality traits

Starch		Protein		Amylose		Amylopectin		Zn		Fe		Ash		Moisture	
Gen	**Means**	**Gen**	**Means**	**Gen**	**Means**	**Gen**	**Means**	**Gen**	**Means**	**Gen**	**Means**	**Gen**	**Means**	**Gen**	**Means**
G4	69.70	G358	11.10	G264	20.70	G48	81.00	G264	58.00	G142	80.80	G264	3.00	G355	16.20
G246	69.60	G161	11.00	G278	20.50	G341	80.90	G142	50.00	G137	79.60	G148	3.00	G184	16.10
G84	69.60	G171	10.90	G142	20.50	G46	80.90	G278	48.20	G166	63.10	G142	2.90	G290	16.10
G175	69.50	G218	10.90	G175	20.30	G1	80.90	G175	46.50	G248	62.90	G1	2.90	G30	16.00
G63	69.50	G88	10.90	G148	20.00	G294	80.90	G148	44.30	G260	61.40	G278	2.70	G225	16.00
G194	69.40	G221	10.80	G319	20.00	G12	80.90	G230	39.00	G172	60.00	G137	2.60	G208	16.00
G123	69.40	G184	10.80	G122	20.00	G2	80.80	G137	38.10	G55	59.70	G175	2.50	G34	16.00
G3	69.40	G293	10.80	G230	20.00	G184	80.80	G355	37.50	G279	59.50	G166	2.40	G90	16.00
G162	69.40	G143	10.70	G163	20.00	G25	80.80	G259	37.50	G23	57.80	G23	2.40	G318	16.00
G41	69.30	G214	10.60	G137	19.90	G356	80.80	G206	37.20	G87	57.80	G48	2.40	G314	16.00
G362	67.90	G63	8.10	G294	19.10	G148	80.00	G186	23.50	G132	15.20	G78	1.30	G135	14.60
G90	67.90	G265	8.00	G1	19.10	G175	79.70	G4	23.30	G121	14.90	G301	1.30	G163	14.50
G264	67.80	G142	8.00	G46	19.10	G142	79.50	G111	23.00	G18	13.10	G244	1.20	G207	14.40
G212	67.60	G94	8.00	G341	19.10	G278	79.50	G289	22.90	G272	13.00	G139	1.20	G278	14.40
G355	67.40	G202	7.70	G48	19.00	G264	79.30	G296	21.90	G6	12.10	G96	1.20	G245	14.30

For the corresponding traits, the first ten landraces in each genotype (Gen) column are the top ten, while the last five are the bottom five.

Table 4.9 Influence of drought on sorghum landrace grain quality concentrations

Environment	Starch	Protein	Amylose	Amylopectin	Zn	Fe	Ash	Moisture
2020								
Jimma	70.77	8.21	19.94	80.06	37.65	44.21	2.00	14.76
Melkassa	67.46	10.75	19.49	80.51	24.83	30.22	1.51	17.09
Miesso	66.80	11.47	19.35	80.65	27.23	29.83	1.49	16.99
2021								
Jimma	69.45	8.99	19.65	80.35	25.57	28.75	1.84	14.89
Melkassa	71.56	7.68	19.82	80.18	24.74	27.10	1.80	13.76
Miesso	66.45	10.22	19.36	80.64	28.32	33.57	2.19	15.53
LSD (0.05)	0.03	0.04	0.01	0.01	0.50	0.58	0.03	0.05

Jimma = wet intermediate; Melkassa = dry lowland (irrigated); Miesso = dry lowland (stress).

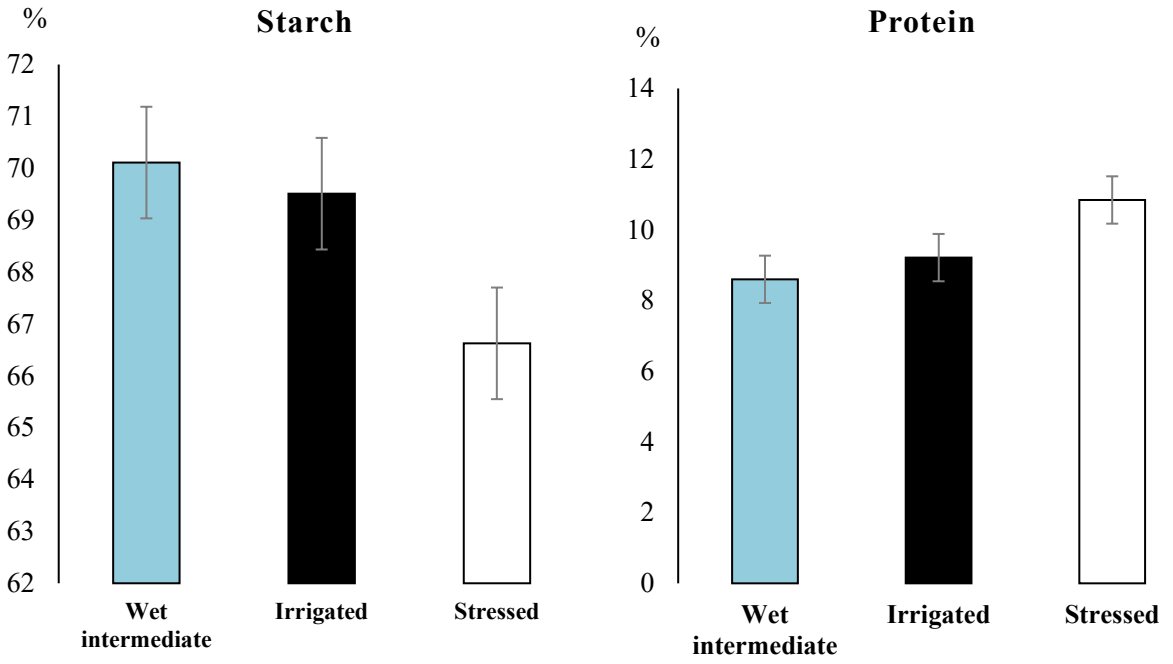


Figure 4.3 Sorghum starch and protein mean concentrations in different locations.

Jimma = wet intermediate, Melkassa = dry lowland (irrigated), Miesso = dry lowland (stress); lowland (> 1 600 m above sea level), intermediate elevation (1 600-1 900 masl), wet (high rainfall and long growing period), dry (low rainfall and short growing period).

Table 4.10 Environmental comparisons for sorghum grain quality traits

Contrast comparisons	Starch	Protein	Amylose	Amylopectin	Zn	Fe	Ash	Moisture
2020	68.34	10.14	19.59	80.41	29.90	34.75	1.67	16.28
2021	69.15	8.97	19.61	80.39	26.21	29.81	1.94	14.73
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Wet intermediate	70.11	8.60	19.80	80.20	61.31	36.48	1.92	14.83
Irrigated	69.51	9.22	19.66	80.34	24.78	28.66	1.65	15.43
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Irrigated	69.51	9.22	19.66	80.34	24.78	28.66	1.65	15.43
Stress	66.63	10.84	19.36	80.64	27.77	31.70	1.84	16.26
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Wet intermediate	70.11	8.60	19.80	80.20	61.31	36.48	1.92	14.83
Stress	66.63	10.84	19.36	80.64	27.77	31.70	1.84	16.26
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	NS	<0.001	<0.001

Jimma = wet intermediate; Melkassa = dry lowland (irrigated); Miesso = dry lowland (stress); NS = not significant.

Zn and Fe concentrations

Across environments, Zn and Fe concentrations (37.65 and 44.21 ppm, respectively) were the highest at JM20 followed by MS21 (28.32 and 33.57 ppm) (Table 4.9) with the lowest concentrations at Melkassa (irrigated) in both years. Jimma (wet intermediate) had higher soil total organic carbon, organic matter, nitrogen and available phosphorus contents than other environments. However, Melkassa had lower total nitrogen and available phosphorus contents than other environments (Table 4.3). Therefore, the high levels of total nitrogen and available phosphorus in the soil helped in promoting Zn and Fe uptake in plants. The wet intermediate environment had higher mean Fe and Zn concentrations than irrigated and stress environments, while irrigated environments had lower mean Fe and Zn concentrations (Table 4.10 and Figure 4.4). In general, Zn and Fe concentrations were higher in 2020 than in 2021. The Zn content in wet intermediate environments was 60% higher than in irrigated environments, while the Fe content in wet intermediate environments was 21% higher than in irrigated environments. The Zn content in wet intermediate environments was 55% higher than in stress environments, while the Fe content in wet intermediate environments was 13% higher than in stress environments.

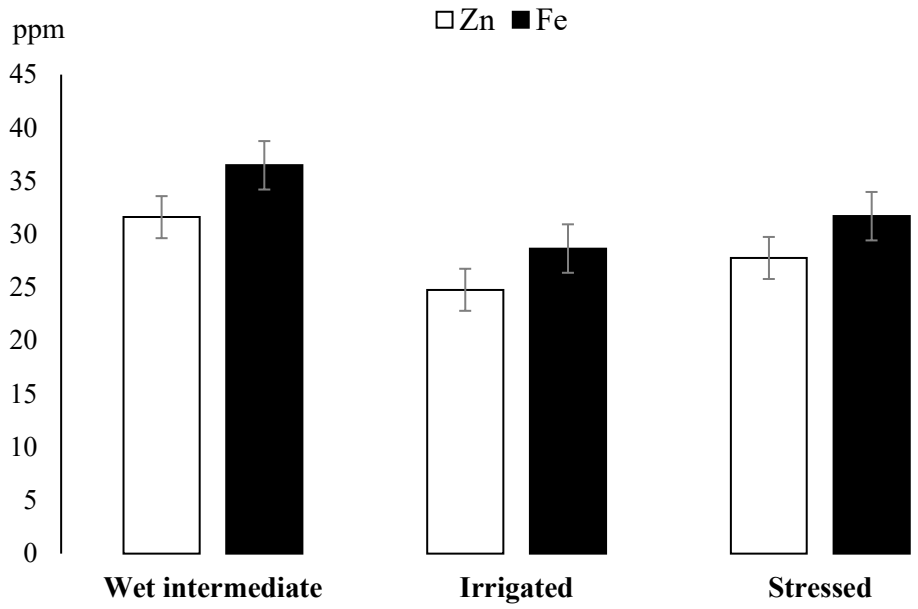


Figure 4.4 Sorghum Fe and Zn mean concentrations in different locations across seasons.

Jimma = wet intermediate; Melkassa = dry lowland (irrigated); Miesso = dry lowland (stress); lowland (> 1 600 m above sea level); intermediate elevation (1 600-1 900 masl); wet (high rainfall and long growing period); dry (low rainfall and short growing period.

4.4.4 Genotype main effects with genotype by environment interaction biplot analysis

"Which-won-where" polygon of GGE biplot

The GGE biplot polygon view identified the highest-performing genotypes by showing interaction patterns among genotypes and environments (Figures 4.5 and 4.6). A polygon was drawn in this GGE biplot by connecting the vertex genotypes, which were far from the origin, with straight lines, and thus all the other genotypes were enclosed within the polygon. Genotypes located at the polygon's vertices showed the best performance or worst in one or more environments. The starch vertex genotypes were G265, G84, G246, G4, G40, G110, G36, G264 and G355. The protein vertices genotypes were G358, G161, G232, G346, G202, G142, G94, G207 and G98. G142, G137, G194, G311, G272, G6, G18, G325, G190 and G119 were selected for the Fe vertex. The Zn vertices were G264, G249, G296, G285, G245 and G135 (Figure 4.5). As a result, these four groups of genotypes were the most responsive to environmental interactions for starch, protein, Fe and Zn, in that order.

In the "which-won-where" GGE biplot, lines from the origin divide the biplot into different sectors and create different mega environments (MGEs). In this study, all traits had two MGEs formed, except for Fe that only had one. For starch, environments JM20, MK20, MK21, MS20 and MS21 formed a joint MGE, whereas environment JM21 was a separate MGE for this trait. For protein, environments JM20, MK20, MK21, MS20 and JMS21 formed an MGE together, whereas environment MS21 was a separate MGE for this trait (Figure 4.5). For Zn, environments JM20, MK20, MK21, JM21 and MS21 formed a joint MGE, whereas environment MS20 formed a separate MGE for this trait, whereas for Fe, all environments formed a single MGE (Figure 4.6).

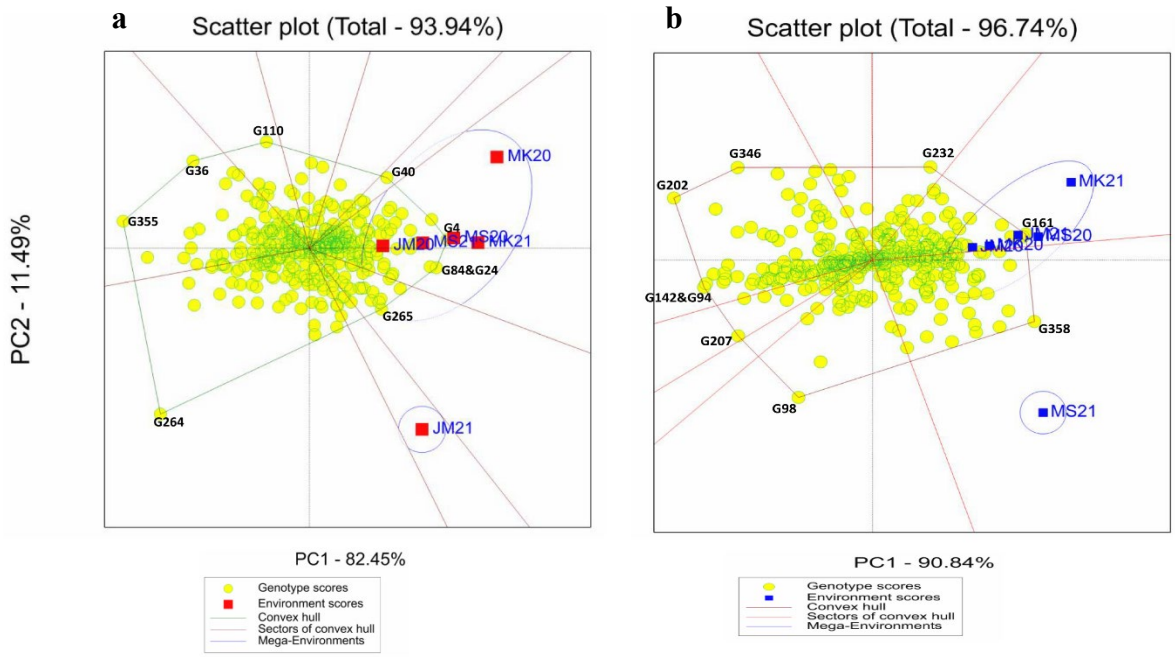


Figure 4.5 "Which-won-where" polygon view of genotype main effects with genotype by environment interaction scatter biplot of the 365 sorghum landraces for (a) starch and (b) protein, displaying landraces with the best performance in each environment and the mega environments. Vertex landraces on convex hulls (polygon) are the greatest for each mega environment for the correlating parameters.

JM20 and JM21 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021, respectively.

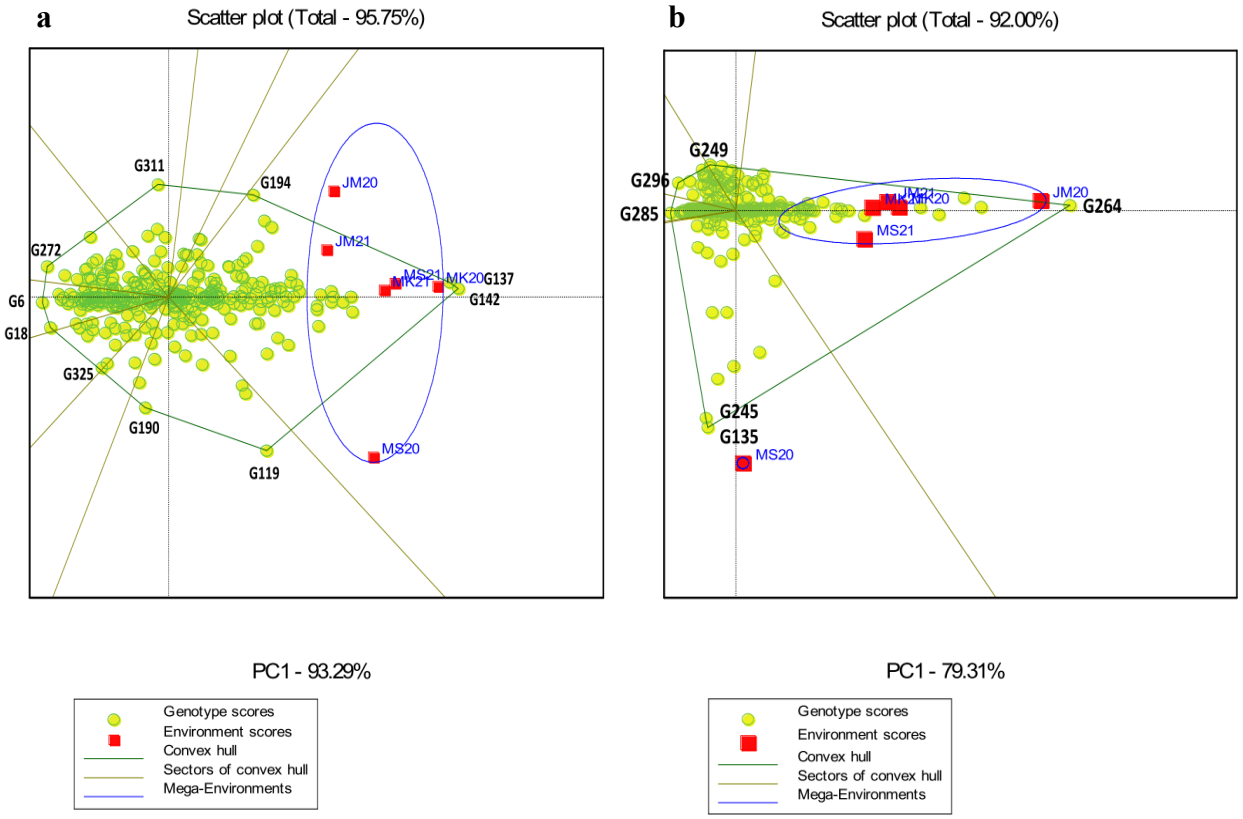


Figure 4.6 "Which-won-where" polygon view of genotype main effects with genotype by environment interaction scatter biplot of the 365 sorghum landraces for (a) Fe and (b) Zn, displaying landraces with the best performance in each environment and the mega environments. Vertex landraces on convex hulls (polygon) are the greatest for each mega environment for the correlating parameters.

JM20 and JM21 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021, respectively.

Genotypes G4 (Bonsa), G246 (ETSL101256) and G84 (ETSL100406) performed well in control and stress conditions for starch concentrations. Whereas genotype G161 (ETSL100834) performed well in control and stress environments for protein concentrations in 2020 and 2021, G358 (IS38353) performed well under stress in 2021 at Miesso. Genotypes G355, G36 and G264 had low starch contents, while genotypes G142, G94, G207, G202 and G98 had low protein contents in all environments (Figure 4.5). Genotype G264 (ETSL101325) performed well in the control and stress conditions for Zn concentrations, whereas G245 (ETSL101255) and G135 (ETSL100798) performed well in the MS20 stress condition. For Fe concentrations, genotypes G137 (ETSL100702) and G142 (ETSL100735) performed well in the control and stress conditions.

Genotypes G272, G6, G18 and G325 had low Fe contents, while the genotypes G285 and G296 had low Zn contents in all environments (Figure 4.6).

Genotype performance and stability

Genotypes should be assessed for average performance and stability over all environments within a single mega-environment. The single-arrowed line on a GGE biplot indicates the average environment coordinate (AEC) and indicates that the mean trait value is higher across environments. A genotype with a greater mean performance has an average environment axis (AEA) in the ranking biplot, represented by a single-arrowhead line passing through the origin. Thus, G246 and G84 had the highest mean starch contents, followed by G4, G175, and so on; G170, G35, G215, G106 and G193 had mean starch contents that were similar to the grand mean; and G355 and G212 had the lowest mean starch contents. In terms of protein, G358, G218, G161, G171 and G221 had the highest mean protein contents, followed by G43, G88, G293, G184 and so on; G170, G35, G215, G106, and G208 had mean protein contents similar to the grand mean; and G202, G63, G142 and G94 had the lowest mean protein contents (Figure 4.7). For Fe concentrations, G137 and G142 had the highest mean Fe contents, followed by G166, G248 and so on. G88 and G162 had mean Fe contents similar to the grand mean and G272, G6 and G18 had the lowest mean Fe contents. G264 had the highest mean Zn content, followed by G142, G278, G175 and G148, and so on; G119 and G90 had mean Zn contents similar to the grand mean (Figure 4.8).

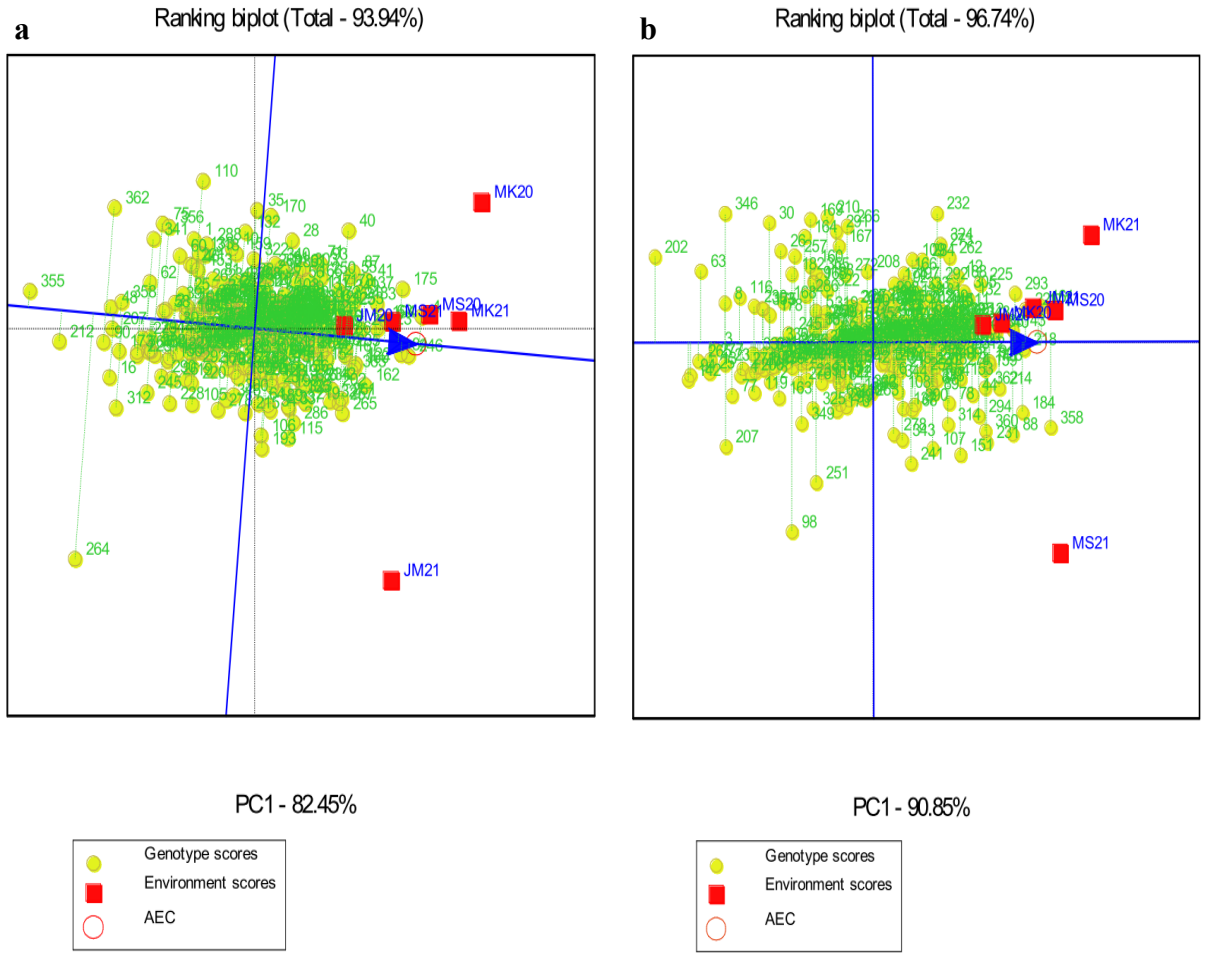


Figure 4.7 Genotype main effects with genotype by environment interaction biplot genotype focus scaling displaying the stability and mean priority of the 356 sorghum landraces for (a) starch and (b) protein.

JM20 and JM21 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021.

The stability of genotypes was assessed using the distance of the vector between genotype positions and the AEA in the ranking biplot. The best-performing and most stable genotypes are those that are far from the origin but on or near the AEA. As a result, G246 and G84 were the most stable genotypes with high mean starch contents and the shortest vector from AEA, whereas G264 and G110 were the least stable genotypes with the longest vector from AEA. The highest mean protein content and shortest AEA vectors were found in G218, while the lowest mean protein contents and longest AEA vectors were found in G210, G232, G98 and G251 (Figure 4.7).

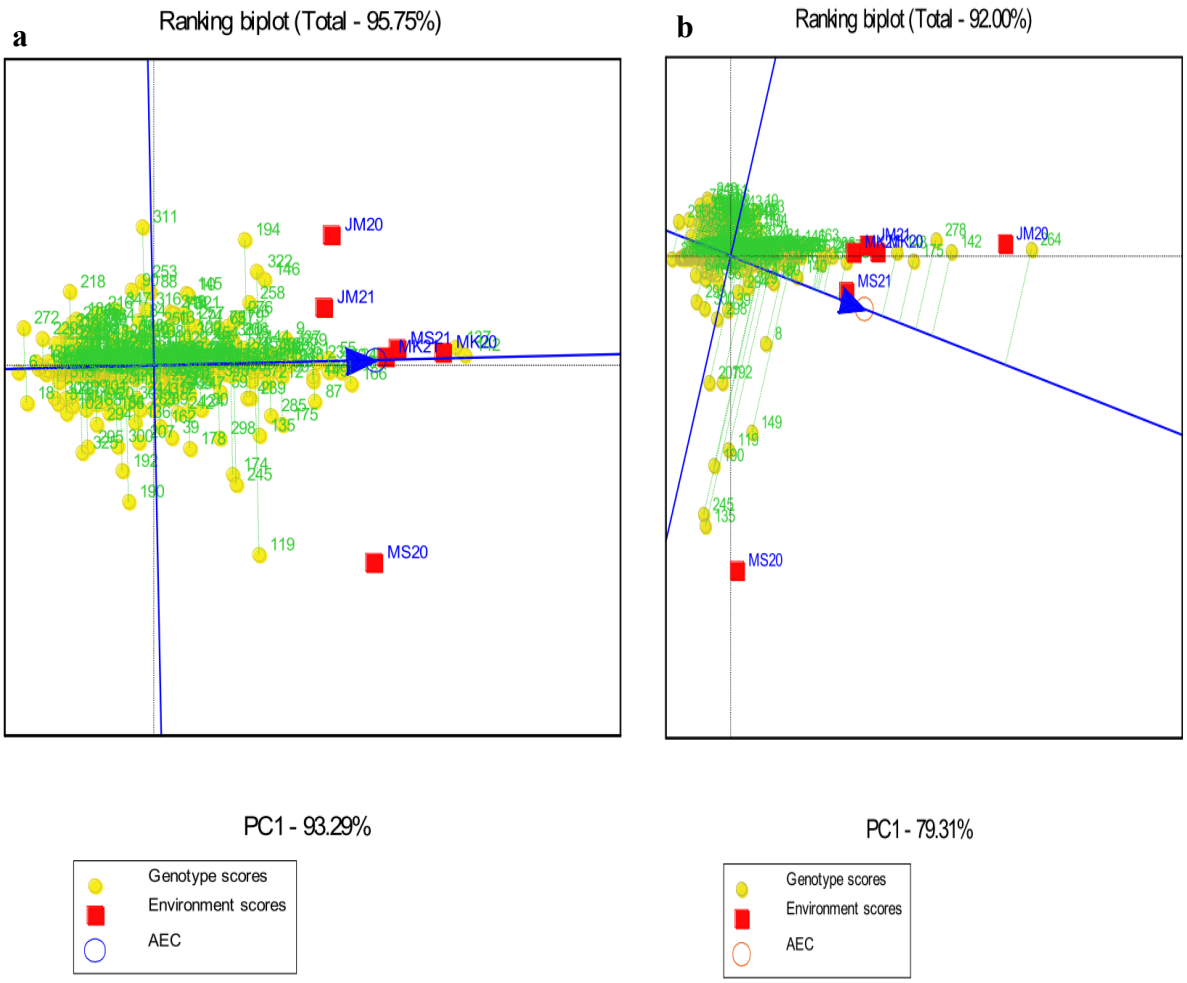


Figure 4.8 Genotype main effects with genotype by environment interaction focus scaling displaying the stability and mean priority of the 356 sorghum landraces for (a) Fe and (b) Zn.

JM20 and JM21 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021.

The highest mean Fe content and shortest AEA vectors were found in G137 and G142 and these genotypes were thus the most stable genotypes with high mean Fe contents and shorter AEA vectors, whereas G311, G194 and G119 were the least stable genotypes with the longest AEA vector. Meanwhile, G148 and G175 had the highest mean Zn contents and the shortest AEA vector, and G245 and G135 had the lowest mean Zn contents and the longest AEA vector (Figure 4.8).

Evaluation of the testing environment

The AEA, also known as the average-tester-axis, contains the AEC and ideal environment (shown by the circle at the end of the arrow in Figures 4.9 and 4.10) of all test environments and the AEA is the line that passes through the biplot origin and ideal environment. The GGE biplot displays AEA, AEC, ideal environment and concentric circles to evaluate the tested environments. The ideal test environment should be both discriminating and representative of the target environments and is on or near the centre of the concentric circles. A smaller angle with the AEA in a test environment is more representative of other test environments while test environments that are close to AEC are more representative of other environments.

In this study, for starch content, if the target environments are distributed into mega-environments, the discriminating but non-representative test environments MK20 and JM21 can be used to select genotypes with specific adaptation. The environment of MK21 was the closest to the ideal environment and thus the best at selecting cultivars adapted to the entire region. It should be noted that more years are required to confirm that a particular test location is ideal (Figure 4.9a). For protein content, MS20 was the most representative and ideal test environment, as well as the closest to the AEC, while MS21 was the most discriminating environment for protein content (Figure 4.9b). Environments with the highest Fe content, MS21, MK20 and MK21, were representative test environments with the smallest angles and were closest to AEC; on the other hand, MS20 and JM20 were discriminating environments for Fe concentrations (Figure 4.10a). The most representative environment for Zn was MS21, while the ideal and discriminating test environments were JM20 and MS20, respectively (Figure 4.10b).

4.4.5 Quality trait correlation

As the dataset was very large, even small correlations were significant (Table 4.11) but only correlations >0.3 are discussed here. Starch was positively and significantly correlated with amylose (r = 39) but negatively and significantly correlated with protein content (r = -0.39). Protein content was negatively and significantly correlated with starch (r = -0.39), Fe (r = -0.31) and Zn (r = -0.34). TSW correlated negatively and significantly with amylose (r = -0.37), Fe (r = -0.55) and positively with moisture (r = 0.33). Fe and Zn were significantly positively correlated (r = 0.54) as well as Fe and ash (r = 0.5) as were Zn and ash (r = 0.82). There was a negative and significant

correlation between grain moisture content and amylose (r = -0.50). On the other hand, grain moisture content had a positive and significant correlation with TSW (r = 0.33), which is expected.

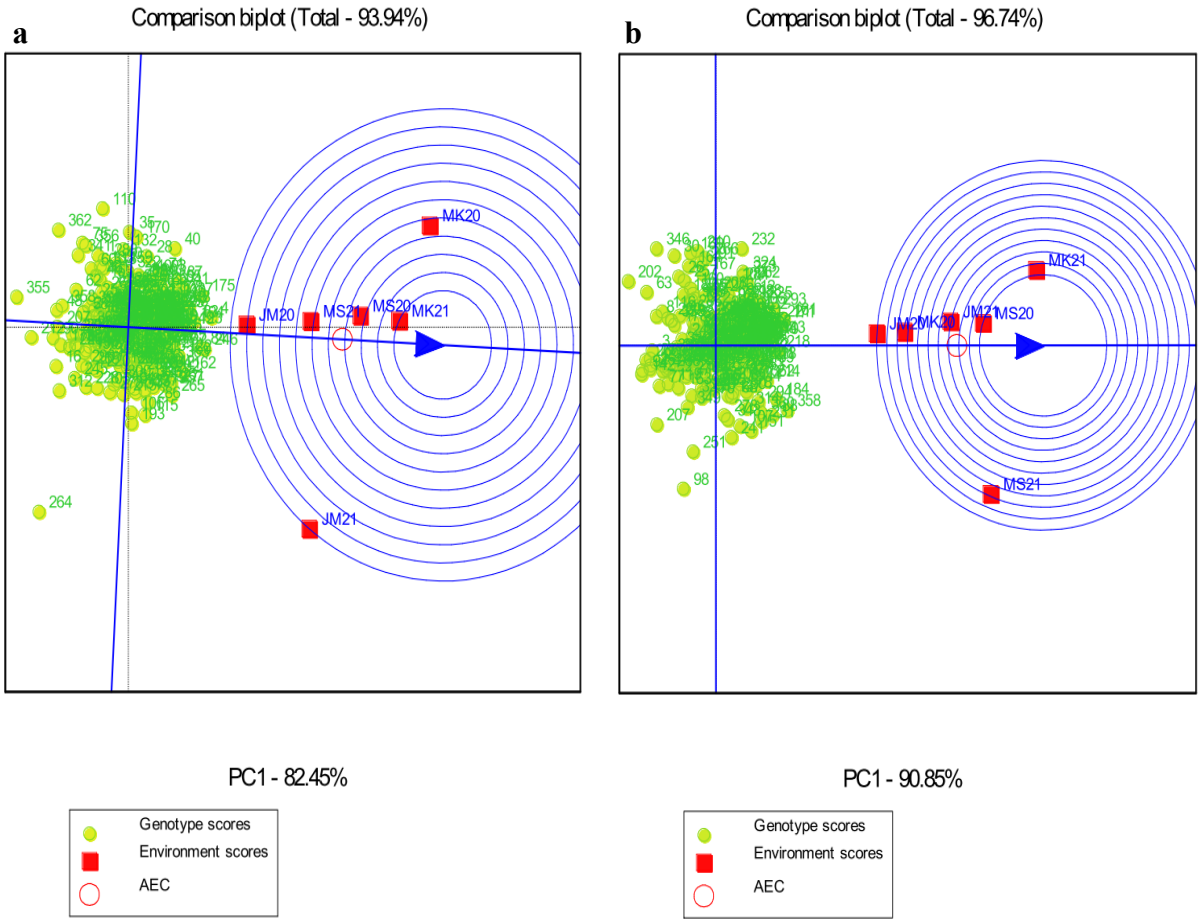


Figure 4.9 The discrimination and representativeness view of the genotype main effects with genotype by environment interaction biplot used to rank test environments relative to an ideal test environment (represented by the centre of the concentric circles) for (a) starch and (b) protein.

JM20 and JM21 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021.

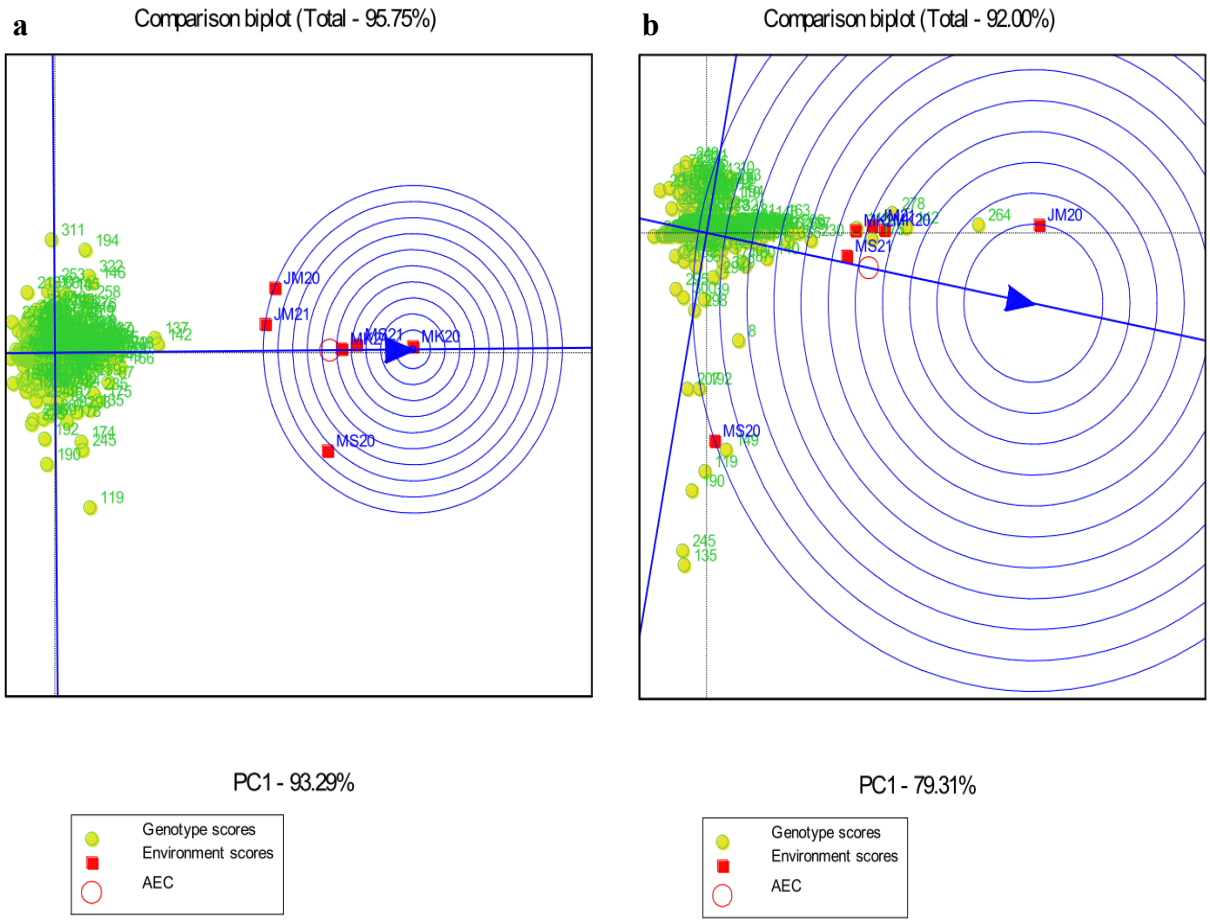


Figure 4.10 The discrimination and representativeness view of the genotype main effects with genotype by environment interaction biplot used to rank test environments relative to an ideal test environment (represented by the centre of the concentric circles) for (a) Fe and (b) Zn.

JM20 and JM21 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021.

Table 4.11 The grain quality correlation coefficients of eight traits from 358 sorghum landraces and checks

Trait	Protein	GY	TSW	Amylose	Fe	Zn	Ash	Moisture
Starch	-0.39**	0.17**	-0.13*	0.39**	0.13*	-0.22**	-0.20**	-0.25**
Protein		-0.11*	0.17**	-0.14**	-0.31**	-0.34**	-0.26**	0.15**
GY			0.18**	0.03	0.01	-0.10	-0.02	0.04
TSW				-0.37**	-0.55**	-0.20**	0.01	0.33**
Amylose					0.20**	0.16**	-0.14**	-0.50**
Fe						0.54**	0.50**	-0.27**
Zn							0.82**	-0.14**
Ash								-0.26**

* $P \leq 0.05$; ** $P \leq 0.01$; GY = grain yield; TSW = thousand seed weight.

4.5 Discussion

Sorghum has great potential due to its ability to adapt to harsh environments as well as evidence of its value as a healthy crop due to its polyphenol content and antioxidant properties. Regarding grain composition with advantageous bioactive compounds and climatic resilience, sorghum has a clear advantage over other cereals. It is critical to understand the grain composition to promote sorghum as a functional food grain. The variety of end-use products that can be made from sorghum grains necessitates different grain traits, which are appropriate for each end-use product (Aruna and Cheruku, 2019). Sorghum has a high phenotypic diversity for grain quality traits, suggesting that it could be bred for specific end-use products. Understanding the genetic variation that underpins the available phenotypic diversity enables breeders to create genotypes with desired grain traits for specific end-uses (Boyles et al., 2017).

The performance of promising genotypes must be repeatable and stable across years to identify and recommend superior genotypes. To achieve this, MET analysis of unbalanced data with factor analysis coupled with mixed models to accurately identify the relative stability and GEI of varieties is crucial. In MET analysis, both Linear mixed model (LMM) and integrating random effects are critical. The LMM provides a flexible modelling framework for analysing MET data, while fitting terms as random effects helps account for the hierarchical structure and variability inherent in MET experiments. LMMs allow for the incorporation of both fixed effects (treatment effects or

covariates) and random effects (environmental effects or experimental errors) into the model, providing a flexible framework for analysing MET data (Smith et al., 2001). The use of an LMM helps account for the correlation among observations within the same environment and across different environments, thereby providing more accurate estimates of treatment effects and their variability (Smith and Cullis, 2018).

In this study, starch and protein concentrations had low mean genetic variance and heritability values in stress environments. If the heritability of the desired traits is high, variability can be taken advantage of through selection. According to a study done on a natural collection of landraces, the BLUP methodology is an effective way to predict additive genetic values and estimate the components of variance (Kruuk, 2004; Smith et al., 2015). Complex traits are impacted by interactions between genetic and environmental factors, which lowers their heritability value. Additionally, GEI leads to various genetic association patterns in various environments (Phuke et al., 2017). Sorghum is the primary source of energy and micronutrients for many people in Africa. The present study was aimed at determining genetic variation, GEI and associations between measured traits, and to determine whether improved varieties with higher levels of Fe and Zn can be developed. In a previous study, efforts were made to identify the variation in mineral content in sorghum grains across environments as well as to understand the variation in grain Fe and Zn concentrations in recombinant inbred line populations (Phuke et al., 2017).

Under drought stress during the flowering stage, the total starch, amylose and amylopectin accumulation all decreased at the mid-late stage of grain filling. Drought stress during the flowering stage decreased starch synthesis enzyme activities, reducing starch accumulation in grains (Sehgal et al., 2018).

Lower accumulation of total starch, amylose and amylopectin was the result of the sorghum being subjected to drought stress during the flowering stage, which is associated with less efficient enzyme activity on sugar nucleotide precursors. Sorghum grain enzymatic activity for SSS, GBSS, SBE and DBE was decreased to varying degrees, preventing starch synthesis during grain filling (Bing et al., 2014). Drought reduces grain size by inhibiting cell proliferation in endosperm cells and the number of starch granules in grain (Nicolas et al., 1985). Drought stress increased total

protein content and kernel hardness but decreased protein digestibility (Kiliç and Yağbasanlar, 2010; Impa et al., 2019), our study indicated that mean protein levels under stress conditions were higher than the non-stress conditions. However, protein digestibility decreased under both stresses, with heat stress being more detrimental than water deficit stress (Impa et al., 2019). In the present study, the mean starch content in wet intermediate environments was 5% higher than in stress environments. However, the mean protein content in stress environments was 21% higher than in wet intermediate environments.

Both salinity and drought stress limit the availability, transport and partitioning of nutrients in the connection between mineral nutrients and plants, which in turn slows plant development. Salinity and drought, however, can have different effects on a plant's mineral concentrations (Hu and Schmidhalter, 2005). A lack of soil water disrupts the balance of minerals in the soil, which prevents plants from growing and developing and affects the final yield of the crop. A lack of water often inhibits the mass flow-dependent mineral nutrient uptake and translocation from the roots to the shoots, which affects all metabolic processes involved in plant physiology (Bhattacharya, 2021).

In this study, mineral concentrations in grains were typically higher under wet intermediate conditions when compared to stress and irrigated conditions. Variation between environments was found in the comparison of GGE biplots for grain Fe and Zn (Phuke et al., 2017). Grain Zn and Fe showed significant GEI, highlighting the need for further research to address the GEI and create sorghum landraces that are high in micronutrients. Fe and Zn concentrations in grain vary according to soil micronutrient concentrations. Another reason for the high GEI for Fe and Zn content could be their heritability (Gregorio, 2002; Stangoulis and Knez, 2022). The difference in the relative performance of genotypes in different environments is a strong indicator of GEI, as well as variation in environmental conditions such as temperature, rainfall and soil type. Starch, the main macronutrient, was reported to have a strong negative correlation with crude protein when grain macronutrients were measured on a percent dry matter basis (Boyles et al., 2016). Fe and Zn contents had a significant positive relationship over environments, indicating the possibility of concurrently selecting for both traits (Phuke et al., 2017), which was also the case in the current study.

The GGE “which-won-where” biplot was applied to distinguish top-performing genotypes by interpreting GEI, MGE clustering and specific adaptation (Yan and Tinker, 2006; Rakshit et al., 2012; Enyew et al., 2021). The unstable genotypes are those that are far away from the biplot origin (vertex genotypes) and perform best or worst in different environments. Vertex genotypes, which are more responsive to environmental change and are thought of as specifically adapted genotypes, are the ones that perform poorly or are the best in some or all the tested environments (Yan et al., 2007). Based on the "which-won-where" biplot, the testing environments were divided into two MGEs with various high-performing genotypes for starch, protein and Zn, but Fe only had one MGE. Environments with the highest starch content for genotypes G4, G84 and G246 in MGE1 were MK20 and MK21, MS20 and MS21, and JM20, while MGE2 had environment JM21 with G265 as the highest starch content genotype. In terms of protein, MGE1 included environments MK21, MK20, MS20, JM21 and JM20 with genotype G161 having the highest protein content, whereas MGE2 only had environment MS21 with the genotype G358 having the highest protein content.

In terms of Zn, MGE1 included environments MK21, MK20, JM21, MS21and JM20 with the genotype with the highest concentration of Zn, G264, whereas in MGE2 environment MS20 had genotypes with the highest concentration of Zn, G245 and G135. The highest Fe-containing genotypes, G137 and G142, were combined into a single MGE for all environments. This suggests that particular genotype adaptations to MGEs occurred and that the GE interaction was successfully exploited (Mohammadi et al., 2015). The most effective way to take advantage of the positive GE interaction is to cluster the target environments into expressive MGEs and choose various genotypes for each MGE (Yan and Tinker, 2005).

The GGE biplots are the best for graphical assessment of the testing environments because they show the heritability of traits in each test environment based on vector length and its genetic correlation with other test environments based on the angle between two test environments, which are two essential components for test environment assessment (Yan and Holland, 2010). Because it takes into account both the effect of genotype and the GE interaction, the model is also known as GGE (Yan, 2001).

The most stable and high-ranking genotypes can be found using the GGE ranking biplot through AEC (Yan, 2001). The highest starch content genotypes in the study were G246, G84, G4 and G175, while landraces with low starch contents like G264, G362, G355 and G212 experienced decreased stability due to the GEI effect. Genotypes G218, G43 and G161 scored highest in terms of protein content in this study. However, the GEI effect decreased the stability of landraces with high protein content, such as G358, G88 and G184. Although genotype G264 had the highest Zn content, the GEI effect made it less stable. The highest Fe content was found in G137 and G142. The GGE biplot statistical model is useful for identifying the top-ranking, stable genotypes across environments and the best genotypes for a given mega environment (Yan and Kang, 2002). The current study found that the GGE biplot could distinguish between environments. Because they were far from the biplot origin in both analyses, the environments had a comparable level of discriminative power. However, the contribution of the environments to genotype stability produced somewhat different results.

Starch has a strong negative correlation with crude protein (Boyles et al., 2016). The protein content of flour and the total starch content were significantly and inversely associated (Labuschagne et al., 2007). Fe and Zn contents had a significant positive relationship over environments, indicating the possibility of concurrently selecting for both traits (Phuke et al., 2017), which was also the case in the current study. Starch had significant negative correlations with protein, and significant positive correlations with amylose. Protein was negatively and significantly correlated with Fe, Zn and starch. So, an increase in protein content will cause a related decrease in Fe and Zn, and simultaneous selection for higher protein content and increased mineral content will not be possible. Fe and Zn were positively and significantly correlated, suggesting that selection for one of the two will lead to an increase in the other as well, which is important as it will make selection for increased Fe and Zn content easier.

4.6 Conclusions

Findings of this study demonstrated that these sorghum landraces are an excellent genetic resource with high variation in grain quality traits such as starch, protein and minerals that should be used to develop new high-quality cultivars with a variety of desirable traits. Stress environments had lower starch but higher protein content than wet intermediate environments. Wet intermediate

environments caused higher mineral concentrations in grains than irrigated and stressed environments. The genotypes G246, G84, G4 and G175 had the highest and most stable starch contents, while G218, G43 and G161 had the highest and most stable protein contents. Genotypes G137 and G142 had the highest and most stable Fe contents, while genotype G264 had the highest and most stable Zn contents. In this study, JM21 was the most discriminating test environment for starch content, while MS21 was the most discriminating for protein content.

4.7 References

Abate, M. and Mekbib, F. (2015). Assessment of genetic variability and character association I Ethiopian low altitude sesame (*Sesamum indicum* L.) genotypes. *Journal of Advanced Studies in Agricultural, Biological, and Environmental Sciences*, 5: 55-66.

Admas, S. and Tesfaye, K. (2017). Genotype-by-environment interaction and yield stability analysis in sorghum (*Sorghum bicolor* (L.) Moench) genotypes in North Shewa, Ethiopia. *Acta Universitatis Sapientiae. Agriculture and Environment,* 9: 82-94.

Aruna, C. and Cheruku, D. (2019). Genetic improvement of grain sorghum. In: Aruna, C., Visarada, K.B.R.S., Bhat, B.V., Tonapi, V.A. (eds). Breeding sorghum for diverse end uses. Woodhead Publishing, pp. 157-173.

Ashraf, M. (2014). Stress-induced changes in wheat grain composition and quality. *Critical Reviews in Food Science and Nutrition*, 54: 1576-1583.

Bhattacharya, A. (2021). Mineral Nutrition of Plants Under Soil Water Deficit Condition: A Review. In: Bhattacharya, A. (ed). Soil Water Deficit and Physiological Issues in Plants. Springer, pp. 287-391.

Bing, Y.I., Zhou, Y.F., Gao, M.Y., Zhang, Z., Yi, H.A.N., Yang, G.D. and Huang, R.D. (2014). Effect of drought stress during the flowering stage on starch accumulation and starch synthesis enzymes in sorghum grains. *Journal of Integrative Agriculture,* 13: 2399-2406.

Boyles, R.E., Cooper, E.A., Myers, M.T., Brenton, Z., Rauh, B.L., Morris, G.P. and Kresovich, S. (2016). Genome-wide association studies of grain yield components in diverse sorghum germplasm. *The Plant Genome,* 9: 1-17.

Boyles, R.E., Pfeiffer, B.K., Cooper, E.A., Rauh, B.L., Zielinski, K.J., Myers, M.T. and Kresovich, S. (2017). Genetic dissection of sorghum grain quality traits using diverse and segregating populations. *Theoretical and Applied Genetics*, 130: 697-716.

Butler, D.G., Cullis, B.R., Gilmour, A.R., Gogel, B.J. and Thompson, R. (2017). ASReml-R /reference manual version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK, pp. 30-49.

Enyew, M., Feyissa, T., Geleta, M., Tesfaye, K., Hammenhag, C. and Carlsson, A.S. (2021). Genotype by environment interaction, correlation, AMMI, GGE biplot, and cluster analysis for grain yield and other agronomic traits in sorghum (*Sorghum bicolor* L. Moench). *Plos One*, 16: e0258211.

Finlay, K.W. and Wilkinson, G.N. (1963). The analysis of adaptation in a plant-breeding program. *Australian Journal of Agricultural Research*, 14: 742-754.

Gemechu, F., Lemessa, K., Melka, T., Gadisa, B., Dekeba, S. and Zewdu, A. (2016). Effect of Climate Change on Agricultural Production and Community Response in Daro Lebu and Mieso District, West Hararghe Zone, Oromia Region National State, Ethiopia. *Journal of Natural Sciences Research*, 6: 30-36.

Gregorio, G.B. (2002). Progress in breeding for trace minerals in staple crops. *The Journal of Nutrition*, 132: 500S-502S.

Horn, L., Shimelis, H., Sarsu, F., Mwadzingeni, L. and Laing, M.D. (2018). Genotype-by-environment interaction for grain yield among novel cowpea (*Vigna unguiculata* L.) selections derived by gamma irradiation. *The Crop Journal*, 6: 306-313.

Hu, Y. and Schmidhalter, U. (2005). Drought and salinity: a comparison of their effects on mineral nutrition of plants. *Journal of Plant Nutrition and Soil Science*, 168: 541-549.

Impa, S.M., Perumal, R., Bean, S.R., Sunoj, V.J. and Jagadish, S.K. (2019). Water deficit and heat stress-induced alterations in grain physicochemical characteristics and micronutrient composition in field-grown grain sorghum. *Journal of Cereal Science*, 86: 124-131.

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

Kang, M.S. (2003). Genotype environment interaction progress and prospects. CAB International, Wallingford, UK, pp. 221-243.

Kaufman, R.C., Wilson, J.D., Bean, S.R., Galant, A.L., Perumal, R.R., Tesso, T. and Shi, Y.C. (2018). Influence of genotype × location interaction on grain sorghum grain chemistry and digestibility. *Agronomy Journal*, 110: 1681-1688.

Kiliç, H. and Yağbasanlar, T. (2010). The effect of drought stress on grain yield, yield components, and some quality traits of durum wheat (*Triticum turgidum* ssp. durum) cultivars. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 38: 164-170.

Kruuk, L.E. (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359: 873-890.

Kumar, N.M., Gangaprasd, S. and Sunil, M.G. (2015). Genetic variability and heritability studies in cowpea [*Vigna unguiculata* (L.) Walp]. *Research Journal of Agricultural Sciences*, 6: 763-766.

Labuschagne, M.T., Geleta, N. and Osthoff, G. (2007). The influence of environment on starch content and amylose to amylopectin ratio in wheat. *Starch-Stärke*, 59: 234-238.

MoA (Ministry of Agriculture) (1998). Agro-ecological zones of Ethiopia: natural resource management and regulatory department. Ethiopian Ministry of Agriculture, Addis Ababa, 1-28

Mohammadi, R., Haghparast, R., Amri, A. and Ceccarelli, S. (2009). Yield stability of rainfed durum wheat and GGE biplot analysis of multi-environment trials. *Crop and Pasture Science*, 61: 92-101.

Mohammadi, M., Noorinia, A.A., Khalilzadeh, G.R. and Hosseinpoor, T. (2015). Application of GGE biplot analysis to investigate GE interaction on barley grain yield. *Current Opinion in Agriculture*, 4: 25.

Nicolas, M.E., Gleadow, R.M. and Dalling, M.J. (1985). Effect of post-anthesis drought on cell division and starch accumulation in developing wheat grains. *Annals of Botany*, 55: 433-444.

Phuke, R.M., Anuradha, K., Radhika, K., Jabeen, F., Anuradha, G., Ramesh, T., and Kumar, A. A. (2017). Genetic variability, genotype × environment interaction, correlation, and GGE biplot analysis for grain iron and zinc concentration and other agronomic traits in RIL population of sorghum (*Sorghum bicolor* L. Moench). *Frontiers in Plant Scienc*e, 8: 712.

Piguet, F. (2003). Ethiopia: Hararghe Food Security is hampered by long-term drought conditions and economic constraints. UNDP Emergencies Unit for Ethiopia, 2-13 March 2003.

Rakshit, S., Ganapathy, K.N., Gomashe, S.S., Rathore, A., Ghorade, R.B., Kumar, M.N. and Patil, J.V. (2012). GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. *Euphytica*, 185: 465-479.

Rao, P.S., Reddy, P.S., Rathore, A., Reddy, B.V. and Panwar, S. (2011). Application of GGE biplot and AMMI model to evaluate sweet sorghum (*Sorghum bicolor*) hybrids for genotype× environment interaction and seasonal adaptation. *Indian Journal of Agricultural Sciences*, 81: 438-444.

Redlands, C.A. (2011). ArcGIS Desktop: Release 10.

Sehgal, A., Sita, K., Siddique, K.H., Kumar, R., Bhogireddy, S., Varshney, R.K. and Nayyar, H. (2018). Drought and/or heat-stress effects on seed filling in food crops: impacts on functional biochemistry, seed yields, and nutritional quality. *Frontiers in Plant Science*, 9: 1705.

Smith, A.B. and Cullis, B.R. (2018). Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. *Euphytica*, 214: 1-19.

Smith, A.B., Borg, L.M., Gogel, B.J. and Cullis, B.R. (2019). Estimation of factor analytic mixed models for the analysis of multi-treatment multi-environment trial data. *Journal of Agricultural, Biological, and Environmental Statistics*, 24: 573-588.

Smith, A.B., Ganesalingam, A., Kuchel, H. and Cullis, B.R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theoretical and Applied Genetics*, 128: 55-72.

Stangoulis, J. C. and Knez, M. (2022). Biofortification of major crop plants with iron and zinc-achievements and future directions. *Plant and Soil*, 474: 57-76.

Thitisaksakul, M., Jiménez, R.C., Arias, M.C. and Beckles, D.M. (2012). Effects of environmental factors on cereal starch biosynthesis and composition. *Journal of Cereal Science*, 56: 67-80.

Worede, F., Mamo, M., Assefa, S., Gebremariam, T. and Beze, Y. (2020). Yield stability and adaptability of lowland sorghum (*Sorghum bicolor* (L.) Moench) in moisture-deficit areas of Northeast Ethiopia. *Cogent Food and Agriculture*, 6: 1736865.

Yan, W. (2001). GGE biplot - A Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agronomy Journal*, 93: 1111-1118.

Yan, W. and Holland, J.B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica*, 171: 355-369.

Yan, W. and Kang, M.S. (2002). GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC press, pp. 63-83.

Yan, W. and Tinker, N.A. (2005). An integrated biplot analysis system for displaying, interpreting, and exploring genotype× environment interaction. *Crop Science,* 45: 1004-1016.

Yan, W. and Tinker, N.A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science*, 86: 623-645.

Yan, W., Kang, M.S., Ma, B., Woods, S. and Cornelius, P.L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science*, 47: 643-653.

Ye, X., Avendaño, S., Dekkers, J.C.M. and Lamont, S.J. (2006). Association of twelve immune-related genes with performance of three broiler lines in two different hygiene environments. *Poultry Science*, 85: 1555-1569.

Zerihun, M., Fox, G., Nega, A., Seyoum A., Minuye, M., Jordan, D., Taddese, T., Assefa, A. (2020). Near-Infrared Reflectance Spectroscopy (NIRS) for Tannin, Starch and Amylase Determination in Sorghum Breeding Programs. *International Journal of Food and Nutritional Science*, 7: 45-50.

Zerihun, M., Fox, G. and Tolessa, K. (2023). Predication of Quality Parameters of Sorghum Grain Using Near-Infrared Spectroscopy (NIRS) Analysis. *Proceedings of Food Science and Nutrition Completed Research Activities*, 109-120. file:///C:/Users/HP/Downloads/Final%20Proceeding%20Janu%202022%20final%20version%20v1%20(2).pdf

CHAPTER 5

GENOME-WIDE ASSOCIATION STUDY OF GRAIN QUALITY IN ETHIOPIAN SORGHUM LANDRACES

5.1 Abstract

The aim of this study was to use GWAS to identify loci linked to the expression of starch, amylose and protein content in 365 Ethiopian sorghum landrace accessions. Loci highly significantly associated with variation for these traits, were identified across the different sorghum landraces. Detected SNPs, linked to starch and protein content, varied across different environments. Similar SNPs identified to be associated with amylose content across different environments, indicated that genes/QTL associated with amylose expression may be less influenced by environmental conditions and may be more consistently expressed across different environments. Across environments, GWAS, using 209,572 SNPs and multiple models, identified five novel loci associated with starch, three with amylose and three with protein.

Candidate genes and proteins showing significant association with diverse traits were identified. The late embryogenesis abundant (LEA) and CCCH-type zinc finger proteins, which encode a stress tolerance response; ABC transporter and sugar carrier proteins, which transport sugar and various molecules; and auxin response factor 12 (ARF12) protein, a transcription factor known for starch synthesis and storage in developing seeds, were found closely located with the significant SNPs detected for starch. The WUSCHEL-related homeobox 6 (WOX6) protein, which acts as a transcription factor and is known to regulate stem cell maintenance and differentiation, an alpha-glucan phosphorylase, involved in amylose synthesis and a zinc finger CCHC domain protein encoding gene, known to regulate gene expression, all located within the proximity of significant SNPs identified for amylose. The long-chain-fatty-acid protein encoding gene, involved in the biosynthesis of fatty acids, TatD family of hydrolases protein, known for plant growth and development, Class III peroxidase 70 precursor and Class III peroxidase 69 precursor proteins, involved in various biological processes like plant defence against pathogens, were found co-localized with the SNPs detected for protein. Identification of loci associated with grain quality provides new insight into the genetic control of these traits, while sorghum landraces with desirable

grain composition can serve as gene sources during breeding. Overall, the findings suggested that loci of grain quality genes play an important role in the chemical and physical properties of sorghum grains.

5.2 Introduction

Sorghum is a cereal grain that is widely consumed in many parts of the world by both humans and livestock (Xiong et al., 2019; Hossain et al., 2022). Sorghum is known for its nutritional value, being rich in starch, protein and dietary fibre, as well as important minerals such as Fe, K and P. It is furthermore a good source of antioxidants and other beneficial compounds, such as polyphenols (Ragaee et al., 2006; Taylor and Emmambux, 2010). Starch is the primary storage carbohydrate in sorghum grains and is composed of two types of glucose polymers, amylose and amylopectin (Bahaji et al., 2014; Seung, 2020).

The protein content of sorghum varies depending on the cultivar, growing conditions and harvest time (Motlhaodi et al., 2018; Bazié et al., 2023). Sorghum protein is of high quality, with a well-balanced amino acid composition (Young and Pellett, 1994). Sorghum has low protein digestibility, which affects its nutritional value and marketability. Improving protein digestibility can help fight hunger and malnutrition and benefit livestock and sorghum farmers (Mofokeng et al., 2017). A study which investigated the effect of altering an alpha-kafirin gene family on sorghum digestibility and protein quality targeted the *k1C* genes with a clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) gene editing technique to develop variants with lower kafirin levels and increased protein quality and digestibility (Li et al., 2018). The edited plants showed significantly increased grain protein digestibility and lysine content. The study identified 42 SNPs that were associated with five forage quality-related traits, including crude protein, neutral detergent fibre, acid detergent fibre, hemicellulose and cellulose contents.

GWAS is a powerful tool for identifying genomic regions associated with complex traits in diverse populations (Saini et al., 2022) and GWAS have been used to identify loci associated with sorghum traits such as grain yield (Girma et al., 2019), drought tolerance (Maina et al., 2022) and disease resistance (Nida et al., 2019; 2021). To perform GWAS, many sorghum lines from different

landrace collections need to be genotyped using a high-density SNP array or whole-genome sequencing (Mundada et al., 2022).

The objective of this study was to identify genomic regions associated with grain quality, indicating a genetic basis for starch, amylose and protein content in Ethiopian sorghum landrace accessions.

5.3 Materials and methods

5.3.1 Plant material and design

Sorghum landraces and improved varieties used in this study, as well as the trial design followed, were as described in section 3.3.1. Planting was done across three sites (Jimma, Melkassa and Miesso) as described in section 4.3.1.

5.3.2 Sorghum grain quality analysis

Analysis of the nutritional quality of sorghum grains collected from six different environments (three locations over two seasons) was done as described in section 4.3.3.

5.3.3 Genotyping-by-sequencing

Results obtained by Girma et al. (2019) were used in this study. DNA extraction, GBS and SNP calling to obtain genetic information for 365 sorghum landraces were performed. The SNP data were extracted from raw data available in the Purdue University Research Repository and a total of 209,572 SNPs were obtained after quality control. The quality control measures involved filtering out individuals with less than 20% missing data, loci with less than 40% missing data and SNPs with a minor allele frequency (MAF) ≥ 0.05. TASSEL 5.0 software (Glaubitz et al., 2014), a commonly used software for genetic analysis, was used. Extracted SNPs were used for downstream analysis such as population structure analysis, association mapping and GWAS to identify SNP loci linked to protein quality.

5.3.4 Linkage disequilibrium analysis

A comprehensive genome-wide linkage disequilibrium (LD) investigation was conducted utilising trait analysis by association, evolution and linkage (TASSEL). This involved assessing the

pairwise squared allele-frequency correlations (r^2) between SNP markers within a sliding window of 50 SNPs. Subsequently, the r^2 values were graphed against physical distance to gauge the degree of LD between locus pairs. To estimate the LD decay rate, the genome-wide LD decay curve line was integrated into the scatterplot through a smoothing spline regression line, a method delineated by Hill and Weir (1988), within the R environment.

5.3.5 Genome wide association study

Grain quality data was generated using NIR and combined ANOVA was performed using the mixed linear model in R software to determine variation in grain quality (starch, amylose and protein) content across different environments. METs were analysed by fitting a linear mixed model using spatial analysis in ASReml (Butler et al., 2017; Smith et al., 2019) to investigate the effect of spatial variability on grain quality content. The methodology employed a combination of statistical and genomic analysis to identify genetic factors associated with sorghum grain quality collected from six different environments.

The BLUP data for sorghum grain quality traits were generated using 365 sorghum landraces, 209,572 high-quality SNPs and multiple models including the general linear model (GLM), mixed linear model (MLM), compressed mixed linear model (CMLM), fixed and random model circulating probability unification (FarmCPU) and Bayesian information and linkage disequilibrium iteratively nested keyway (BLINK) in BLUP were implemented in R using the GAPIT package version 3 (Wang and Zhang, 2021). Genomic regions associated with grain quality composition (starch, amylose and protein) were identified and gene annotation databases like Phytozome were used to find candidate genes within these regions.

5.4 Results

5.4.1 Multi environment trial analysis and genetic variation

Data obtained in Chapter 4 indicated that Ethiopian sorghum landrace accessions exhibited significant variability in its starch, amylose and protein contents across different environments. This data was used for GWAS in the current chapter.

5.4.2 Linkage disequilibrium

The average linkage disequilibrium (LD) analysis examines the non-random association of alleles at different loci on the same chromosome. The LD across the ten chromosomes shows a relatively consistent pattern, with r^2 values ranging from 0.10 to 0.14, and an overall mean of 0.12 (Table 5.1). LD was statistically significant for 44% of the marker pairs on average (269,457 pairs; mean $r^2 = 0.26$; $p \leq 0.01$) (Table 5.1). Approximately 17% of these significant pairs (107,365 pairs) were physically linked ($r^2 > 0.2$), with an average r^2 value of 0.50. Chromosome 3 exhibited the highest number of marker pairs showing significant LD (30,860 pairs; $r^2 = 0.27$), while chromosome 9 displayed the lowest (23,438 pairs; $r^2 = 0.25$) (Table 5.1). Notably, chromosomes 1, 2 and 3 displayed the strongest LD among significant marker pairs (mean $r^2 = 0.27$), while chromosomes 4, 6 and 10 exhibited the weakest LD (mean $r^2 = 0.24$) (Table 5.1). At the genomic level, the overall r^2 value stood at 0.12. The LD decay curve began at an r^2 value of 0.47 and reached half-decay at 0.24. The half-decay line intersected the LD decay curve at a distance of 7.6 kilobases (kb) (Figure 5.1).

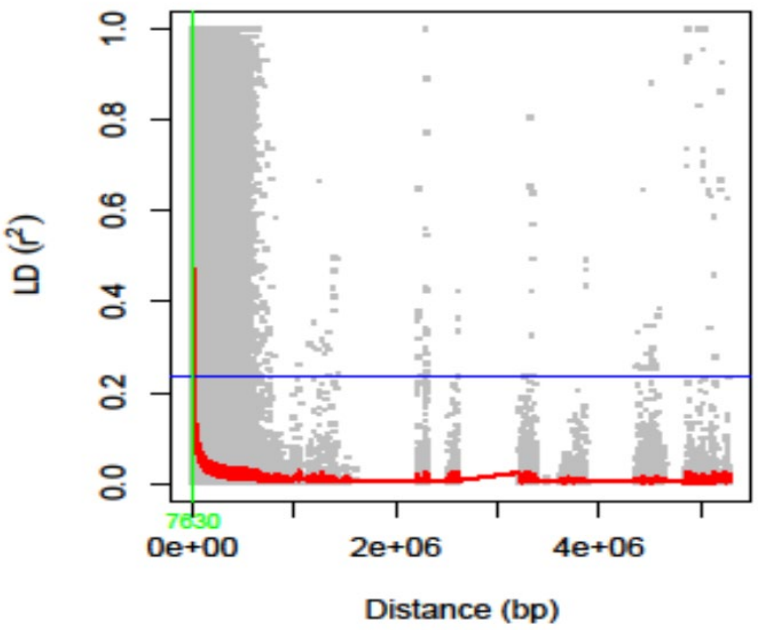


Figure 5.1 The scatter plot of genome-wide linkage disequilibrium (LD) decay determined based on the r^2 values of the marker pairs. The red curve line is the smoothing spline regression model fitted to LD decay. The horizontal blue line is the half decay r^2 value of the genome ($r^2 = 0.24$), whereas the vertical green line is the genetic distance between markers in base pair (bp) (7.6 kb) at the intersection between the half decay and the LD decay curve.

Table 5.1 A summary of the results of linkage disequilibrium (LD) analyses among SNP marker pairs, including the number of significant marker pairs for each chromosome

Chr	Total number of pairs	Average r^2 for all pairs	D in kb	Number of significant pairs*	Average r^2 for all significant pairs	D in kb	Number of PLP (r^2 >0.2) **	Average r^2 for PLP	D in kb	Complete LD (r^2 = 1)
1	59821	0.12	59.44	25999 (43)	0.27	54.25	11344 (19)	0.49	46.8	644
2	62561	0.13	47.53	28802 (46)	0.27	43.27	12393 (20)	0.51	36.9	972
3	61417	0.14	58.48	30860 (50)	0.27	50.92	13303 (22)	0.51	40.6	821
4	62817	0.12	43.99	29580 (47)	0.24	39.14	10876 (17)	0.5	30.4	716
5	61962	0.11	50.97	25202 (41)	0.25	46.99	9897 (16)	0.49	42.38	685
6	61159	0.10	57.79	22603 (37)	0.24	50.71	8230 (13)	0.48	42.94	522
7	61699	0.13	43.12	28818 (47)	0.26	39.56	11823 (19)	0.5	31.1	922
8	63500	0.12	42.68	28024 (44)	0.26	38.33	10957 (17)	0.5	33.87	815
9	61414	0.11	41.6	23438 (38)	0.25	36.76	8991 (15)	0.52	30.83	549
10	62708	0.11	52.4	26131 (42)	0.24	48.01	9551 (15)	0.51	36.14	618
T/M	619058	0.12	49.80	269457 (44)	0.26	44.79	107365 (17)	0.50	37.20	6620

Chr = Chromosome; D = Physical distance; kb = kilobase; PLP = physically linked pairs; r^2 = squared allele-frequency correlations between alleles at the two loci; T/M = Total/Mean; Significant marker pairs = $P \leq 0.01$; * Numbers in the brackets are the percentage of the total number of marker-pairs; ** Numbers in the brackets are the percentage of the number of significant marker pairs.

5.4.2 Model comparison for genome-wide association study analysis

BLUP for measured traits of each sorghum landrace across six environments was used and compared using five different association mapping (AM) models, ranging from simple to complex (Figure 5.2). The five different statistical multiple models, GLM, MLM, CMLM, FarmCPU and BLINK, were tested for their ability to identify true genetic associations with grain quality traits while minimising false positives and negatives. The quantile-quantile (QQ) plots were used to visualise false associations and is a graphical tool used to compare the observed distribution of *p*-values from a statistical test with the expected distribution under the null hypothesis of no association. If the observed *p*-values deviate from the expected distribution, it can indicate the presence of false positives or negatives (Weine et al., 2021).

Results indicated that the GLM model generated false positives and the line generated from the model deviated from the expected diagonal line for amylose and protein. In contrast, the MLM and CMLM models produced *p*-values that plotted close to the expected line but only detected significant associations for amylose. Both the BLINK and FarmCPU models performed well, producing *p*-values that were plotted closest to the expected diagonal line and identifying significant associations for the grain quality traits. These complex models identified highly significant SNPs after adjustments to the Holm-Bonferroni test (HB p-value ≤ 0.05) (Massimiliano et al., 2018) across all traits. By reducing *p*-value inflations, these models were able to control false positive and negative error rates.

5.4.3 Genome-wide associations studies across environments

The GWAS study identified 67 different SNPs significantly associated with starch (21), amylose (37) and protein (9) after adjustment (HB p-value ≤ 0.05) using the BLINK and FarmCPU models (Appendix 4). Furthermore, there was detection of different SNPs in the ability of BLINK and FarmCPU to detect significantly associated markers when covariates were included (Table 5.2).

Because of the highly significant interaction detected between genotypes and environments in the previous chapters, GWAS was conducted separately for each environment. Using BLUP values, 11 different significant SNPs were identified for starch (5), amylose (3) and protein (3).

Significantly associated SNPs linked to starch, amylose and protein detected in the individual environments using GWAS, are presented in Table 5.2.

Five SNPs, S01_73075918 (chromosome 1), S01_31764844 (chromosome 1), S04_53545349 (chromosome 4), S06_48180471 (chromosome 6) and S06_59687350 (chromosome 6), were significantly associated with starch content and were detected across environments. Only S06_59687350 was detected in all three locations; S01_31764844 was detected at both Jimma and Melkassa, while SNPs S01_73075918 and S06_48180471 were only detected at Jimma and S04_53545349 was detected at Melkassa.

Three SNPs, S01_67415726, S07_62671638 and S10_21605004, located on chromosomes 1, 7 and 10, respectively, were significantly associated with amylose content. All three SNPs were detected in all locations (Jimma, Melkassa and Miesso) across both years (2020 and 2021), except for S01_67415726 and S10_21605004 which were not detected in 2020 in Jimma.
Three SNPs, S01_1332500, S02_3743197 and S09_3030812, located on chromosomes 1, 2 and 10, respectively, were significantly associated with protein content and both were only detected under dry lowlands (Melkassa and Miesso).

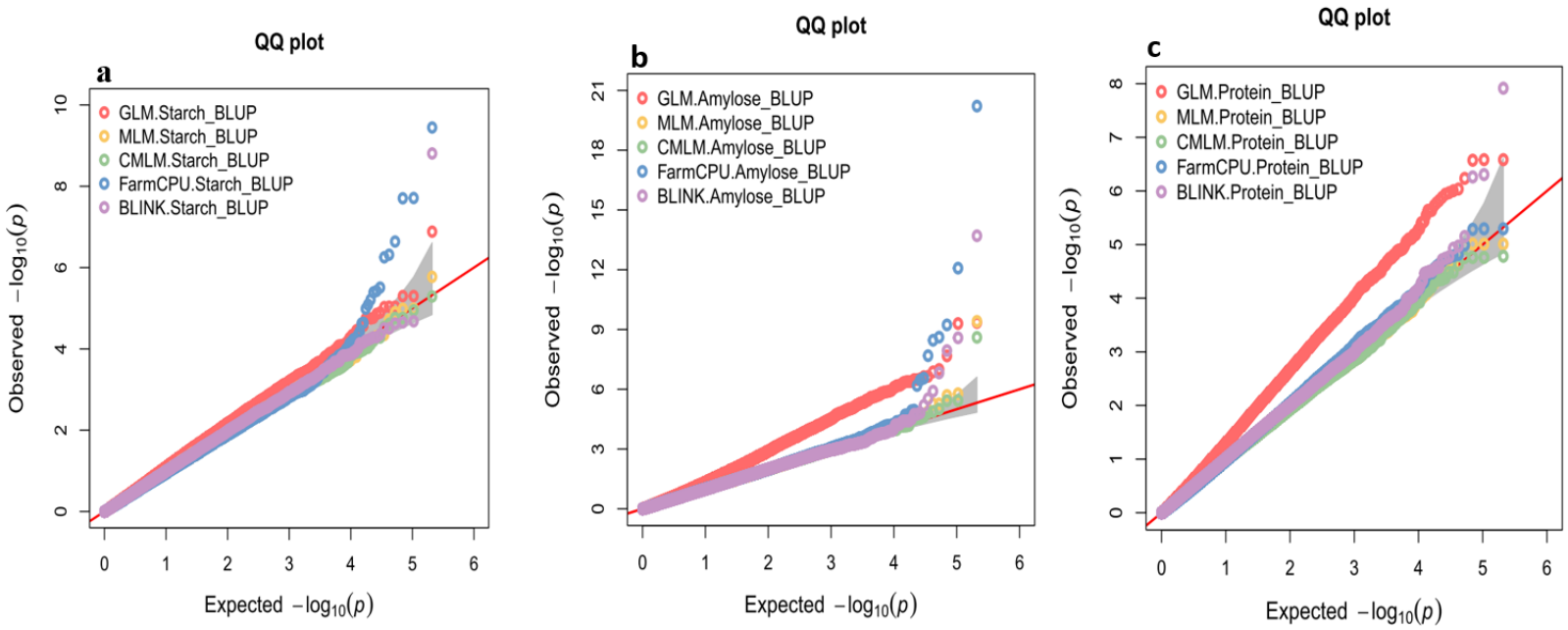


Figure 5.2 Genome-wide association analysis of grain quality across environments using different statistical models, best linear unbiased prediction analysis and quantile-quantile (QQ) plots for starch (a), amylose (b) and protein (c).

GLM = general linear model; MLM = mixed linear model; CMLM = compressed mixed linear model, FarmCPU = fixed and random model circulating probability unification; BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; BLUP= best linear unbiased prediction.

Table 5.2 Summary of SNPs consistently and significantly associated with amylose, protein and composition in sorghum landraces across years and locations

Trait	Year	Location	Model	SNP	Chr	Pos	*p-value*	Effect	Associated gene	Putative gene description
Starch	2020	Jimma	FarmCPU	S01_31764844	1	31764844	4.78E-09	-0.08	*Sobic.001G260000*	Zinc finger, CCCH-type
Starch	2021	Melkassa	FarmCPU	S01_31764844	1	31764844	1.32E-07	-0.16	*Sobic.001G260000*	Zinc finger, CCCH-type
Starch	2020	Jimma	FarmCPU	S01_73075918	1	73075918	*4.07E-07*	0.07	*Sobic.001G454100*	Sugar carrier protein C
Starch	2020	Melkassa	FarmCPU	S04_53545349	4	53545349	2.54E-05	0.17	*Sobic.004G182000*	Late embryogenesis abundant (LEA)
Starch	2021	Jimma	BLINK	S06_48180471	6	48180471	*2.38E-09*	-0.19	*Sobic.006G114100*	ABC transporter G family member 31 (ABCG31)
Starch	2020	Jimma	FarmCPU	S06_48180471	6	48180471	*3.96E-07*	-0.07	*Sobic.006G114100*	ABC transporter G family member 31 (ABCG31)
Starch	2021	Jimma	FarmCPU	S06_48180471	6	48180471	*6.14E-10*	-0.15	*Sobic.006G114100*	ABC transporter G family member 31 (ABCG31)
Starch	2020	Jimma	BLINK	S06_59687350	6	59687350	*7.24E-10*	-0.09	*Sobic.006G262100*	Auxin response factor 12 (ARF12)
Starch	2021	Melkassa	BLINK	S06_59687350	6	59687350	*6.87E-10*	-0.21	*Sobic.006G262100*	Auxin response factor 12 (ARF12)
Starch	2021	Miesso	BLINK	S06_59687350	6	59687350	*1.15E-09*	-0.15	*Sobic.006G262100*	Auxin response factor 12 (ARF12)
Starch	2021	Melkassa	FarmCPU	S06_59687350	6	59687350	*5.73E-07*	-0.13	*Sobic.006G262100*	Auxin response factor 12 (ARF12)
AMY	2021	Jimma	BLINK	S01_67415726	1	67415726	*2.88E-11*	0.12	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2020	Melkassa	BLINK	S01_67415726	1	67415726	*2.38E-09*	0.09	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2021	Melkassa	BLINK	S01_67415726	1	67415726	*3.07E-09*	0.1	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2020	Miesso	BLINK	S01_67415726	1	67415726	*1.22E-09*	0.09	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2021	Miesso	BLINK	S01_67415726	1	67415726	*9.51E-11*	0.09	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)

Table 5.2 Summary of SNPs consistently and significantly associated with amylose, protein and composition in sorghum landraces across years and locations (continued)

Trait	Year	Location	Model	SNP	Chr	Pos	*p-value*	Effect	Associated gene	Putative gene description
AMY	2021	Jimma	FarmCPU	S01_67415726	1	67415726	*4.89E-08*	0.08	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2020	Melkassa	FarmCPU	S01_67415726	1	67415726	*3.44E-08*	0.06	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2021	Melkassa	FarmCPU	S01_67415726	1	67415726	*4.24E-08*	0.06	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2020	Miesso	FarmCPU	S01_67415726	1	67415726	*3.89E-10*	0.07	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2021	Miesso	FarmCPU	S01_67415726	1	67415726	*2.03E-08*	0.07	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	2020	Jimma	BLINK	S07_62671638	7	62671638	*8.74E-09*	0.06	*Sobic.007G193900*	Os08g0531200
AMY	2021	Jimma	BLINK	S07_62671638	7	62671638	*6.43E-17*	0.13	*Sobic.007G193900*	Os08g0531200
AMY	2020	Melkassa	BLINK	S07_62671638	7	62671638	*9.68E-15*	0.11	*Sobic.007G193900*	Os08g0531200
AMY	2021	Melkassa	BLINK	S07_62671638	7	62671638	*1.42E-14*	0.11	*Sobic.007G193900*	Os08g0531200
AMY	2020	Miesso	BLINK	S07_62671638	7	62671638	*1.09E-20*	0.14	*Sobic.007G193900*	Os08g0531200
AMY	2021	Miesso	BLINK	S07_62671638	7	62671638	*1.50E-16*	0.11	*Sobic.007G193900*	Os08g0531200
AMY	2020	Jimma	FarmCPU	S07_62671638	7	62671638	*5.70E-14*	0.09	*Sobic.007G193900*	Os08g0531200
AMY	2021	Jimma	FarmCPU	S07_62671638	7	62671638	*6.13E-18*	0.14	*Sobic.007G193900*	Os08g0531200
AMY	2020	Melkassa	FarmCPU	S07_62671638	7	62671638	*1.69E-28*	0.15	*Sobic.007G193900*	Os08g0531200
AMY	2021	Melkassa	FarmCPU	S07_62671638	7	62671638	*1.96E-28*	0.16	*Sobic.007G193900*	Os08g0531200
AMY	2020	Miesso	FarmCPU	S07_62671638	7	62671638	*7.47E-27*	0.14	*Sobic.007G193900*	Os08g0531200
AMY	2021	Miesso	FarmCPU	S07_62671638	7	62671638	*4.89E-19*	0.12	*Sobic.007G193900*	Os08g0531200
AMY	2021	Jimma	BLINK	S10_21605004	10	21605004	*2.32E-09*	0.09	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2020	Melkassa	BLINK	S10_21605004	10	21605004	*3.83E-10*	0.08	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2021	Melkassa	BLINK	S10_21605004	10	21605004	*4.77E-10*	0.09	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein

Table 5.2 Summary of SNPs consistently and significantly associated with amylose, protein and composition in sorghum landraces across years and locations (continued)

Trait	Year	Location	Model	SNP	Chr	Pos	*p-value*	Effect	Associated gene	Putative gene description
AMY	2020	Miesso	BLINK	S10_21605004	10	21605004	*1.03E-10*	0.08	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2021	Miesso	BLINK	S10_21605004	10	21605004	*4.54E-11*	0.08	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2021	Jimma	FarmCPU	S10_21605004	10	21605004	*1.00E-10*	0.08	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2020	Melkassa	FarmCPU	S10_21605004	10	21605004	*2.73E-13*	0.07	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2021	Melkassa	FarmCPU	S10_21605004	10	21605004	*3.82E-13*	0.08	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2020	Miesso	FarmCPU	S10_21605004	10	21605004	*3.48E-14*	0.07	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	2021	Miesso	FarmCPU	S10_21605004	10	21605004	*1.30E-07*	0.06	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
Protein	2020	Melkassa	BLINK	S01_1332500	1	1332500	3.00E-06	0.11	*Sobic.001G014900*	Long-chain-fatty-acid
Protein	2020	Melkassa	BLINK	S02_3743197	2	3743197	*2.98E-08*	0.19	*Sobic.002G038200*	TatD family like protein
Protein	2021	Melkassa	BLINK	S02_3743197	2	3743197	*6.48E-07*	0.3	*Sobic.002G038200*	TatD family like protein
Protein	2020	Miesso	BLINK	S02_3743197	2	3743197	*3.54E-09*	0.3	*Sobic.002G038200*	TatD family like protein
Protein	2020	Melkassa	BLINK	S09_3030812	9	3030812	*5.49E-12*	0.32	*Sobic.009G033500*	Class III peroxidase 70 precursor
Protein	2020	Melkassa	BLINK	S09_3030812	9	3030812	*5.49E-12*	0.32	*Sobic.009G033400*	Class III peroxidase 69 precursor

Holm-Bonferroni with adjustment = HB p.value ≤ 0.05; AMY = amylose; SNP = single nucleotide polymorphism; Chr = Chromosome; Pos = position on chromosome; BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; FarmCPU = fixed and random model circulating probability unification
The candidate genes were identified using the reference genome data retrieved from the sorghum genome browser *Sorghum bicolor* v3.1.1, which is accessible on Phytozome (https://phytozome-next.jgi.doe.gov/info/Sbicolo).

5.4.4 Genome wide association study for grain quality traits in sorghum

Starch

Major QTL regions within 100 kb upstream and downstream of QTL peaks have been investigated to identify candidate genes associated with the variations. A total of 21 SNPs were found to be significantly (HB *p*-value ≤ 0.05) associated with starch content. The SNPs distributed across the ten sorghum chromosomes (Appendix 4) of which five were highly significant and appeared as major peaks. The S01_31764844 SNP was located within the gene region of *Sobic.001G260000*. The S01_31764844 SNP was located within the gene region of *Sobic.001G260000*, which encodes a protein known as zinc finger, CCCH-type. SNP S06_59687350 was detected by both models within the *Sobic.006G262100* gene on chromosome 6, with a functional annotation for ARFs through selection of target genes as transcription factors. FarmCPU detected an additional four closely linked SNPs; S01_73075918, S01_31764844, S04_53545349 and S06_48180471. The GWAS approach mapped S01_73075918 and S01_31764844 both to chromosome 1 and S04_53545349 to chromosome 4, while S06_48180471 and S06_59687350 both mapped to chromosome 6 (Figure 5.3 and Table 5.3).

Genes that were close to the identified SNPs were considered potential candidate genes. The S01_73075918 SNP was located within the gene region of *Sobic.001G454100*, which encodes a putative sugar carrier protein. While the exact function of this gene in sorghum grain starch metabolism is not fully understood, this protein may play a role in the transport of sugars into and out of the starch granule during grain development and maturation. The S04_53545349 SNP was located within the gene region of *Sobic.004G182000*, which is annotated as encoding a LEA hydroxyproline-rich glycoprotein-related protein (HRGP).

The S06_48180471 SNP was located within the gene region of *Sobic.006G114100*. The *Sobic.006G114100* gene encodes the ABC transporter G family member 31 (ABCG31) protein in sorghum. The S06_59687350 SNP was located within the gene region of *Sobic.006G262100*, which encode a protein known as ARF12.

Amylose

A total of 37 SNPs were significantly associated with amylose across environments in all models, with a significance level of HB *p*-value ≤ 0.05 (Appendix 4). Among these significant SNPs, the three SNPs located as major peaks at 64.4 on chromosome 1, 62.6 on chromosome 7 and 21.6 Mpb on chromosome 10 were consistently detected by both BLINK and FarmCPU models across environments (Figure 5.3) in BLUP analysis. The SNP on chromosome 7, S07_62671638, showed significant associations with amylose for all five models tested: GLM, MLM, CMLM, FarmCPU and BLINK (Table 5.2). FarmCPU detected five additional highly significant SNPs, which are located at different loci on chromosomes 1, 6 and 10 (Figure 5.3). The SNP positions were used to examine the sorghum genome annotation database to find candidate genes closely linked to these SNPs.

The nearest gene to SNP S01_67415726, located on chromosome 1, was *Sobic.001G386600*, which encodes a protein with similarity to WOX6. The nearest gene to SNP S07_62671638, located on chromosome 7, was *Sobic.007G193900*. This gene encodes the Os08g0531200 protein, a putative alpha-glucan phosphorylase. SNP S10_21605004, located on chromosome 10 of the sorghum, was nearest to the *Sobic.010G137333* gene. The zinc finger CCHC domain-containing protein encoded by the *Sobic.010G137333* gene.

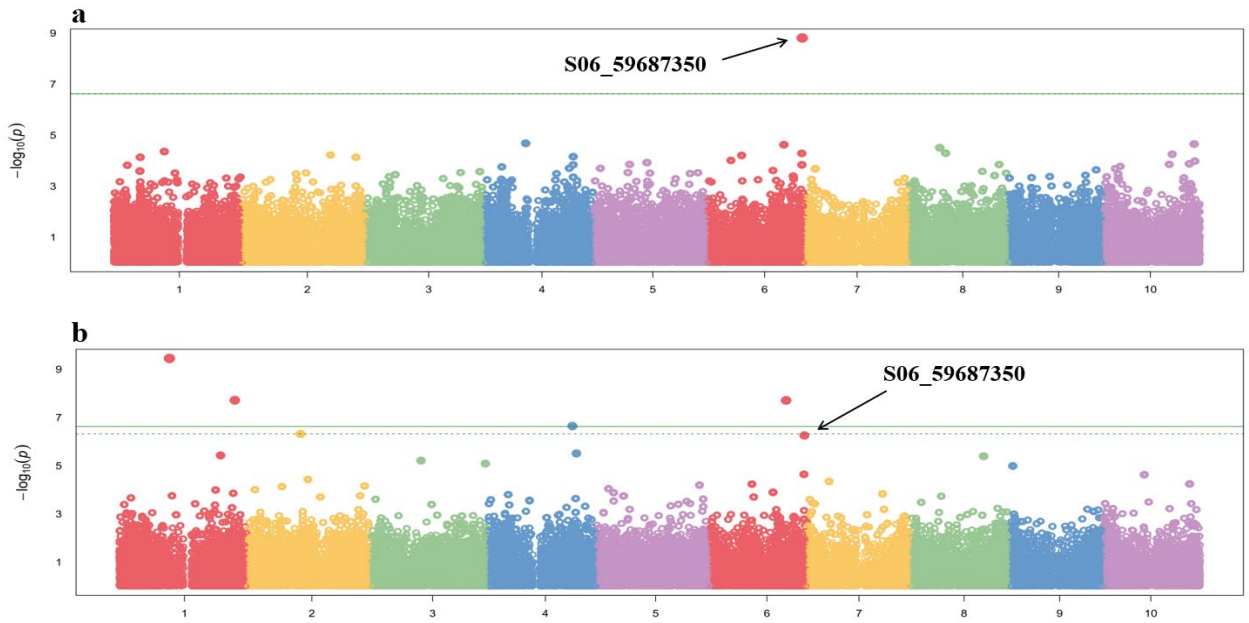


Figure 5.3 Genome-wide association study for sorghum landrace grain quality using the best linear unbiased prediction (BLUP). Manhattan plots for starch using (a) the Bayesian information and linkage disequilibrium iteratively nested keyway (BLINK) and (b) the fixed and random model circulating probability unification (FarmCPU) models.

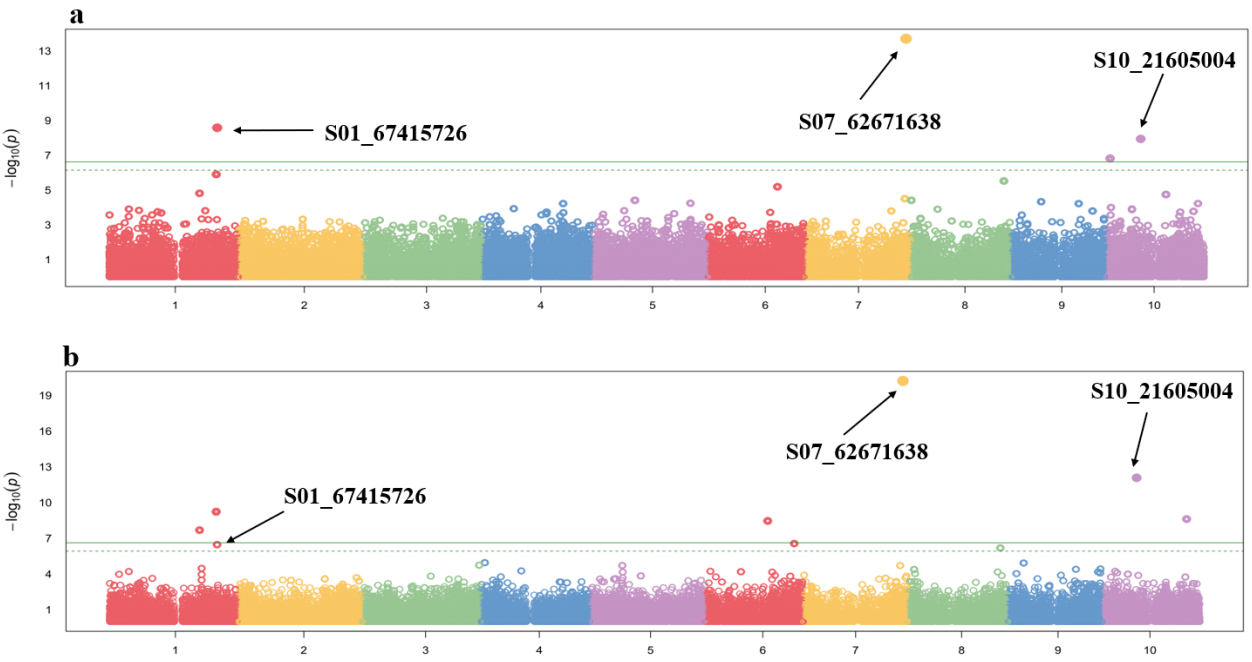


Figure 5.4 Genome-wide association study for sorghum landraces' grain quality using the best linear unbiased prediction (BLUP). Manhattan plots for amylose using (a) the Bayesian information and linkage disequilibrium iteratively nested keyway (BLINK) and (b) the fixed and random model circulating probability unification (FarmCPU) models.

Protein

A total of nine SNPs were significantly associated with protein content across environments with a significance level of HB *p*-value ≤ 0.05 (Appendix 4). Three SNPs; S01_1332500, S02_3743197 and S09_3030812 were significantly associated with protein content (Figure 5.5 and Table 5.3). FarmCPU, CMLM and MLM failed to detect any significant association with protein content using BLUP and GWAS analysis, while GLM detected two and BLINK three SNPs (HB *p*-value ≤ 0.05). Genes *Sobic.001G014900, Sobic.02G038200* and *Sobic.09G033500* were candidate genes located close to SNPs identified on chromosomes 1, 2 and 9, respectively. *Sobic.001G014900* is a gene found in sorghum on chromosome 1, encoding a long-chain-fatty-acid enzyme. *Sobic.02G038200* is a candidate gene located on chromosome 2, encoding a protein that belongs to the TatD family of hydrolases. *Sobic.09G033500* and *Sobic.009G033400* are candidate genes located on chromosome 9, encoding proteins similar to Class III peroxidase 70 precursor and Class III peroxidase 69 precursor, respectively, which are members of the peroxidase enzyme family.

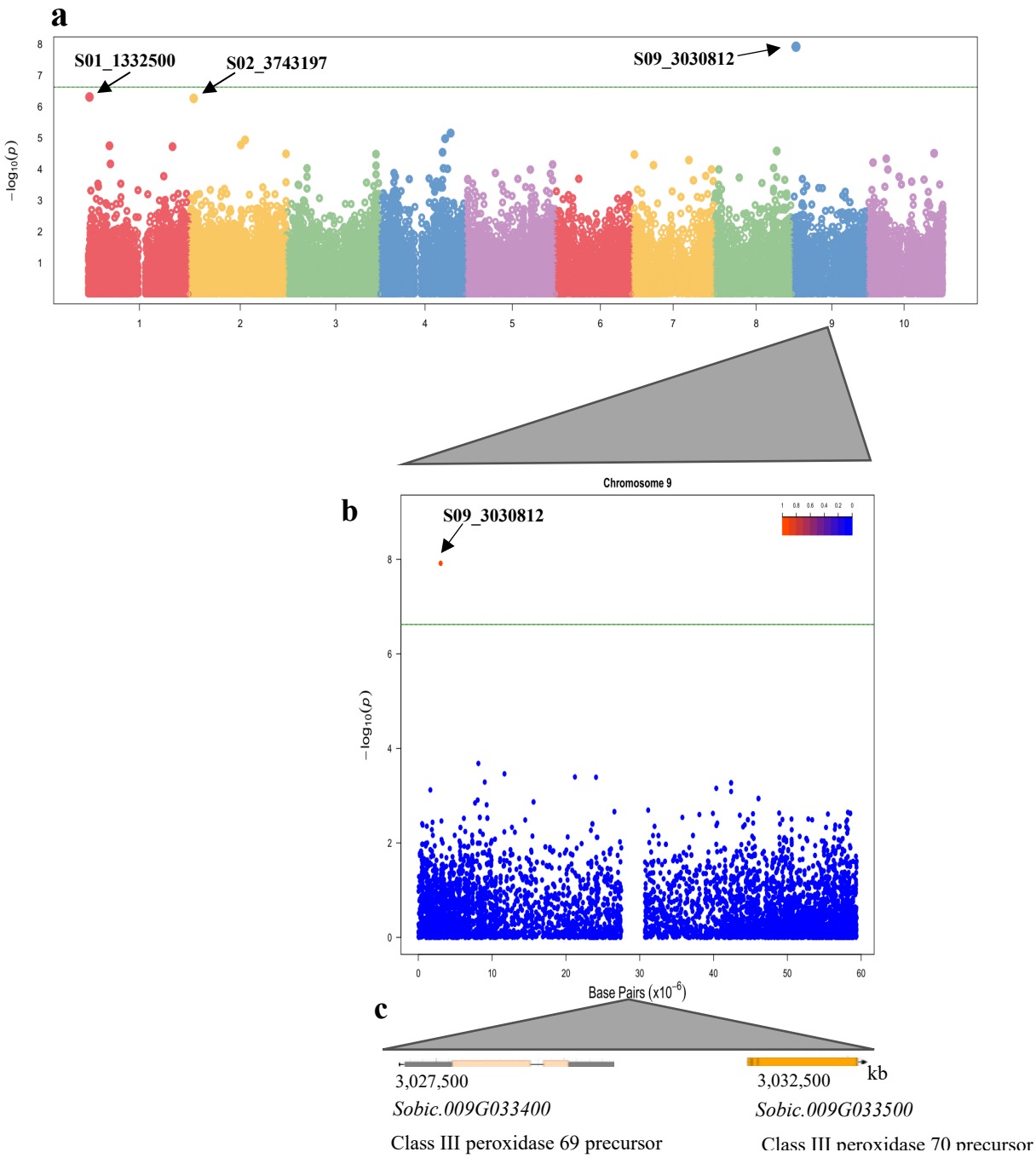


Figure 5.5 Candidate Class III peroxidase 69 precursor and Class III peroxidase 70 precursor genes associated with protein content in sorghum landraces. (a) Manhattan plot of genome-wide association study for proteins using the Bayesian information and linkage disequilibrium iteratively nested keyway (BLINK) model. The significant SNP linked to protein content is indicated with an arrow. (b) Manhattan plot of chromosome 9 indicating the genomic location of the peroxidase genes at 30.3 Mbp. (c) The specifics of the significant locus at indicate the location of candidate genes for the significant SNP (S09_3030812).

Table 5.3 Summary of significant SNPs, associated genes and putative gene descriptions consistently associated with grain quality content in sorghum landrace collections in the best linear unbiased prediction (BLUP)

Trait	Model	SNP	Chr	Pos	*p*-value	HB *p*-value ≤0.05	Effect	Associated gene	Putative gene description
Starch	FarmCPU	S01_31764844	1	31764844	3.58E-10	8E-05	-0.16	*Sobic.001G260000*	Zinc finger, CCCH-type
Starch	FarmCPU	S01_73075918	1	73075918	1.94E-08	1E-03	0.14	*Sobic.001G454100*	Sugar carrier protein C
Starch	FarmCPU	S04_53545349	4	53545349	2.29E-07	1E-02	0.08	*Sobic.004G182000*	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein-related protein
Starch	FarmCPU	S06_48180471	6	48180471	1.97E-08	1E-03	-0.12	*Sobic.006G114100*	ABC transporter G family member 31 (ABCG31)
Starch	GLM	S06_59687350	6	59687350	1.31E-07	0.03	-0.16	*Sobic.006G262100*	Auxin response factor 12 (ARF12)
Starch	BLINK	S06_59687350	6	59687350	1.56E-09	0.00	-0.16	*Sobic.006G262100*	Auxin response factor 12 (ARF12)
Starch	FarmCPU	S06_59687350	6	59687350	5.59E-07	2E-02	-0.11	*Sobic.006G262100*	Auxin response factor 12 (ARF12)
AMY	BLINK	S01_67415726	1	67415726	2.62E-09	3E-04	0.08	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	FarmCPU	S01_67415726	1	67415726	3.43E-07	9E-03	0.06	*Sobic.001G386600*	WUSCHEL-related homeobox 6 (WOX6)
AMY	CMLM	S07_62671638	7	62671638	2.49E-09	0.00	0.13	*Sobic.007G193900*	Os08g0531200
AMY	MLM	S07_62671638	7	62671638	3.90E-10	8E-05	1E-01	*Sobic.007G193900*	Os08g0531200
AMY	GLM	S07_62671638	7	62671638	5.03E-10	5E-05	0.12	*Sobic.007G193900*	Os08g0531200
AMY	BLINK	S07_62671638	7	62671638	1.98E-14	4E-09	0.1	*Sobic.007G193900*	Os08g0531200
AMY	FarmCPU	S07_62671638	7	62671638	6.11E-21	1E-15	0.13	*Sobic.007G193900*	Os08g0531200
AMY	BLINK	S10_21605004	10	21605004	1.15E-08	8E-04	0.07	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
AMY	FarmCPU	S10_21605004	10	21605004	8.32E-13	9E-08	0.07	*Sobic.010G137333*	Zinc finger CCHC domain-containing protein
Protein	BLINK	S01_1332500	1	1332500	4.94E-07	0.04	0.15	*Sobic.001G014900*	Long-chain-fatty-acid
Protein	GLM	S02_3743197	2	3743197	2.61E-07	0.00	0.28	*Sobic.002G038200*	Hydrolase, the TatD family like protein
Protein	BLINK	S02_3743197	2	3743197	5.47E-07	0.04	0.21	*Sobic.002G038200*	Hydrolase, the TatD family like protein
Protein	BLINK	S09_3030812	9	3030812	1.21E-08	0.00	0.33	*Sobic.009G033500*	Class III peroxidase 70 precursor
Protein	BLINK	S09_3030812	9	3030812	1.21E-08	0.00	0.33	*Sobic.009G033400*	Class III peroxidase 69 precursor

Holm-Bonferroni with adjustment = HB *p*.value ≤ 0.05; AMY = amylose; SNP = single nucleotide polymorphism; Chr = Chromosome; GLM = general linear model; MLM = mixed linear model; CMLM = compressed mixed linear model; FarmCPU = fixed and random model circulating probability unification; BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; BLUP= best linear unbiased prediction; Pos = position on chromosome.

5.5 Discussion

Sorghum is an important cereal crop grown globally for human consumption, animal feed and industrial uses (Zheng et al., 2011). Grain quality is an essential trait for sorghum breeding programmes and there has been increasing interest in identifying genetic markers associated with grain quality traits through GWAS and candidate gene approaches (Rhodes et al., 2014; Boyles et al., 2017). GWAS involves screening many genetic markers across the entire genome of unrelated individuals in a population to identify markers associated with a trait of interest (Yang et al., 2015; Alqudah et al., 2020). The candidate gene approach, on the other hand, involves selecting genes known to be involved in the pathway or function of a trait of interest and screening for genetic variation within those genes (Alenezi et al., 2023). This study presents findings from GWAS of a large collection of Ethiopian sorghum landraces, which enabled the identification of novel loci associated with sorghum grain quality traits.

Linkage disequilibrium

Understanding the LD pattern is vital for shaping association studies and molecular breeding strategies (Thornsberry and Buckler, 2003). In this research, LD decay was assessed at both chromosome and genome levels using SNP data from 365 sorghum landraces. The average r^2 value across the genome was 0.12, with LD initiating at an r^2 value of 0.47 and reaching half-decay ($r^2 = 0.24$) within 7.6 kb. These LD decay findings align closely with those reported for Ethiopian sorghum accessions (Wondimu et al., 2023), falling within the range of 10-15 kb. However, they lower previously published values of $r^2 < 0.1$ within 150 kb (Morris et al., 2013) and 100 kb (Bouchet et al., 2012). The average r^2 across sorghum chromosomes demonstrated a consistent LD decay rate, ranging from 0.10 to 0.14, consistent with prior research on sorghum (Wang et al., 2013; Hu et al., 2019; Enyew et al., 2022). LD generally declined rapidly with distance ($r^2 < 0.1$ within 1500 bp), but rates of decline were highly variable among genes. This rapid decline probably reflects large effective population sizes in maize during its evolution and high levels of recombination within genes (Remington et al., 2001).

Evaluation of genome-wide association study models

Five AM models, namely GLM (Price et al., 2006), MLM (Yu et al., 2006), CMLM (Zhang et al., 2010), FarmCPU (Liu et al., 2016) and BLINK (Huang et al., 2019) were compared for efficacy in detecting markers linked to grain quality traits. The BLINK and FarmCPU models were the most efficient in detecting linked markers in the Ethiopian sorghum landrace accessions. The two multiple models, MLM and CMLM, were extremely conservative, as no

significant SNP markers were detected. In contrast, the GLM model detected an excessive number of linked markers.

The BLINK and FarmCPU models, using more complex methods of Holm-Bonferroni adjustments (HB p-value ≤ 0.05) (Massimiliano et al., 2018), detected 67 linked SNP markers. The BLINK and FarmCPU models may be more efficient for complex populations and also improve computing speed and increase the capacity to analyse big genomic data sets, while CMLM, MLM and GLM may be more efficient for simpler populations or different types of traits (Kaler et al., 2020; Wang and Zhang, 2021; Nida et al., 2021). Malik et al. (2019) compared the performance of three GWAS models, MLM, CMLM and FarmCPU, for association analysis of yield and straw quality traits in wheat; FarmCPU outperformed MTAs for plant height, yield and harvest index. Both real and simulated data analyses revealed that BLINK outperformed FarmCPU in terms of statistical power and computing time. BLINK is the most recently developed GWAS model and eliminates problems associated with FarmCPU. BLINK replaces the FarmCPU's random effect component with the fixed effect model using Bayesian information criteria (Huang et al., 2019).

Based on the comparison of QQ plots and HB p-values ≤ 0.05, BLINK and FarmCPU consistently outperformed CMLM, MLM and GLM for all sorghum grain quality traits analysed, suggesting that these models were suited for analysing the data under consideration for this study. A comparison of QQ plots using different models for all quality traits revealed that the BLINK and FarmCPU models controlled false positives and negatives better. Furthermore, BLINK and FarmCPU identified SNP markers that were closer to the published location of genes controlling these traits than the other models. Instead of a large peak area, as in other models, the BLINK and FarmCPU models provided the most significant marker that was always present and closest to the published genes.

Identification of linked markers across environments

In this study, significant SNP markers were detected across environments, but some of them are still poorly defined as candidate genes in the phytozome database. The majority of SNP markers in plants have not been discovered and defined as candidate genes in the phytozome (Ganal et al., 2009; Hassanyar et al., 2023). During the GWAS analysis, the two multiple models FarmCPU and BLINK detected the most significantly associated SNPs across environments. The GWAS was conducted separately for each environment due to the highly

significant interaction between genotypes and environments. The GWAS identified SNPs that were significantly associated with starch, amylose and protein across all environments. However, some SNPs were only detected in certain environments while other SNPs were detected in all environments, indicating the differential expression of genes/QTL linked to these SNPs under different environmental conditions.

The five significant SNPs linked to starch content, S01_31764844, S01_73075918, S04_53545349, S06_48180471 and S06_59687350, were detected across environments differently. All SNPs were detected at Jimma except for S04_53545349, which was only detected at Melkassa in 2020. Two of these SNPs, S01_73075918 and S06_48180471, were only detected at Jimma; at Miesso, only one SNP, S06_59687350, was detected.

The three SNPs linked to protein content were only detected under dry lowland conditions but not at Jimma (intermediate), S01_1332500 and S02_3743197, were only detected at Melkassa, while S09_3030812 was detected at Melkassa and Miesso. The three SNPs (S01_67415726, S07_62671638 and S10_21605004) linked to amylose content were detected in all environments, but the two SNPs (S01_67415726 and S10_21605004) were not detected in 2020 in Jimma. This indicates that the detected genes/QTL linked to amylose content may be less influenced by environmental factors and may have a more consistent expression across different environments.

In the trial growing seasons (2020 and 2021), Jimma (wet intermediate) and Melkassa (irrigated) received consistent moisture, an intermediate temperature and had better soil macronutrients and organic matter than Miesso (stress), as described in section 3.3.3 and Tables 4.3 and 4.4. As a result, more genes/QTL were expressed at Jimma and Melkassa than at Miesso for starch, amylose and protein content. In 2011, more genes/QTL were expressed for all traits across all sites than in 2020. This suggests that the expression of genes/QTL linked to these SNPs may be influenced by specific environmental factors, such as climate, rain fall, soil type and management practices, which can affect starch, amylose and protein contents differently in different regions. The QTLs are differentially expressed due to the GEI under two soil moisture treatments (Messmer et al., 2009). Starch synthesis and phytohormone biosynthesis are regulated by post-anthesis moderate soil drying treatment and differentially expressed miRNAs in inferior spikelets (Teng et al., 2022). Zhao et al. (2023) reported on QTL linked to agronomic traits in rice that were differentially expressed in different environments and it was

predicted that the biological process was related to salt and water stresses and light intensity in the annotation.

Genome-wide associations and candidate genes

This study used the genome assembly version *Sorghum bicolor* v3.1.1 from Phytozome (Goodstein et al., 2012) and identified putative proteins and candidate genes through reference genome data retrieved from the sorghum genome browser *Sorghum bicolor* v3.1.1, available on Phytozome. This ensures transparency and clarity regarding the genomic reference used for alignment and subsequent analyses.

Starch

Five novel loci associated with major effects on sorghum starch biosynthesis have been detected. The first SNP on chromosome 1 has been linked to the *Sobic.001G260000* gene, which has a functional annotation to a putative zinc finger, CCCH-type proteins in sorghum and plays a role in plant growth and stress response. The CCCH-type zinc finger proteins have important roles in plant growth, development and stress adaptation, including heat and salt tolerance (Xu et al., 2023; Zhang et al., 2023).

The second SNP located on chromosome 1 was linked to the *Sobic.001G454100* gene with functional annotation to a putative sugar carrier protein C and may play a role in transporting sugars within the sorghum plant. During grain development, sugars are synthesised in leaves and transported to the developing grain, where it is converted to starch and stored in the endosperm (Li et al., 2022). Sugar transporters, such as the putative carrier protein encoded by this gene, are likely involved in this process. Starch is the major component of cereal grains, including sorghum, and is synthesised and stored in specialised organelles called amyloplasts. Transport of sugars to and from these organelles is critical for starch synthesis and accumulation. Given its similarity to a sugar carrier protein, expression of this gene may play a role in transporting sugars into amyloplasts for starch synthesis, or in exporting sugars out of amyloplasts for use in other metabolic processes (Dupont, 2008).

The third SNP located on chromosome 4 was linked to the *Sobic.004G182000* gene and has been annotated to encode a LEA hydroxyproline-rich glycoprotein (HRGP)-related protein. Similarly, Sukumaran et al. (2012) detected the starch synthesis gene *SSIIb* on chromosome 4 in sorghum collections. These LEA proteins are known to be involved in various stress

responses, including drought, salinity and cold (Shinozaki and Yamaguchi-Shinozaki, 1997) and show higher concentrations during the later stages of seed development (Liu et al*.,* 2013). LEA3 proteins are involved in both abiotic and biotic stress tolerance in maize (Liu et al., 2013), aflatoxin resistance in maize grains (Chen et al., 2002; 2007) and resistance to head blight in wheat (Liu et al., 2019) and barley (Huang et al., 2016). Nida et al. (2021) indicated that this locus could also impact the *LEA3* gene, which encodes a protein that accumulates in sorghum seeds. On the other hand, HRGPs are a group of cell wall proteins that have been implicated in plant growth and development, as well as stress responses (Shailasree et al., 2004). They are characterised by a high hydroxyproline (a modified amino acid) content and are often glycosylated (addition of sugar molecules on side chains) (Showalter, 1993). These proteins could thus possibly protect sorghum plants from damage caused by environmental stresses by maintaining the stability of cellular components such as membranes and proteins. Sequence variation in this gene could affect the level or activity of the LEA3 protein, which could impact the ability of seeds to withstand stress. Expression of the *Sobic.004G182000* gene may play a role in regulating starch composition or granule structure under stress conditions in sorghum.

The fourth SNP located on chromosome 6 was linked to the *Sobic.006G114100* gene that encodes the ABCG31 protein. ABC transporters are a large family of proteins that transport various molecules, including lipids, sugars and secondary metabolites (Yazaki, 2005; 2006) across cell membranes in both prokaryotes and eukaryotes (Saurin et al., 1999). The specific function of ABCG31 in sorghum starch content is not well characterised in literature. In plants, ABC transporters have been shown to play important roles in various physiological processes, such as the transport of nutrients, defence against pathogens and regulation of hormone levels. There is some evidence that ABC transporters may be involved in the transport of sucrose, which is an important precursor for starch biosynthesis (Elferink et al., 2001; Chandravanshi et al., 2019). Previously discovered QTL in the sorghum QTL atlas (Mace et al., 2019) indicates that the ABC transporter is co-located with a starch synthesis QTL, identified using the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) sorghum mini core collection, landraces and Chinese cultivars (Chen et al., 2019). Therefore, ABCG31 may play a role in the regulation of starch biosynthesis in sorghum by transporting sucrose or other related molecules.

The fifth SNP, also located on chromosome 6, was linked to the *Sobic.006G262100* gene in sorghum that encodes for a protein that belongs to the ARF12 family. ARFs are transcription factors that regulate gene expression in response to the plant hormone auxin. Various studies indicated that ARFs play a role in plant development, including seed development and storage compound synthesis, such as starch, in seeds (Santos-Mendoza et al., 2008; Shirley et al., 2019; Mishra et al., 2022). It is therefore possible that this gene and its encoded protein, ARF12, may play a role in regulating starch metabolism in sorghum seeds. ARFs have been shown to directly or indirectly regulate the expression of genes involved in starch synthesis in several plant species. The expression of these genes is thought to be regulated by the binding of ARF transcription factors to specific regulatory elements in their promoter regions (Pratt and Zhang, 2021).

In sorghum, starch biosynthesis and metabolism are regulated by a complex network of genes and enzymes. The functions of genes and enzymes in sorghum include the regulation of starch synthesis and degradation, control of starch granule size and structure, and control of the influence of environmental and developmental factors on starch metabolism (Chen et al., 2012; Thitisaksakul et al., 2012).

Amylose

Three major novel loci were detected to be associated with amylose content. The first gene, *Sobic.001G386600*, encoded a protein within the WOX6 family this is a transcription factor involved in regulating stem cell activity and organ development in plants (Ikeda *et al*., 2009; van der Graaff et al., 2009) and can control the expression of genes involved in starch synthesis (Huang et al., 2021; Cai et al., 2023). Therefore, *Sobic.001G386600* may also play a role in regulating amylose content in sorghum.

The second gene, *Sobic.007G193900* encodes the Os08g0531200 protein, a putative alpha-glucan phosphorylase, which is involved in amylose synthesis in rice. Alpha-glucan phosphorylase is an enzyme that catalyses the addition of glucose units to the growing chain of amylose, a type of starch made up of long chains of glucose molecules (Mueller et al., 2009; Ubiparip et al., 2018). Dhaka et al. (2021) identified enzyme-encoding genes *ADP glucose phosphorylase* and *phosphorylase* involved in starch metabolism in foxtail millet (*Setaria italica*). *Phosphorylase* catalyses starch biosynthesis in rice endosperm at low temperatures (Satoh et al., 2008). Amylose is a type of starch composed of glucose units joined by alpha-1,4

glycosidic bonds. Its synthesis is catalysed by the enzyme granule-bound starch synthase (GBSS) (Nakamura, 2002; Ball and Morell, 2003). In the context of amylose synthesis in sorghum, the zinc finger CCHC domain-containing protein encoded by the *Sobic.010G137333* gene may play a role in regulating the expression of genes involved. Zinc finger proteins are known to have a diverse range of functions, including DNA binding, transcriptional regulation, RNA packaging and protein-protein interactions (Leon and Roth, 2000; Laity et al., 2001). They are also involved in a variety of biological processes, such as development, differentiation and response to environmental stress (De Storme and Geelen, 2014; Jiang et al., 2022; Wu et al., 2022).

In general, the function of specific genes and proteins involved in amylose synthesis and regulation in sorghum is an area of active research and there may be ongoing studies to investigate the role of the *Sobic.010G137333* gene and zinc finger CCHC domain containing protein in this process. Similar research findings have been reported on the synthesis of amylopectin in cereal grains (Jeon et al., 2010; Butardo et al., 2020), mostly expressed in leaves and some in the early developing endosperm.

Protein

In this study, three new protein-associated loci encoding the *Sobic.001G014900* on chromosome 1, *Sobic.02G038200* on chromosome 2, and the *Sobic.09G033500* and *Sobic.009G033400* genes on chromosome 9. Similarly, Boyles et al. (2017) identified the loci on chromosomes 1, 2 and 9 related to sorghum protein content. Ayalew et al. (2022) consistently detected two significant QTLs, one on chromosome 1 (115.2-119.2 cM) and the other on chromosome 2 (120.3-123.2 cM), across four or five of the six environments and in the combined analysis of sorghum protein content. Kamal et al. (2023) identified the loci and candidate genes located on chromosome 9 in sorghum protein content. The various studies have revealed additional candidate genes linked to protein content on chromosome 2. These consist of the putative homolog of alpha-amylase 3 (AMY3), encoded by *Sb02g023790* gene (Rhodes et al., 2017) and the candidate gene *Sobic.002G217100,* which encoded for serine/threonine-protein kinases involved in the signal cascade for nitrogen metabolism in plants (Champigny, 1995; Li et al., 2018), observed in diverse sorghum accessions.

The search for previously detected QTL in the sorghum QTL atlas (Mace et al*.,* 2019) revealed that the *Sobic.001G014900* gene, found in *Sorghum bicolor,* is co-localised with another protein content QTL identified using BTx623/Rio bi-parental lines (Murray et al., 2008). The

Sobic.001G014900 gene catalyses the formation of long-chain fatty acids, a small protein that carries the fatty acid during the biosynthesis process (Black and DiRusso, 2007; Beld et al., 2015) and are essential components of cellular membranes and energy storage molecules.

The protein encoded by the *Sobic.002G038200* gene on chromosome 2 belongs to the TatD family of hydrolases. TatD hydrolases are enzymes that hydrolyse nucleic acids and are involved in various cellular processes, including DNA repair and RNA processing (Chen et al., 2014). These enzymes are involved in the breakdown of damaged or misfolded proteins, which is important for maintaining cellular homeostasis and plays an important role in protein quality control. Hydrolase has been identified to play an important role in several processes in *Arabidopsis thaliana* including chemical defence against herbivory, lignification, hydrolysis of cell wall-derived oligosaccharides during germination and control of active phytohormone levels (Xu et al., 2004; Ahn et al., 2007). The TatD protein family is also known to be involved in the regulation of programmed cell death, which is an important process for the development and maintenance of plant tissues (Kim et al., 2008; Reape and McCabe, 2008). Thus, the TatD protein family appears to play a vital role in plant growth and development, as well as in response to environmental stresses.

The *Sobic.009G033500* and *Sobic.009G033400* genes located on chromosome 9 in sorghum encode a protein that is similar to the Class III peroxidase 70 precursor and Class III peroxidase 69 precursor, respectively. Peroxidases are enzymes that catalyse the oxidation of various substrates using hydrogen peroxide (Banci, 1997; Nicell and Wright, 1997). They play important roles in a wide range of biological processes, including plant defence, lignin biosynthesis and oxidative stress response. In sorghum, peroxidase may play a role in lignification, which is the process by which lignin is deposited in cell walls to provide structural support and rigidity to plant tissues. Lignin is a complex polymer that is synthesised from monolignols and peroxidases are involved in the oxidation of these monolignols to form the lignin polymer (Whetten and Sederoff, 1995; Liu et al., 2018). The *Sobic.009G033500* and *Sobic.009G033400* genes and their encoded enzymes may be involved in plant defence mechanisms. Peroxidases have been shown to play a role in plant defence against pathogens by generating reactive oxygen species, which can damage or kill invading microorganisms (Mehdy, 1994; Almagro et al., 2009). The *Sobic.09G033500* and *Sobic.009G033400* genes may also be involved in stress response by regulating the expression of stress-responsive genes, because peroxidase may be involved in the response to abiotic stress such as drought and high

salinity, by regulating the levels of reactive oxygen species and protecting the plant from oxidative damage.

5.6 Conclusions

Genome-wide association mapping is a powerful tool that enables the identification of genomic regions associated with important traits. In this case, Ethiopian sorghum landrace accessions were used to identify SNPs that are associated with grain quality traits and thus several tightly linked genes, which are involved in stress tolerance, transport of sugar molecules, biosynthesis, gene regulation and transcription factors.

The LEA proteins are known to play a role in stress tolerance in plants, particularly in response to drought and cold stress. The ABC transporter and sugar carrier proteins are involved in the transport of sugars across cell membranes, which is important for plant growth and development. The alpha-glucan phosphorylase and long-chain fatty acid biosynthesis enzymes are involved in the production of carbohydrates and lipids, respectively, which are important for energy storage and metabolism. The zinc finger CCHC domain and Class III peroxidase proteins are involved in regulating gene expression and plant growth, while the transcription factors ARF12 and WOX6 are involved in the regulation of plant development and growth. Overall, these genes were identified to be associated with starch, amylose and protein grain quality traits in Ethiopian sorghum landraces and provided valuable information for understanding the genetic basis of these important traits and for developing improved sorghum varieties with enhanced grain quality. This is a useful reference for understanding the molecular mechanism and helping sorghum grain quality improvement.

Recent advances in genomics, online bioinformatic data repositories, high-throughput phenotypic screening techniques such as NIR and the decreasing cost of next-generation sequencing have enabled breeders to incorporate grain quality parameters into their breeding programmes.

5.7 References

Ahn, Y.O., Zheng, M., Bevan, D.R., Esen, A., Shiu, S.H., Benson, J. and Shih, M.C. (2007). Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 35. *Phytochemistry*, 68: 1510-1520.

Alenezi, W.M., Fierheller, C.T., Serruya, C., Revil, T., Oros, K.K., Subramanian, D.N. and Tonin, P.N. (2023). Genetic analyses of DNA repair pathway associated genes implicate new candidate cancer predisposing genes in ancestrally defined ovarian cancer cases. *Frontiers in Oncology*, 13: 609.

Almagro, L., Gómez Ros, L.V., Belchi-Navarro, S., Bru, R., Ros Barceló, A. and Pedreño, M. A. (2009). Class III peroxidases in plant defense reactions. *Journal of Experimental Botany*, 60: 377-390.

Alqudah, A.M., Sallam, A., Baenziger, P.S. and Börner, A. (2020). GWAS: fast-forwarding gene identification and characterization in temperate cereals: lessons from barley - a review. *Journal of Advanced Research*, 22: 119-135.

Bahaji, A., Li, J., Sánchez-López, Á.M., Baroja-Fernández, E., Muñoz, F.J., Ovecka, M. and Pozueta-Romero, J. (2014). Starch biosynthesis, its regulation and biotechnological approaches to improve crop yields. *Biotechnology Advances*, 32: 87-106.

Ball, S.G. and Morell, M.K. (2003). From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. *Annual Review of Plant Biology*, 54: 207-233.

Banci, L. (1997). Structural properties of peroxidases. *Journal of Biotechnology*, 53: 253-263.

Bazié, D., Dibala, C.I., Kondombo, C.P., Diao, M., Konaté, K., Sama, H. and Dicko, M.H. (2023). Physicochemical and nutritional potential of fifteen sorghum cultivars from Burkina Faso. *Agriculture*, 13: 675.

Beld, J., Lee, D.J. and Burkart, M.D. (2015). Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Molecular BioSystems*, 11: 38-59.

Black, P.N. and DiRusso, C.C. (2007). Yeast acyl-CoA synthetases at the crossroads of fatty acid metabolism and regulation. *Biochimicaet Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1771: 286-298.

Boyles, R.E., Pfeiffer, B.K., Cooper, E.A., Rauh, B.L., Zielinski, K.J., Myers, M.T. and Kresovich, S. (2017). Genetic dissection of sorghum grain quality traits using diverse and segregating populations. *Theoretical and Applied Genetics*, 130: 697-716.

Butardo, V.M., Luo, J., Li, Z., Gidley, M.J., Bird, A.R., Tetlow, I.J. and Rahman, S. (2020). Functional genomic validation of the roles of soluble starch synthase IIa in Japonica rice endosperm. *Frontiers in Genetics*, 11: 289.

Butler, D.G., Cullis, B.R., Gilmour, A.R., Gogel, B.J. and Thompson, R. (2017). ASReml-R /reference manual version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK, pp. 30-49.

Cai, Y., Huang, L., Song, Y., Yuan, Y., Xu, S., Wang, X. and Wang, Y. (2023). LAZY3 interacts with LAZY2 to regulate tiller angle by modulating shoot gravity perception in rice. *Plant Biotechnology Journal,* 21: 1217-1228.

Champigny, M.L. (1995). Integration of photosynthetic carbon and nitrogen metabolism in higher plants. *Photosynthesis Research*, 46: 117-127.

Chandravanshi, M., Sharma, A., Dasgupta, P., Mandal, S.K. and Kanaujia, S.P. (2019). Identification and characterization of ABC transporters for carbohydrate uptake in *Thermus thermophilus* HB8. *Gene*, 696: 135-148.

Chen, B.R., Wang, C.Y., Ping, W., Zhu, Z.X., Ning, X.U., Shi, G.S. and Huang, R.D. (2019). Genome-wide association study for starch content and constitution in sorghum (*Sorghum bicolor* (L.) Moench). *Journal of Integrative Agriculture*, 18: 2446-2456.

Chen, Y.C., Li, C.L., Hsiao, Y.Y., Duh, Y. and Yuan, H.S. (2014). Structure and function of TatD exonuclease in DNA repair. *Nucleic Acids Research*, 42: 10776-10785.

Chen, Y., Wang, M. and Ouwerkerk, P.B. (2012). Molecular and environmental factors determining grain quality in rice. *Food and Energy Security*, 1: 111-132.

Chen, Z.Y., Brown, R.L., Damann, K.E. and Cleveland, T.E. (2002). Identification of unique or elevated levels of kernel proteins in aflatoxin-resistant maize genotypes through proteome analysis. *Phytopathology*, 92: 1084-1094.

Chen, Z.Y., Brown, R.L., Damann, K.E. and Cleveland, T.E. (2007). Identification of maize kernel endosperm proteins associated with resistance to aflatoxin contamination by *Aspergillus flavus*. *Phytopathology*, 97: 1094-1103.

De Storme, N. and Geelen, D. (2014). The impact of environmental stress on male reproductive development in plants: biological processes and molecular mechanisms. *Plant, Cell and Environment*, 37: 1-18.

Dhaka, A., Muthamilarasan, M. and Prasad, M. (2021). A comprehensive study on core enzymes involved in starch metabolism in the model nutricereal, foxtail millet (*Setaria italica* L.). *Journal of Cereal Science*, 97: 103153.

Dupont, F.M. (2008). Metabolic pathways of the wheat (*Triticum aestivum*) endosperm amyloplast revealed by proteomics. *BMC Plant Biology*, 8: 1-18.

Elferink, M.G., Albers, S.V., Konings, W.N. and Driessen, A.J. (2001). Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. *Molecular Microbiology*, 39: 1494-1503.

Enyew, M., Feyissa, T., Carlsson, A.S., Tesfaye, K., Hammenhag, C., Seyoum, A. and Geleta, M. (2022). Genome-wide analyses using multi-locus models revealed marker-trait associations for major agronomic traits in *Sorghum bicolor*. *Frontiers in Plant Science*, 13: 999692.

Ganal, M.W., Altmann, T. and Röder, M.S. (2009). SNP identification in crop plants. *Current Opinion in Plant Biology*, 12: 211-217.

Girma, G., Nida, H., Seyoum, A., Mekonen, M., Nega, A., Lule, D. and Mengiste, T. (2019). A large-scale genome-wide association analysis of Ethiopian sorghum landrace collection reveal loci associated with important traits. *Frontiers in Plant Science*, 10: 691.

Glaubitz, J.C., Casstevens, T.M., Lu, F., Harriman, J., Elshire, R.J., Sun, Q. and Buckler, E.S. (2014). TASSEL-GBS: a high-capacity genotyping by sequencing analysis pipeline. *PloS One*, 9: e90346.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar D.S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, 40: D1178-D1186.

Hassanyar, A.K., Nie, H., Li, Z., Lin, Y., Huang, J., Woldegiorgis, S.T. and Su, S. (2023). Discovery of SNP molecular markers and candidate genes associated with Sacbrood virus resistance in *Apis ceranacerana* larvae by whole-genome resequencing. *International Journal of Molecular Sciences*, 24*:* 6238.

Hossain, M.S., Islam, M.N., Rahman, M.M., Mostofa, M.G. and Khan, M.A.R. (2022). Sorghum: A prospective crop for climatic vulnerability, food and nutritional security. *Journal of Agriculture and Food Research*, 8: 100300.

Huang, L., Wang, W., Zhang, N., Cai, Y., Liang, Y., Meng, X. and Wang, Y. (2021). LAZY2 controls rice tiller angle through regulating starch biosynthesis in gravity-sensing cells. *New Phytologist*, 231: 1073-1087.

Huang, M., Liu, X., Zhou, Y., Summers, R.M. and Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience*, 8: 154.

Huang, Y., Li, L., Smith, K.P. and Muehlbauer, G.J. (2016). Differential transcriptomic responses to *Fusarium graminearum* infection in two barley quantitative trait loci associated with Fusarium head blight resistance. *BMC Genomics*, 17: 1-16.

Hu, Z., Olatoye, M.O., Marla, S. and Morris, G.P. (2019). An integrated genotyping-by-sequencing polymorphism map for over 10,000 sorghum genotypes. *Plant Genome*, 12: 180044.

Ikeda, M., Mitsuda, N. and Ohme-Takagi, M. (2009). Arabidopsis WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *The Plant Cell*, 21: 3493-3505.

Jeon, J.S., Ryoo, N., Hahn, T.R., Walia, H. and Nakamura, Y. (2010). Starch biosynthesis in cereal endosperm. *Plant Physiology and Biochemistry*, 48: 383-392.

Jiang, Y., Liu, L., Pan, Z., Zhao, M., Zhu, L., Han, Y. and Zhang, M. (2022). Genome-wide analysis of the C2H2 zinc finger protein gene family and its response to salt stress in ginseng, *Panax ginseng* Meyer. *Scientific Reports*, 12: 10165.

Kaler, A.S., Gillman, J.D., Beissinger, T. and Purcell, L.C. (2020). Comparing different statistical models and multiple testing corrections for association mapping in soybean and maize. *Frontiers in Plant Science*, 10: 1794.

Kamal, N.M., Gorafi, Y.S.A., Tomemori, H., Kim, J.S., Elhadi, G.M.I. and Tsujimoto, H. (2023). Genetic variation for grain nutritional profile and yield potential in sorghum and the possibility of selection for drought tolerance under irrigated conditions. *BMC Genomics*, 24: 515.

Kim, K.S., Min, J.Y. and Dickman, M.B. (2008). Oxalic acid is an elicitor of plant programmed cell death during *Sclerotinia sclerotiorum* disease development. *Molecular Plant-Microbe Interactions*, 21: 605-612.

Laity, J.H., Lee, B.M. and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Current Opinion in Structural Biology*, 11: 39-46.

Leon, O. and Roth, M. (2000). Zinc fingers: DNA binding and protein-protein interactions. *Biological Research*, 33: 21-30.

Li, A., Jia, S., Yobi, A., Ge, Z., Sato, S.J., Zhang, C. and Holding, D.R. (2018). Editing of an alpha-kafirin gene family increases, digestibility and protein quality in sorghum. *Plant Physiology*, 177: 1425-1438.

Li, P., Wang, L., Liu, H. and Yuan, M. (2022). Impaired SWEET-mediated sugar transportation impacts starch metabolism in developing rice seeds. *The Crop Journal*, 10: 98-108.

Li, J., Tang, W., Zhang, Y.W., Chen, K.N., Wang, C., Liu, Y., Zhan, Q., Wang, C., Wang, S.B., Xie, S.Q. and Wang, L. (2018). Genome-wide association studies for five forage quality-related traits in sorghum (*Sorghum bicolor)*. *Frontiers in Plant Science*, 9: 1146.

Liu, J., Li, L., Foroud, N. A., Gong, X., Li, C. and Li, T. (2019). Proteomics of bulked rachides combined with documented QTL uncovers genotype nonspecific players of the Fusarium head blight responses in wheat. *Phytopathology*, 109: 111-119.

Liu, Q., Luo, L. and Zheng, L. (2018). Lignins: biosynthesis and biological functions in plants. *International Journal of Molecular Sciences*, 19: 335.

Liu, X., Huang, M., Fan, B., Buckler, E.S. and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genetics*, 12: e1005767.

Liu, Y., Wang, L., Xing, X., Sun, L., Pan, J., Kong, X. and Li, D. (2013). ZmLEA3, a multifunctional group 3 LEA protein from maize (*Zea mays* L.), is involved in biotic and abiotic stresses. *Plant and Cell Physiology*, 54: 944-959.

Mace, E., Innes, D., Hunt, C., Wang, X., Tao, Y., Baxter, J. and Jordan, D. (2019). The sorghum QTL Atlas: a powerful tool for trait dissection, comparative genomics, and crop improvement. *Theoretical and Applied Genetics*, 132: 751-766.

Maina, F., Harou, A., Hamidou, F. and Morris, G.P. (2022). Genome-wide association studies identify putative pleiotropic locus mediating drought tolerance in sorghum. *Plant Direct*, 6: 413.

Malik, P.L., Janss, L., Nielsen, L.K., Borum, F., Jørgensen, H., Eriksen, B. and Rasmussen, S. K. (2019). Breeding for dual-purpose wheat varieties using marker–trait associations for biomass yield and quality traits. *Theoretical and Applied Genetics*, 132: 3375-3398.

Massimiliano, G., Zirilli, A., Cozzucoli, P.C. and Angela, A. (2018). Bonferroni-Holm and permutation tests to compare health data: methodological and applicative issues. *BMC Medical Research Methodology*, 18:1-9.

Mehdy, M.C. (1994). Active oxygen species in plant defence against pathogens. *Plant Physiology*, 105: 467.

Messmer, R., Fracheboud, Y., Bänziger, M., Vargas, M., Stamp, P. and Ribaut, J.M. (2009). Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. *Theoretical and Applied Genetics*, 119: 913-930.

Mishra, B.S., Sharma, M. and Laxmi, A. (2022). Role of sugar and auxin crosstalk in plant growth and development. *Physiologia Plantarum*, 174: e13546.

Mofokeng, A.M., Shimelis, H. and Laing, M. (2017). Breeding strategies to improve sorghum quality. *Australian Journal of Crop Science*, 11: 142-148.

Motlhaodi, T., Bryngelsson, T., Chite, S., Fatih, M., Ortiz, R. and Geleta, M. (2018). Nutritional variation in sorghum [*Sorghum bicolor* (L.) Moench] accessions from southern Africa revealed by protein and mineral composition. *Journal of Cereal Science*, 83: 123-129.

Mueller, M., Takemasa, R., Schwarz, A., Atomi, H. and Nidetzky, B. (2009). "Short-chain" α-1, 4-glucan phosphorylase having a truncated N-terminal domain: functional expression and characterization of the enzyme from *Sulfolobus solfataricus*. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1794: 1709-1714.

Mundada, P.S., Kadam, S.B., Pable, A.A. and Barvkar, V.T. (2022). Recent Advances and Applicability of GBS, GWAS and GS in Millet Crops. In: Sonah, H., Goyal, V., Shivaraj, S.M., Deshmukh, R.K. (eds). Genotyping by Sequencing for Crop Improvement. John Wiley & Sons, Ltd. pp. 270-294.

Murray, S.C., Sharma, A., Rooney, W.L., Klein, P.E., Mullet, J.E., Mitchell, S.E. and Kresovich, S. (2008). Genetic improvement of sorghum as a biofuel feedstock: I. QTL for stem sugar and grain nonstructural carbohydrates. *Crop Science*, 48: 2165-2179.

Nakamura, Y. (2002). Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue. *Plant and Cell Physiology*, 43: 718-725.

Nicell, J.A. and Wright, H. (1997). A model of peroxidase activity with inhibition by hydrogen peroxide. *Enzyme and Microbial Technology*, 21: 302-310.

Nida, H., Girma, G., Mekonen, M., Lee, S., Seyoum, A., Dessalegn, K. and Mengiste, T. (2019). Identification of sorghum grain mold resistance loci through genome wide association mapping. *Journal of Cereal Science*, 85: 295-304.

Nida, H., Girma, G., Mekonen, M., Tirfessa, A., Seyoum, A., Bejiga, T. and Mengiste, T. (2021). Genome-wide association analysis reveals seed protein loci as determinants of variations in grain mold resistance in sorghum. *Theoretical and Applied Genetics*, 134: 1167-1184

Pratt, I.S. and Zhang, B. (2021). Genome-wide identification of ARF transcription factor gene family and their expression analysis in sweet potato. *International Journal of Molecular Sciences*, 22: 9391.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38: 904-909.

Ragaee, S., Abdel-Aal, E.S.M. and Noaman, M. (2006). Antioxidant activity and nutrient composition of selected cereals for food use. *Food Chemistry*, 98: 32-38.

Reape, T.J. and McCabe, P.F. (2008). Apoptotic-like programmed cell death in plants. *New Phytologist*, 180: 13-26.

Remington, D.L., Thornsberry, J.M., Matsuoka, Y., Wilson, L.M., Whitt, S.R., Doebley, J., Kresovich, S., Goodman, M.M. and Buckler IV, E.S. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences*, 98: 11479-11484.

Rhodes, D.H., Hoffmann Jr, L., Rooney, W.L., Ramu, P., Morris, G.P. and Kresovich, S. (2014). Genome-wide association study of grain polyphenol concentrations in global sorghum [*Sorghum bicolor* (L.) Moench] germplasm. *Journal of Agricultural and Food Chemistry*, 62: 10916-10927.

Rhodes, D.H., Hoffmann, L., Rooney, W.L., Herald, T.J., Bean, S., Boyles, R., Brenton, Z.W. and Kresovich, S. (2017). Genetic architecture of kernel composition in global sorghum germplasm. *BMC Genomics,* 18:1-8.

Saini, D.K., Chopra, Y., Singh, J., Sandhu, K.S., Kumar, A., Bazzer, S. and Srivastava, P. (2022). Comprehensive evaluation of mapping complex traits in wheat using genome-wide association studies. *Molecular Breeding*, 42:1-52.

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M. and Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. *The Plant Journal*, 54: 608-620.

Satoh, H., Shibahara, K., Tokunaga, T., Nishi, A., Tasaki, M., Hwang, S.K. and Nakamura, Y. (2008). Mutation of the plastidial α-glucan phosphorylase gene in rice affects the synthesis and structure of starch in the endosperm. *The Plant Cell*, 20: 1833-1849.

Saurin, W., Hofnung, M. and Dassa, E. (1999). Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *Journal of Molecular Evolution*, 48: 22-41.

Seung, D. (2020). Amylose in starch: towards an understanding of biosynthesis, structure and function. *New Phytologist*, 228: 1490-1504.

Shailasree, S., Kini, K.R., Deepak, S., Kumudini, B.S. and Shetty, H.S. (2004). Accumulation of hydroxyproline-rich glycoproteins in pearl millet seedlings in response to *Sclerospora graminicola* infection. *Plant Science*, 167: 1227-1234.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (1997). Gene expression and signal transduction in water-stress response. *Plant Physiology*, 115: 327.

Shirley, N.J., Aubert, M.K., Wilkinson, L.G., Bird, D.C., Lora, J., Yang, X. and Tucker, M.R. (2019). Translating auxin responses into ovules, seeds and yield: Insight from Arabidopsis and the cereals. *Journal of Integrative Plant Biology*, 61: 310-336.

Showalter, A.M. (1993). Structure and function of plant cell wall proteins. *The Plant Cell*, 5: 9.

Smith, A.B., Borg, L.M., Gogel, B.J. and Cullis, B.R. (2019). Estimation of factor analytic mixed models for the analysis of multi-treatment multi-environment trial data. *Journal of Agricultural, Biological and Environmental Statistics*, 24: 573-588.

Taylor, J.R. and Emmambux, M.N. (2010). Developments in our understanding of sorghum polysaccharides and their health benefits. *Cereal Chemistry*, 87: 263-271.

Teng, Z., Chen, Y., Yuan, Y., Peng, Y., Yi, Y., Yu, H. and Ye, N. (2022). Identification of microRNAs regulating grain filling of rice inferior spikelets in response to moderate soil drying post-anthesis. *The Crop Journal*, 10: 962-971.

Thitisaksakul, M., Jiménez, R.C., Arias, M.C. and Beckles, D.M. (2012). Effects of environmental factors on cereal starch biosynthesis and composition. *Journal of Cereal Science*, 56: 67-80.

Thornsberry, J., and Buckler, E. (2003). Structure of linkage disequilibrium in plants. *Annual Revive Plant Biology*, 54: 357-374.

Ubiparip, Z., Beerens, K., Franceus, J., Vercauteren, R. and Desmet, T. (2018). Thermostable alpha-glucan phosphorylases: characteristics and industrial applications. *Applied Microbiology and Biotechnology*, 102: 8187-8202.

van der Graaff, E., Laux, T. and Rensing, S.A. (2009). The WUS homeobox-containing (WOX) protein family. *Genome Biology*, 10: 1-9.

Wang, J. and Zhang, Z. (2021). GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genomics, Proteomics and Bioinformatics*, 19: 629-640.

Wang, Y.H., Upadhyaya, H.D., Burrell, A.M., Sahraeian, S.M.E., Klein, R.R., and Klein, P.E. (2013). Genetic structure and linkage disequilibrium in a diverse, representative collection of the C4 model plant, *sorghum bicolor*. *G3: Genes Genomes Genetics*, 3: 783-793.

Weine, E., McPeek, M.S. and Abney, M. (2021). Application of Equal Local Levels to Improve QQ Plot Testing Bands with R Package qqconf. *Journal of Statistical Software,* 106: 10

Whetten, R. and Sederoff, R. (1995). Lignin biosynthesis. *The Plant Cell*, 7: 1001.

Wondimu, Z., Dong, H., Paterson, A.H., Worku, W. and Bantte, K. (2023). Genome-wide association study reveals genomic loci influencing agronomic traits in Ethiopian sorghum (*Sorghum bicolor* (L.) Moench) landraces. *Molecular Breeding*, 43: 1-15.

Wu, Z., Shen, S., Wang, Y., Tao, W., Zhao, Z., Hu, X. and Yu, P. (2022). Genome-wide identification and expression analysis of the Zinc Finger Protein gene subfamilies under drought stress in *Triticum aestivum*. *Plants*, 11: 2511.

Xiong, Y., Zhang, P., Warner, R.D. and Fang, Z. (2019). Sorghum grain: From genotype, nutrition and phenolic profile to its health benefits and food applications. *Comprehensive Reviews in Food Science and Food Safety*, 18: 2025-2046.

Xu, W., Jian, S., Li, J., Wang, Y., Zhang, M. andXia, K. (2023). Genomic identification of CCCH-Type zinc finger protein genes reveals the role of HuTZF3 in tolerance of heat and salt stress of Pitaya (*Hylocereus polyrhizus*). *International Journal of Molecular Sciences*, 24: 6359.

Xu, Z., Escamilla-Treviño, L., Zeng, L., Lalgondar, M., Bevan, D., Winkel, B. and Esen, A. (2004). Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 1. *Plant Molecular Biology*, 55: 343-367.

Yang, J., Jiang, H., Yeh, C.T., Yu, J., Jeddeloh, J.A., Nettleton, D. and Schnable, P.S. (2015). Extreme-phenotype genome-wide association study (XP-GWAS): a method for identifying trait-associated variants by sequencing pools of individuals selected from a diversity panel. *The Plant Journal*, 84: 587-596.

Yazaki, K. (2005). Transporters of secondary metabolites. *Current Opinion in Plant Biology*, 8: 301-307.

Yazaki, K. (2006). ABC transporters are involved in the transport of plant secondary metabolites. *FEBS Letters*, 580: 1183-1191.

Young, V.R. and Pellett, P.L. (1994). Plant proteins concerning human protein and amino acid nutrition. *The American Journal of Clinical Nutrition*, 59: 1203S-1212S.

Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F. and Buckler, E.S. (2006). A unified mixed-model method for association mapping that accounts formultiple levels of relatedness. *Nature Genetics*, 38: 203-208.

Zhang, Q., Zhang, J., Wei, F., Fu, X., Wei, H., Lu, J. and Wang, H. (2023). The CCCH-Type Zinc Finger Protein GhC3H20 enhances salt stress tolerance in *Arabidopsis thaliana* and cotton through ABA Signal Transduction Pathway. *International Journal of Molecular Sciences*, 24: 5057.

Zhang, Z., Ersoz, E., Lai, C.Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A. and Buckler, E.S. (2010). The mixed linear model approach is adapted for genome-wide association studies. *Nature Genetics*, 42: 355-360.

Zhao, Q., Shi, X.S., Wang, T., Chen, Y., Yang, R., Mi, J. and Zhang, Y.M. (2023). Identification of QTNs, QTN-by-environment interactions and their candidate genes for grain size traits in main crop and ratoon rice. *Frontiers in Plant Science*, 14: 1119218.

Zheng, L.Y., Guo, X.S., He, B., Sun, L.J., Peng, Y., Dong, S.S. and Jing, H.C. (2011). Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*). *Genome Biology*, 12: 1-15.

CHAPTER 6

GENOME-WIDE ASSOCIATION STUDY OF MINERAL COMPOSITION IN ETHIOPIAN SORGHUM LANDRACES

6.1 Abstract

Sorghum is a cereal grain that is widely consumed in many parts of the world and is known for its nutritional value due to being rich in minerals such as P, K and Mg. Deficiencies in micronutrients in sorghum, including ash, Fe and Zn and this can lead to poor nutritional quality for animals and human consumption, causing detrimental health effects. GWAS using 365 diverse Ethiopian sorghum landrace accessions identified SNPs significantly linked to ash, Fe and Zn contents. Different SNPs linked to ash, Fe and Zn content were detected in different environments. A higher number of SNPs linked to Fe and Zn content were detected in Jimma and Melkassa compared to the other test environment, Miesso. Candidate genes with significant associations with mineral traits were identified. Candidate genes *Sobic.002G266800*, *Sobic.007G193500* and *Sobic.008G068300* are tightly linked with the significant SNPs detected for ash, an important indicator of mineral composition in sorghum, as it is primarily composed of minerals such as Ca, K and Mg. *Sobic.002G266800* plays a role in the biosynthesis of flavonoid glycosides in sorghum, which in turn contributes to the plant's pigmentation and defence strategies. *Sobic.007G193500* is annotated as teosinte glume architecture1 (tga1) and is involved in the development of teosinte glumes, which influences the structure, composition and properties of glumes, as well as the processing and nutritional quality of the grain. The *Sobic.010G218500* gene, significantly linked to Fe content, encoded a RISBZ5 transcription factor protein belonging to the basic leucine zipper (bZIP) family of proteins. In plants, bZIP transcription factors are involved in various stress responses. GWAS using multiple models, detected genes across environments linked to Zn content encoded proteins from the zinc finger protein (ZNF) family, which are involved in Zn ion uptake and transport.

6.2 Introduction

Sorghum is a cereal crop grown in tropical, subtropical and temperate climates. Starch, protein, micronutrients and crude fibre are abundant in sorghum grain. Sorghum is a staple crop for most people in sub-Saharan Africa, due to being one of the least expensive energy and micronutrient sources (Parthasarathy Rao et al., 2006). Fe and Zn are essential micronutrients for human health and deficiencies in these minerals can lead to serious health problems. Zn is required for effective immune function, wound healing and growth and development (Bhowmik et al., 2010; Roohani et al., 2013), while Fe is necessary for the formation of haemoglobin in red blood cells, which carries oxygen throughout the body (Gupta, 2014). In sorghum, ash content is an important indicator of mineral composition, as it is composed primarily of minerals such as Ca, K and Mg (Pereira et al., 2012).

GWAS can be used to investigate the genetic basis of various traits in sorghum, including ash content and mineral compositions such as Fe and Zn. Several genomic regions have been identified to be associated with ash content and Fe and Zn composition; in sorghum (Kotla et al., 2019) and maize (Hindu et al., 2018). These regions contain genes that are involved in the transport and metabolism of minerals in plants. Studies have furthermore shown that there is genetic variation in ash content and mineral composition in sorghum (Paiva et al., 2017; Motlhaodi et al., 2018).

Sorghum is a staple food in Ethiopia and improving its Fe and Zn content through breeding could have important health benefits for the population. Understanding the genetic basis of ash content and mineral composition could help breeders develop sorghum varieties with improved nutritional quality. It could also provide insights into the molecular mechanisms underlying the uptake and transport of minerals in plants. A GWAS study of Ethiopian sorghum landrace accessions could potentially identify genetic variants associated with ash content and Fe and Zn composition in the grain. The objective of this study was to identify genomic regions associated with ash, Fe and Zn content, indicating a genetic basis for mineral composition of Ethiopian sorghum landraces.

6.3 Materials and methods

6.3.1 Plant material and design

Sorghum landraces and improved varieties used in this study, as well as the trial design followed, were as described in section 3.3.1. Planting was done across three sites (Jimma, Melkassa and Miesso) as described in section 4.3.1.

6.3.2 Sorghum mineral composition analyses

The 200 grams of samples were collected from 365 sorghum accessions in six environments for grain quality analysis as described in section 4.3.3.

6.3.3 Genotyping-by-sequencing

Genotyping-by-sequencing was done as described in section 5.3.3.

6.3.4 Data analysis

Data analysis (ANOVA, spatial analysis, BLUP, identification of gene regions and gene annotation) was done as described in section 5.3.4.

6.4 Results

6.4.1 Phenotypic variation

Data obtained in Chapter 4 indicated that Ethiopian sorghum landrace showed high variability in mineral content (ash, Fe and Zn) across environments and this data was used for GWAS in the current chapter.

6.4.2 The quantile-quantile plot analysis

BLUP analysis of ash, Fe and Zn of each sorghum landrace accessions across six environments was used for this study. The QQ plots were used to determine false associations (Figure 6.1). Results indicated that the GLM, MLM and CMLM models generated false positives and negatives. Complex and multi-locus models (BLINK and FarmCPU), expected to control false positives due to population structure and family relatedness, identified highly significant SNPs after Holm-Bonferroni (HB p-value ≤ 0.05) adjustments across traits as described in section 5.4.2 and 5.5.1. These complex models reduced p-value inflations, resulting in controlled false positive and negative error rates. Multiple significant SNP marker associations were identified on the same chromosome by the BLINK and FarmCPU models.

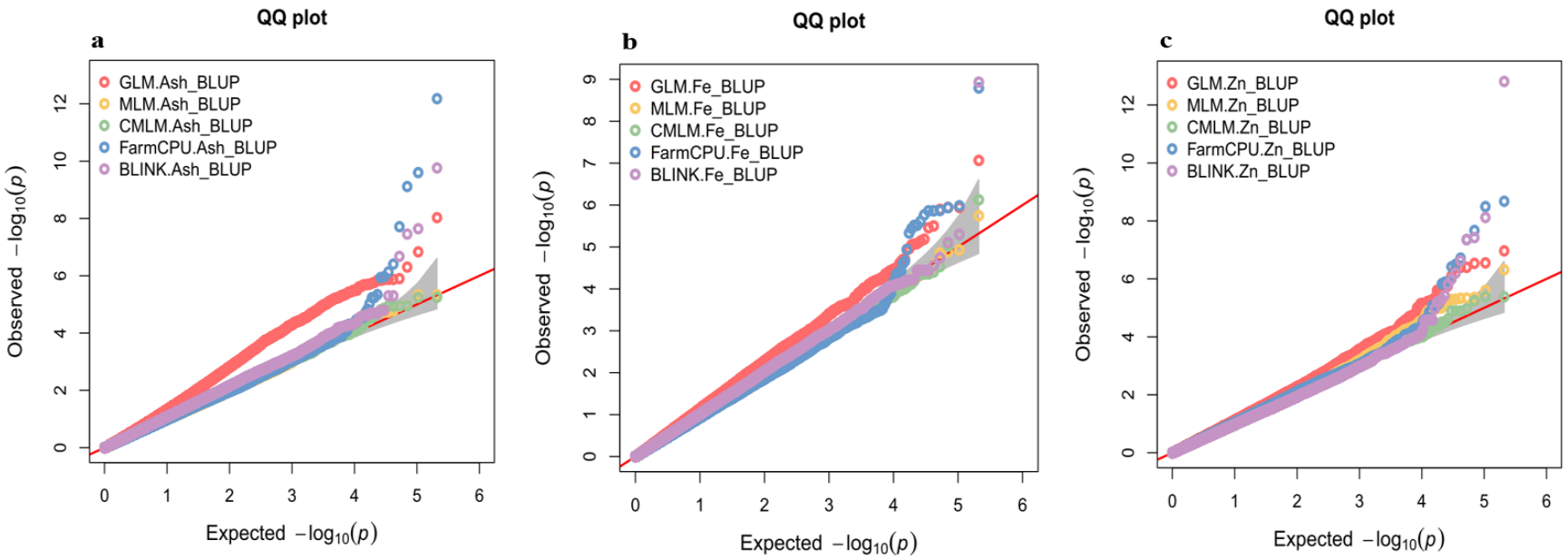


Figure 6.1 Genome-wide association analysis of grain mineral composition across environments using different statistical models, best linear unbiased prediction analysis and quantile-quantile (QQ) plots for ash (a), Zn (b) and Fe (c)

GLM = general linear model; MLM = mixed linear model; CMLM = compressed mixed linear model; FarmCPU = fixed and random model circulating probability unification; BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; BLUP= best linear unbiased prediction.

6.4.3 Genome-wide association studies across environments

Major QTL regions within 100 kb upstream and downstream of QTL peaks have been investigated to identify candidate genes associated with the variations. The GWAS analysis was carried out using the GLM, MLM, CMLM, FarmCPU and BLINK models, but the most significant SNPs associated with ash, Fe and Zn in individual environments were detected using the FarmCPU and BLINK models. The GWAS study identified 121 different SNPs significantly associated with ash (49), Fe (36) and Zn (36) after adjustment (HB p-value ≤ 0.05) using the two models (Appendix 5). Because of the highly significant interaction between genotypes and environments, GWAS was conducted separately for each environment (Tables 6.1 and 6.2).

In BLUP analysis, significant SNPs were identified that were linked; 12 with ash, two with Fe and 18 with Zn content (Table 6.3). Among these significant SNPs, three linked to ash and four with Zn that were commonly and consistently detected by both BLINK and FarmCPU models in BLUP analysis; however, SNP S05_65910190 was detected by BLINK and FarmCPU models while other SNP S10_56127541 was only detected by FarmCPU in Fe content (Figures 6.3 to 6.5 and Tables 6.4); these SNPs were performed in individual environments (Table 6.2).

For ash content, Melkassa had the highest number of significant SNPs (14) for ash in 2021, while no significant SNPs were detected at Miesso in 2020. The phenotypic variation (R^2) in ash ranged from 5 to 50. For Fe, the highest number of significant SNPs (13) were identified at Melkassa in 2020, while only one significant SNP was detected at Miesso in 2020. The phenotypic variation (R^2) for Fe ranged from 6 to 50. For Zn content, the highest number of significant SNPs (19) were detected at Melkassa for both seasons, while no significant SNP was detected at Miesso in 2020. The phenotypic variation (R^2) for Zn content ranged from 5 to 50 across environments (Table 6.1)

The highest number of highly significant and consistent SNPs associated with ash, Fe and Zn content were detected at Melkassa and Jimma for both seasons. Three SNPs, S02_65113562, S07_62611661 and S08_7974942, located on chromosomes 2, 7 and 10, respectively, were significantly associated with ash content. In Miesso, no SNPs were detected linked with the ash content. SNP S02_65113562 was detected at Jimma and Melkassa over the two years, while SNP S08_7974942 was detected at both locations in 2021.

Two SNPs, S05_65910190 and S10_56127541 were linked to Fe content; S05_65910190 was detected in all locations and over both years, except at Jimma in 2020, while SNP S10_56127541 was detected at Jimma in 2020 and Miesso in 2021.

Four SNPs, S01_30708847 (chromosome 1), S03_2052347 (chromosome 3), S08_51315455 (chromosome 8) and S08_9366659 (chromosome 8), were significantly associated with Zn content and were detected consistently across environments by both BLINK and FarmCPU models. SNP S08_51315455 was detected only at Jimma and SNP S01_30708847 was detected only at Melkassa in 2020. SNP S08_9366659 was detected at Jimma and Melkassa over the two years, while SNP S03_2052347 was detected at Melkassa over the two years and at Miesso in 2021; SNP S01_30708847 was detected only at Melkassa in 2020 for Zn content (Table 6.2).

Table 6.1 Summary of genome-wide association study data obtained for sorghum ash, Fe and Zn content including number of significant single nucleotide polymorphisms, *p*-value range and range of phenotypic variation explained

Trait	Envt	Number of SNPs (HB *p*-value ≤ 0.05)	*p*-value		Range of R^2	
			Min	Max	Min R^2	Max R^2
Ash	JM20	13	7.86E-10	4.51E-07	5	30
	JM21	12	6.16E-09	1.59E-06	10	30
	MK20	2	1.83E-09	3.08E-09	18	27
	MK21	14	2.64E-15	2.53E-06	23.5	25.5
	MS20	0				
	MS21	11	2.42E-09	2.32E-06	50	50
Fe	JM20	12	5.03E-17	1.38E-06	13	20
	JM21	2	9.40E-08	4.20E-07	50	50
	MK20	13	6.87E-12	3.01E-06	22	34
	MK21	10	1.86E-08	1.65E-06	16	26
	MS20	1	1.60E-08	1.60E-08	50	50
	MS21	11	1.76E-12	1.56E-06	6	20
Zn	JM20	10	7.86E-10	4.51E-07	5	30
	JM21	14	6.16E-09	1.59E-06	10	30
	MK20	19	1.83E-09	3.08E-09	18	27
	MK21	19	2.64E-15	2.53E-06	23.5	25.5
	MS20	0				
	MS21	4	2.42E-09	2.32E-06	50	50

JM20 and JM21 = Jimma 2020 and 2021; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021; Max = Maximum, Min = Minimum; Envt = Environment; R^2 = phenotypic variation; HB= Holm-Bonferroni with adjustment.

Table 6.2 Summary of single nucleotide polymorphisms consistently and significantly associated with micronutrient composition in sorghum landraces across years and locations

Trait	Year	Location	Model	SNP	Chr	Pos	*P*-value	MAF	HB *p*-value ≤ 0.05	Effect	Associated gene
Ash	2020	Jimma	GLM	S02_65113562	2	65113562	1.20E-08	0.26	0.00	0.16	*Sobic.002G266800*
Ash	2021	Jimma	BLINK	S02_65113562	2	65113562	1.42E-07	0.26	0.01	0.09	*Sobic.002G266800*
Ash	2020	Melkassa	BLINK	S02_65113562	2	65113562	3.08E-09	0.26	0.00	0.1	*Sobic.002G266800*
Ash	2021	Melkassa	BLINK	S02_65113562	2	65113562	5.23E-09	0.26	0.00	0.1	*Sobic.002G266800*
Ash	2021	Melkassa	FarmCPU	S02_65113562	2	65113562	1.26E-06	0.26	0.02	0.06	*Sobic.002G266800*
Ash	2020	Jimma	FarmCPU	S07_62611661	7	62611661	1.93E-10	0.33	0.00	-0.1	*Sobic.007G193500*
Ash	2021	Jimma	BLINK	S08_7974942	8	7974942	6.16E-09	0.41	0.00	-0.13	*Sobic.008G068300*
Ash	2021	Jimma	FarmCPU	S08_7974942	8	7974942	5.65E-07	0.41	0.02	-0.08	*Sobic.008G068300*
Ash	2021	Melkassa	BLINK	S08_7974942	8	7974942	2.45E-07	0.41	0.01	-0.11	*Sobic.008G068300*
Ash	2021	Melkassa	FarmCPU	S08_7974942	8	7974942	1.52E-14	0.41	0.00	-0.11	*Sobic.008G068300*
Fe	2021	Jimma	BLINK	S05_65910190	5	65910190	9.40E-08	0.24	0.02	3.36	*Sobic.005G177500*
Fe	2020	Melkassa	BLINK	S05_65910190	5	65910190	7.34E-09	0.24	0.00	5.92	*Sobic.005G177500*
Fe	2020	Melkassa	FarmCPU	S05_65910190	5	65910190	1.57E-09	0.24	0.00	4.06	*Sobic.005G177500*
Fe	2021	Melkassa	BLINK	S05_65910190	5	65910190	9.33E-10	0.24	0.00	4.78	*Sobic.005G177500*
Fe	2021	Melkassa	FarmCPU	S05_65910190	5	65910190	1.86E-08	0.24	0.00	3.43	*Sobic.005G177500*
Fe	2020	Miesso	BLINK	S05_65910190	5	65910190	1.60E-08	0.24	0.00	4.9	*Sobic.005G177500*
Fe	2021	Miesso	BLINK	S05_65910190	5	65910190	5.10E-08	0.24	0.00	4.46	*Sobic.005G177500*
Fe	2020	Jimma	BLINK	S10_56127541	10	56127541	3.11E-09	0.15	0.00	3.97	*Sobic.010G218500*
Fe	2020	Jimma	FarmCPU	S10_56127541	10	56127541	1.47E-07	0.15	0.01	3.28	*Sobic.010G218500*
Fe	2021	Miesso	BLINK	S10_56127541	10	56127541	1.37E-08	0.15	0.00	4.55	*Sobic.010G218500*
Fe	2021	Miesso	FarmCPU	S10_56127541	10	56127541	1.76E-12	0.15	0.00	4.81	*Sobic.010G218500*

Table 6.2 Summary of single nucleotide polymorphisms consistently and significantly associated with micronutrient composition in sorghum landraces across years and locations (continued)

Trait	Year	Location	Model	SNP	Chr	Pos	*P*-value	MAF	HB *p*-value ≤ 0.05	Effect	Associated gene
Zn	2020	Melkassa	BLINK	S01_30708847	1	30708847	7.56E-07	0.25	0.02	-1.37	*Sobic.001G258400*
Zn	2020	Melkassa	BLINK	S03_2052347	3	2052347	8.28E-08	0.15	0	-1.32	*Sobic.003G023500*
Zn	2020	Melkassa	FarmCPU	S03_2052347	3	2052347	6.17E-08	0.15	0	-1.07	*Sobic.003G023500*
Zn	2021	Melkassa	BLINK	S03_2052347	3	2052347	3.14E-07	0.15	0.01	-1.05	*Sobic.003G023500*
Zn	2021	Melkassa	FarmCPU	S03_2052347	3	2052347	2.69E-07	0.15	0.01	-0.85	*Sobic.003G023500*
Zn	2021	Miesso	BLINK	S03_2052347	3	2052347	2.08E-12	0.15	0	-2.15	*Sobic.003G023500*
Zn	2020	Jimma	BLINK	S08_51315455	8	51315455	4.07E-09	0.41	0	2.49	*Sobic.008G110200*
Zn	2020	Jimma	FarmCPU	S08_51315455	8	51315455	1.22E-09	0.41	0	2.27	*Sobic.008G110200*
Zn	2020	Jimma	BLINK	S08_9366659	8	9366659	7.31E-10	0.36	0	-3.26	*Sobic.008G071050*
Zn	2020	Jimma	FarmCPU	S08_9366659	8	9366659	4.60E-11	0.36	0	-3.01	*Sobic.008G071050*
Zn	2021	Jimma	BLINK	S08_9366659	8	9366659	1.36E-08	0.36	0	-1.41	*Sobic.008G071050*
Zn	2020	Melkassa	BLINK	S08_9366659	8	9366659	2.51E-09	0.36	0	-1.45	*Sobic.008G071050*
Zn	2021	Melkassa	BLINK	S08_9366659	8	9366659	1.42E-09	0.36	0	-1.24	*Sobic.008G071050*
Zn	2021	Melkassa	FarmCPU	S08_9366659	8	9366659	4.03E-07	0.36	0.01	-0.85	*Sobic.008G071050*

Fe = Iron; Zn = Zinc; SNP = single nucleotide polymorphism; Chr = Chromosome; Pos = position on chromosome; MAF = marker allele frequency; HB p-value ≤0.05 = Holm-Bonferroni with adjustment; GLM = General linear model; BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; FarmCPU = fixed and random model circulating probability unification.

6.4.4 GWAS for mineral composition in sorghum

GWAS, using the BLINK and FarmCPU models, identified 439 different SNPs with significant association with ash, Fe and Zn at −log10 $p \geq 3.5$ ($p \leq 0.0003$), which were reduced to 32 (13 using BLINK and 19 using FarmCPU) after Holm-Bonferroni error rate adjustment (HB p-value ≤ 0.05) in BLUP analysis. BLINK identified four SNPs link to ash, one to Fe and eight to Zn content. FarmCPU identified eight SNPs associated with ash and one link to Fe and 10 to Zn content (Table 6.3).

Table 6.3 Number of single nucleotide polymorphism markers identified to be significantly linked to ash, Fe and Zn content in sorghum, using five association models in best linear unbiased prediction analysis before and after Holm-Bonferroni (HB p-value ≤ 0.05) adjustment

	BLINK		**FarmCPU**		**CMLM**		**MLM**		**GLM**	
	−log10	HB	−log10	HB	−log10	HB	−log10	HB	−log10	HB
Trait	$p \geq 3.5$	$p \leq 0.05$	$p \geq 3.5$	$p \leq 0.05$	$p \geq 3.5$	$p \leq 0.05$	$p \geq 3.5$	$p \leq 0.05$	$p \geq 3.5$	$p \leq 0.05$
Ash	114	4	77	8	73	0	74	0	731	217
Fe	66	1	31	1	40	0	60	0	157	1
Zn	69	8	82	10	74	0	116	0	225	12

p-value threshold (–log10 $p \geq 3.5$) without adjustment and a Holm-Bonferroni (HB) = p-value ≤ 0.05 with adjustment. BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; FarmCPU = fixed and random model circulating probability unification; CMLM = compressed mixed linear model; MLM = mixed linear model; GLM = general linear model.

SNP density

A total of 209,572 SNPs were identified in the 361 sorghum landraces using GBS Chromosome 2 had the most SNP markers (26,036), while chromosome 7 had the least (15,471). The highest SNP density was observed on chromosome 2 with 6.4 SNP markers per 10 kb, whereas the lowest density was on chromosome 7 with 4.7 SNP markers per 10 kb. The average density was 5.8 markers per 10 kb (Table 6.4). Overall, the genotyping results were of high quality in this study. The density of markers in the Manhattan plots of association mapping varied substantially along chromosomes (Figure 6.2). Across mapping populations and germplasm diversity investigations, at least two major chromosome regions with high marker density have been consistently detected in sorghum chromosomes 2, whereas the lowest was in chromosome 7.

Table 6.4 SNP density per chromosome calculated for individual 10 kb non-overlapping windows

Chromosome	Number of SNPs	Chromosome length (kb)	SNP density (SNP/10kb)
1	25785	41218.19	6.30
2	26036	40852.89	6.40
3	24924	39475.24	6.30
4	20963	36468.00	5.70
5	23312	37509.51	6.20
6	19727	35990.80	5.50
7	15471	32990.27	4.70
8	18329	33977.96	5.40
9	16903	31669.38	5.30
10	18122	30029.81	6.00
Whole genome	209572	36018.21	5.80

SNPs = Single nucleotide polymorphisms; kb = kilobase

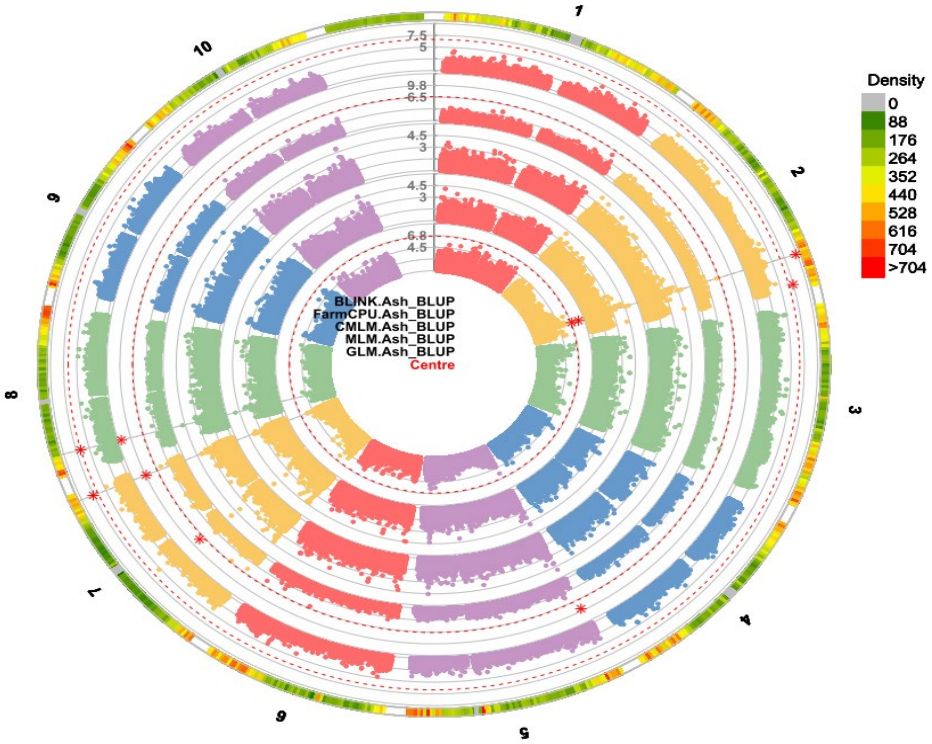


Figure 6.2 Density of SNPs across the genome. Manhattan plots of association mapping for ash. The highest SNPs coverage was observed on chromosome 2, while the lowest was on chromosome 7.

BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; FarmCPU = fixed and random model circulating probability unification; CMLM = compressed mixed linear model; MLM = mixed linear model; GLM = general linear model; BLUP= best linear unbiased prediction.

Ash

A total of 12 SNPs, at a significance level of HB *p*-value ≤ 0.05, were detected to be significantly associated with ash content using the FarmCPU and BLINK models in BLUP analysis. FarmCPU detected eight and BLINK four SNPs linked to ash content (Table 6.3). Ten regions were identified across the genome as having a significant association with ash content (Figure 6.2). Among these significant SNPs, three (S02_65113562, S07_62611661 and S08_7974942; Figure 6.3, Table 6.5) were commonly and consistently detected by both the FarmCPU and BLINK models and were located on chromosomes 2, 7 and 8. The SNP S02_65113562 was identified in the region of chromosome 2 with the highest SNP density, whereas S07_62611661 was found in the region with the lowest SNP density on chromosome 7. The first SNP, S02_65113562, was closely linked to the *Sobic.002G266800* gene that encodes for a flavonoid 3-O-glucosyltransferase (UFGT), a protein belonging to the UDP-glycosyltransferase (UGT) family of enzymes. The second SNP, S07_62611661 was linked to the *Sobic.007G193500* gene with a functional annotation to teosinte glume architecture1 (tga1) while the third SNP, S08_7974942 was linked to the *Sobic.008G068300* gene with a functional annotation to the Os05g0278500 protein (Table 6.5).

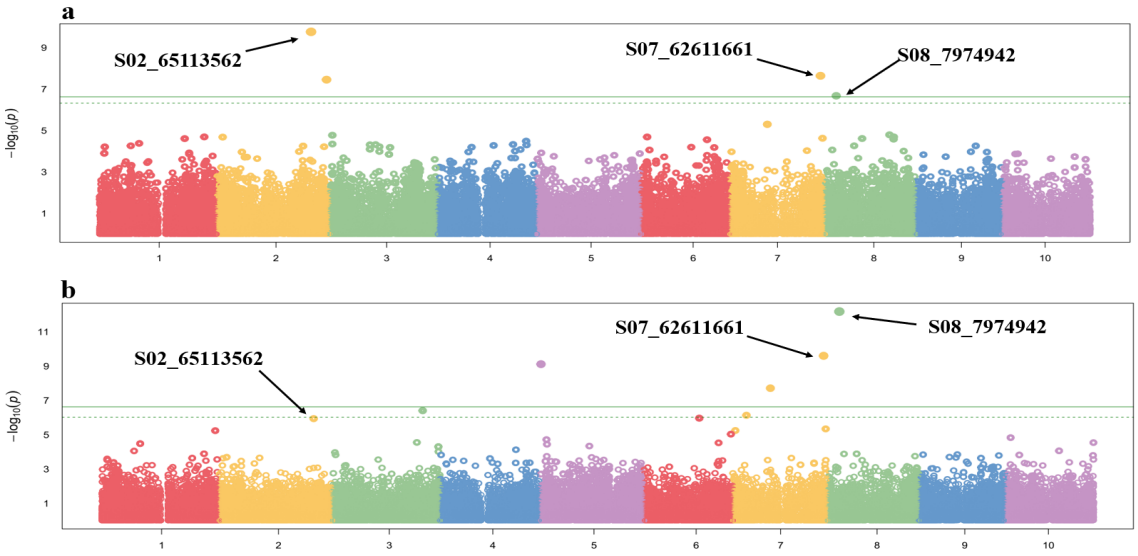


Figure 6.3 Genome-wide association study sorghum landrace grain mineral composition using the best linear unbiased prediction (BLUP). Manhattan plots for ash using (a) the Bayesian information and linkage disequilibrium iteratively nested keyway (BLINK) and (b) the fixed and random model circulating probability unification (FarmCPU) models.

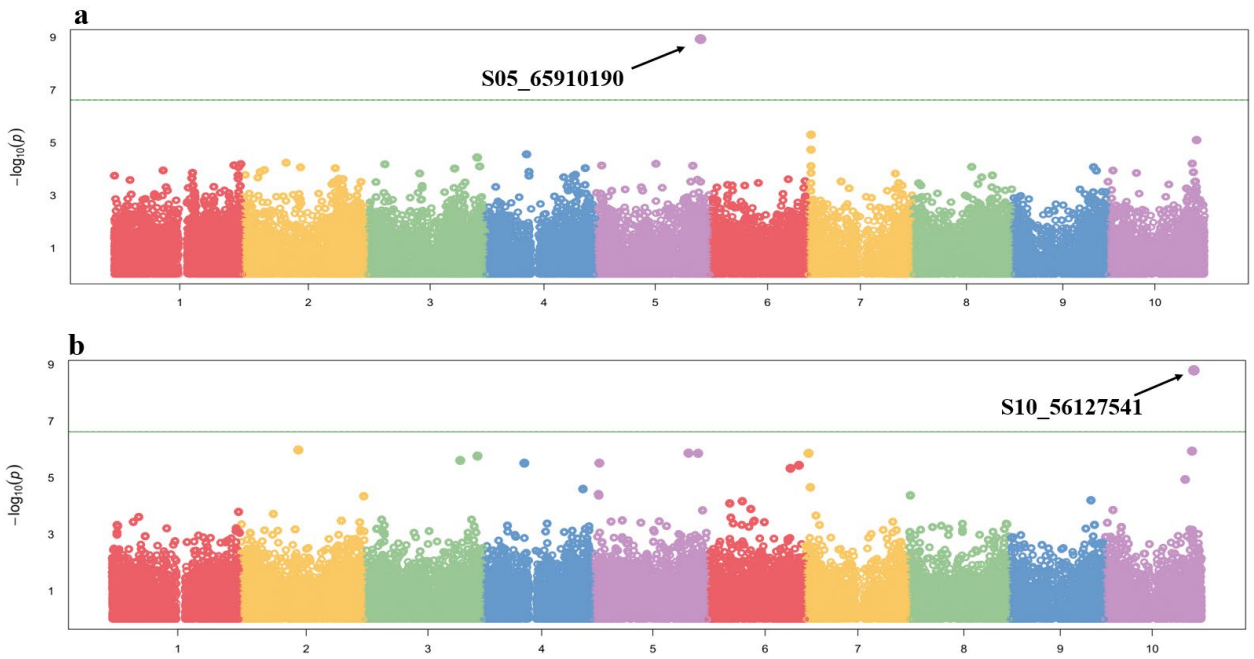


Figure 6.4 Genome-wide association study sorghum landrace grain mineral composition using the best linear unbiased prediction. Manhattan plot for Fe using the Bayesian information and linkage disequilibrium iteratively nested keyway (a) and fixed and random model circulating probability unification (b) models.

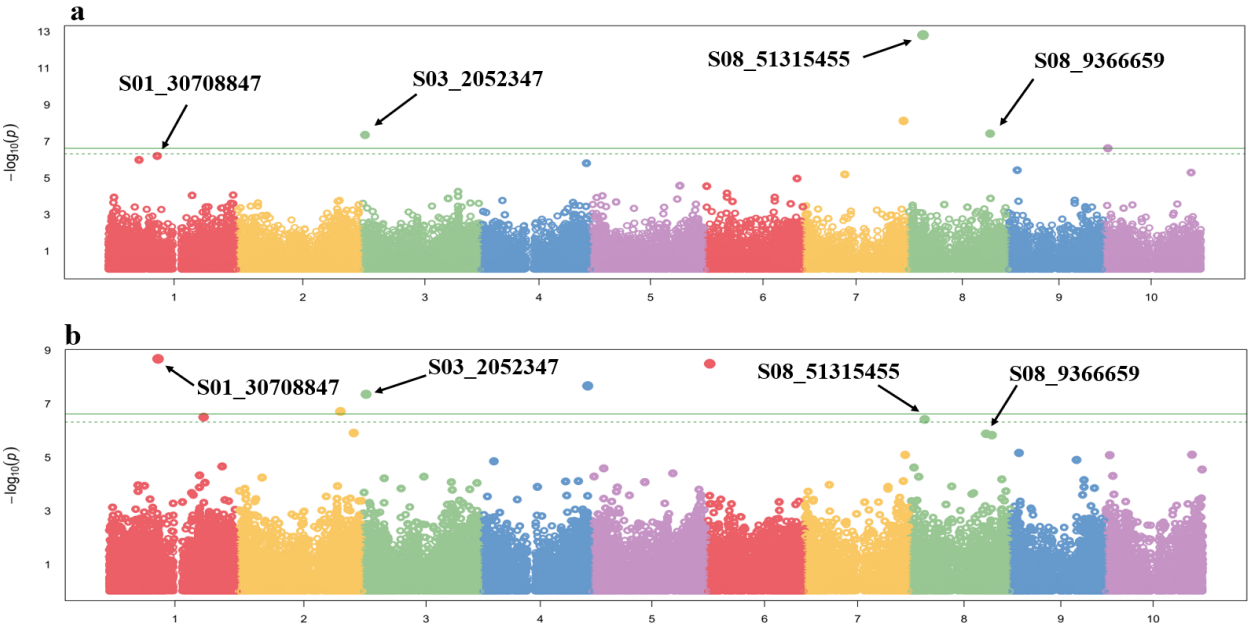


Figure 6.5 Genome-wide association study sorghum landrace grain mineral composition using the best linear unbiased prediction. Manhattan plots Zn using the Bayesian information and linkage disequilibrium iteratively nested keyway (a) and fixed and random model circulating probability unification (b) models. Significant markers detected by both models are indicated by arrows.

Table 6.5 Summary of SNPs consistently and significantly associated with micronutrient composition in sorghum landrace collections

Trait	Model	SNP	Chr	Pos	*p*-value	MAF	HB *p*-value ≤ 0.05	Effect	Associated gene	Putative gene description
Ash	GLM	S02_65113562	2	65113562	9.35E-09	0.26	0.00	0.11	*Sobic.002G266800*	Flavonoid 3-O-glucosyltransferase
Ash	BLINK	S02_65113562	2	65113562	1.72E-10	0.26	0.00	0.10	*Sobic.002G266800*	Flavonoid 3-O-glucosyltransferase
Ash	FarmCPU	S02_65113562	2	65113562	0.00	0.26	0.04	0.04	*Sobic.002G266800*	Flavonoid 3-O-glucosyltransferase
Ash	BLINK	S07_62611661	7	62611661	2.28E-08	0.33	0.00	-0.07	*Sobic.007G193500*	Teosinte glume architecture1 (tga1)
Ash	FarmCPU	S07_62611661	7	62611661	2.51E-10	0.33	0.00	-0.07	*Sobic.007G193500*	Teosinte glume architecture1 (tga1)
Ash	BLINK	S08_7974942	8	7974942	2.11E-07	0.41	0.01	-0.10	*Sobic.008G068300*	Os05g0278500 protein
Ash	FarmCPU	S08_7974942	8	7974942	6.59E-13	0.41	0.00	-0.10	*Sobic.008G068300*	Os05g0278500 protein
Fe	GLM	S05_65910190	5	65910190	8.60E-08	0.24	0.02	4.94	*Sobic.005G177500*	Os05g0247500 protein
Fe	BLINK	S05_65910190	5	65910190	1.17E-09	0.24	0.00	4.94	*Sobic.005G177500*	Os05g0247500 protein
Fe	FarmCPU	S05_65910190	5	65910190	1.37E-06	0.24	0.05	2.72	*Sobic.005G177500*	Os05g0247500 protein
Fe	FarmCPU	S10_56127541	10	56127541	1.61E-09	0.15	0.00	3.48	*Sobic.010G218500*	RISBZ5
Zn	BLINK	S01_30708847	1	30708847	6.27E-07	0.25	0.02	-1.33	*Sobic.001G258400*	Glutaredoxin-C6
Zn	FarmCPU	S01_30708847	1	30708847	2.10E-09	0.25	0.00	-1.47	*Sobic.001G258400*	Glutaredoxin-C6
Zn	GLM	S03_2052347	3	2052347	4.09E-07	0.15	0.01	-1.84	*Sobic.003G023500*	Os06g0116300
Zn	BLINK	S03_2052347	3	2052347	4.44E-08	0.15	0.00	-1.28	*Sobic.003G023500*	Os06g0116300
Zn	FarmCPU	S03_2052347	3	2052347	4.41E-08	0.15	0.00	-1.15	*Sobic.003G023500*	Os06g0116300
Zn	BLINK	S08_51315455	8	51315455	3.75E-08	0.41	0.00	1.05	*Sobic.008G110200*	Pollen-specific protein SF3
Zn	FarmCPU	S08_51315455	8	51315455	1.48E-06	0.41	0.03	0.75	*Sobic.008G110200*	Pollen-specific protein SF3
Zn	GLM	S08_9366659	8	9366659	1.08E-07	0.36	0.01	-1.88	*Sobic.008G071050*	The 3' exoribonuclease family of proteins
Zn	BLINK	S08_9366659	8	9366659	1.54E-13	0.36	0.00	-1.76	*Sobic.008G071050*	The 3' exoribonuclease family of proteins
Zn	FarmCPU	S08_9366659	8	9366659	3.84E-07	0.36	0.01	-1.06	*Sobic.008G071050*	The 3' exoribonuclease family of proteins

HB = Holm-Bonferroni (adjusted); SNP = single nucleotide polymorphism; Chr = Chromosome, Pos = position on chromosome; MAF = minor allele frequency; GLM = general linear model; BLINK = Bayesian information and linkage disequilibrium iteratively nested keyway; FarmCPU = fixed and random model circulating probability.

Fe

At a significance level of HB *p*-value ≤ 0.05, two SNPs S05_65910190 and S10_56127541 were significantly associated with Fe using the GLM, FarmCPU and BLINK models. The S05_65910190 SNP was detected using all three models, while the S10_56127541 SNP was detected only using FarmCPU in BLUP analysis. The two significant SNPs, S05_65910190 and S10_56127541, were detected on chromosomes 5 and 10, respectively (Figure 6.4 and Table 6.5). The *Sobic.005G177500* gene, which fell within the S05_65910190 SNP region, encodes an Os05g0247500 protein that is a member of the plant chitinase (CHI) gene family and is widely implicated in pathogen defence response. The *Sobic.010G218500* gene, located within the S10_56127541 SNP region, is known to encode a RISBZ5 protein. The RISBZ5 is a transcription factor belonging to the basic leucine zipper (bZIP) family of proteins.

Zn

The significantly associated SNPs, S01_30708847, S03_2052347, S08_51315455 and S08_9366659, were identified likely to be linked Zn content in sorghum (Figure 6.5 and Table 6.5). Various of the candidate genes identified to be linked to these SNPs did not have any direct association with known Zn traits. Some of these may be false positive associations, while others may be novel associations yet to be discovered as causal genes for Zn variation.

The S01_30708847 SNP was associated with the *Sobic.001G258400* gene, which is a member of the Glutaredoxin-C6 family of proteins. SNPs S03_2052347 was located close to the *Sobic.003G023500* gene. This gene encodes for a protein that belongs to the family of zinc finger proteins, Os06g0116300. The S08_51315455 SNP was linked to the *Sobic.008G110200* gene, encoding a putative protein in the sorghum genome with sequence similarity to the pollen-specific protein SF3. The S08_9366659 SNP was linked to the *Sobic.008G071050* gene predicted to encode a protein similar to the 3' exoribonuclease family, domain 1 containing protein.

6.5 Discussion

In this study, GWAS was conducted to identify genetic markers associated with grain ash, Fe and Zn contents in sorghum. The BLINK and FarmCPU models, using more complex methods of Holm-Bonferroni adjustments (HB p-value ≤ 0.05), identified 32 SNPs linked to mineral composition in sorghum landrace collections in BLUP analysis.

Identification of loci across environments

Genetic diversity within a population can influence SNP frequencies across environments. If a population has high genetic diversity, it may exhibit greater variation in SNP frequencies across different environments (Muir et al., 2008). In the current study, the highest number of SNPs linked to Fe and Zn content were detected in Melkassa and Jimma for 2020 and 2021 compared to Miesso. When comparing SNPs across environments within the same population of crops, it is important to consider factors such as environmental conditions, management practices, and genetic diversity. Environmental factors such as soil type, temperature, moisture and nutrient availability can affect the expression of genes related to mineral nutrient uptake and metabolism, which in turn, can lead to changes in SNP frequencies over time and across environments. As a result, SNPs associated with nutrient content may vary across different environments even within the same population.

The presence or absence of specific SNPs linked to genes related to nutrient transport and metabolism can influence the accumulation of ash, Fe and Zn in sorghum landrace collections. Tang et al. (2017) reported for rice landraces grown in different regions, variations in the *OsNRAMP5* gene, which encodes a transporter protein that facilitates the uptake of Fe and Zn in rice, were associated with differences in nutrient content across environments. A similar study on maize landraces grown in different regions found that variations in the *ZmZIP4* gene, which encodes a transporter protein that facilitates the uptake of Zn in maize, were associated with differences in Zn content across environments (Li et al., 2013). The current study suggested that the presence or absence of specific SNPs can contribute to variation in nutrient content across different environments within the same sorghum landrace collections. Understanding these genetic variations can help breeders develop crops that are better adapted to specific environments and that produce higher nutrient yields.

Genome-wide association and identification of candidate genes

SNP density

The two major chromosome regions with high marker density have been consistently detected in sorghum chromosomes 2, with the lowest in chromosome 7. Similarly, Li et al. (2018) found the region with the lowest SNP density on chromosome 7 in sorghum. The density of DNA polymorphisms (SNPs) in sorghum genomes varied widely, with higher density in euchromatic regions and lower density in pericentromeric regions. The research showed that genes encoded in heterochromatic pericentromeric DNA had lower SNP density and higher levels of SNP density situated in euchromatic DNA (Evans et al., 2013). The density of markers in the consensus map varied substantially along chromosomes, even within intervals of 20 cM or less, most likely due to changes in recombination frequency and/or genetic diversity (Erayman et al., 2004; Saintenac et al., 2011). The estimates of DNA content of euchromatic and heterochromatic areas based on metaphase chromosomal measurements are predicated on the assumption that sorghum chromosomes are uniformly compressed at this point (Kim et al., 2005). While centromeric and pericentromeric regions form a repressive chromatin environment in general, some of these repetitive elements are expressed at low levels in specific tissues or developmental stages and processed by the RNAi pathway (May et al., 2005; Slotkin et al., 2009; Slotkin, 2010).

Ash

Three SNPs were identified to be significantly linked to ash content in sorghum and were significantly associated with three different sorghum genes: *Sobic.002G266800*, *Sobic.007G193500* and *Sobic.008G068300.*

Sobic.002G266800 is annotated as a flavonoid 3-O-glucosyltransferase (UFGT), which catalyses the transfer of glucose from uridine diphosphate (UDP) glucose to hydroxyl groups of flavonoids, resulting in the formation of flavonoid glycosides. Hatcher and Kruger (1997) studied the simple phenolic acid composition of Canadian wheat flours for extremely significant connections with ash concentration, colour and polyphenol oxidase activity. The study identified that the polyphenol (tannin) had a relationship with ash content in sorghum cultivars (Tasie and Gebreyse, 2020).

Flavonoids are a class of secondary metabolites that are found in many plant species and have been associated with a variety of biological functions, including pigmentation, ultraviolet protection and defence against herbivores and pathogens (Lattanzio et al., 2009). Plant secondary metabolism is a vast reservoir of natural chemical diversity that includes a wide range of compounds and enzymes as well as a diverse set of mechanisms for gene regulation and metabolite and enzyme transport. Nida et al. (2019) reported that flavonoid biosynthesis genes like *dihydroflavonol 4-reductase 3 (DFR3)* were found to be overexpressed in mold-resistant sorghum lines. Many findings have been reported where plants synthesise a wide range of secondary compounds (flavonoids and phenolics) and, have developed a chemical defence against environmental stresses (Bartwal et al., 2013) and pathogens (Zaynab et al., 2018), including resistance to grain mold in sorghum (Esele et al., 1993; Melake-Berhan et al., 1996; Menkir et al., 1996; Audilakshmi et al., 1999; Nida et al., 2019).

In sorghum, flavonoid glycosides are major pigments responsible for the red, purple and black colouration of various tissues, including the pericarp, seed coat and vegetative organs. They have also been implicated in plant-microbe interactions, as some flavonoid glycosides can act as signalling molecules or antimicrobial compounds. Therefore, it is likely that *Sobic.002G266800* plays a role in the biosynthesis of flavonoid glycosides in sorghum, which in turn contributes to the plant's pigmentation and defence strategies.

Flavonoids are a type of secondary plant phenolic compound with powerful antioxidant and chelating properties. Flavonoids are found primarily in foods as glycosides and polymers, which are degraded to varying degrees in the digestive tract (Heim et al., 2002). As a result, flavonoids with antioxidant properties may be useful as complementary therapies (Wang et al., 2021). Plants have evolved complex defence mechanisms to protect themselves against pathogens and secondary metabolites play a crucial role in these defence responses by the production of antimicrobial compounds **(**alkaloids, terpenoids and phenolic) (Zaynab et al., 2018). These secondary metabolites can directly inhibit the growth and development of pathogens (Kumar et al., 2020).

The glumes can be attributed to the presence of secondary compounds, such as anthocyanins. Anthocyanins are pigments that can produce red, purple, or blue colours in plants (Goodman et al., 2004; Khoddami et al., 2017). They are known for their antioxidant properties and can be involved in defence mechanisms against herbivores and protection against environmental stress.

Sobic.007G193500 is annotated as teosinte glume architecture1 (tga1) and is involved in the development of teosinte glumes, which are modified leaves that protect the reproductive structures of the plant. This gene may play a role in the development of sorghum glumes, the outermost protective structures that surround the grain. The mature grain is made up of organs, tissues, cell types and glume cover, all of which differ greatly in their structure, composition and properties, influencing the processing and nutritional quality of the grain (Gubatz et al*.,* 2010). Naked grains with low glume coverage significantly improved threshing efficiency and seed quality in sorghum (Xie et al., 2022). On the other hand, *Sobic.008G068300* is a functional annotation of the Os05g0278500 protein found in rice and is involved in cell wall biosynthesis (Deng et al., 2022).

Most cell wall components vary significantly in their cell wall synthesis gene functions and development and have an intimate influence on plant growth, development and physiology (Lin et al., 2016). Grain glumes are the protective outer coverings of cereal grains. These structures of grain glum are important for protecting the developing seeds from environmental factors and pests (Shen et al., 2023). The relationships involving grain glumes are primarily related to the reproductive success and survival of the plant. Ficco et al. (2020) investigated the relationship between kernel shape and size and ash distribution in durum wheat. They found that thicker kernels had a more uniform distribution of ash content and thickness could be a useful trait for predicting milling quality. Therefore, *Sobic.008G068300* may play a similar role in sorghum, potentially affecting the strength or structure of the plant cell walls, including the glume cover of sorghum.

Fe

Significant SNPs linked to candidate genes associated with sorghum Fe content were identified. Two significant SNPs and candidate genes were identitifed: S05_65910190 detected the *Sobic.005G177500* gene on chromosome 5 and S10_56127541 detected the *Sobic.010G218500* gene on chromosome 10. However, Shakoor et al. (2016) discovered the SNP S1_19766414 and candidate gene *Sobic.001G213400* on chromosome 7 for Fe concentration in sorghum.

The SNP found on chromosome 5 was linked to the *Sobic.005G177500* gene known to encode the Os05g0247500 protein. This protein is a member of the plant chitinase (CHI) family and is widely

implicated in defence response. Chitinases (CHIs) (EC 3.2.1.14) are lytic enzymes that catalyse chitin degradation (Henrissat, 1991). Chitin, the second most abundant biopolymer, is a vital component of insects and fungi (Mack et al., 2015). Chitin is a component of the cell wall of most fungi (Miya et al., 2007), or lipopolysaccharide is a component of the envelope of all gram-negative bacteria (Ranf et al., 2015).

Wang et al. (2020) identified the CHI gene family in rice and Arabidopsis, as well as potential candidates for improving plant resistance using a transgenic approach. Plant CHIs are members of a large gene family and some play critical roles in plant defence against pathogens through the creation of harmful reactive oxygen species (ROS), making it critical to understand their function for application in genetic analysis and breeding. Because CHIs break down chitin, it can function as a defence-related enzyme that inhibits fungal growth.

Several genes involved in Fe homeostasis in plants are regulated in response to biotic stresses, indicating that infection disrupts Fe homeostasis (Liu et al., 2007; Segond et al., 2009; Calla et al., 2014). Aznar et al. (2015) reported that Fe can catalyse the formation of deleterious ROS. Plants may use Fe to increase local oxidative stress in defence responses against pathogens because it can stimulate the creation of harmful ROS.

Furthermore, Fe deficiency and excess amounts can disrupt the homeostasis of a plant cell, resulting in a decrease in photosynthetic activity, respiration, and increased accumulation of Na^{+} and Ca^{-} ions, culminating in an excess generation of ROS (Tripathi et al., 2018). Plants that are stressed by Fe deficiency may express more genes related to Fe absorption and transport. Zhang et al. (2019) reported that plants have developed several mechanisms for effective Fe uptake in response to limited soil Fe availability. These mechanisms enable plants to more effectively adapt to Fe-deficient environments. These processes include the uptake of Fe from the soil, the movement of Fe from roots to shoots and the storage of Fe within cells.

The second SNP identified to be significantly linked to Fe content, was linked to the *Sobic.010G218500* gene encoding a protein from the RISBZ5 transcription factor belonging to the

basic leucine zipper (bZIP) family of proteins. In plants, bZIP transcription factors are involved in various stress responses, including those to cold, salinity and drought stress (Liu et al., 2012).

Plants are severely affected when exposed to excessive salinity because it causes a wide range of physiological, biochemical and molecular alterations (Tester and Davenport, 2003; Khan et al., 2009). Drought harms cell division and expansion, root elongation, leaf size, root proliferation and shoot development suppression in crop plants (Sharp et al., 2004; Yamaguchi et al., 2010).

Drought causes an increase in the production of ROS as a result of changes in electron transportation pathways (Asada, 2006; Waraich et al., 2011). Whereas plants have an abundance of mechanisms to combat increased ROS levels during abiotic stress conditions, plants appear to purposefully generate ROS as signalling molecules to control various processes such as pathogen defence, programmed cell death and stomatal behaviour (Apel and Hirt, 2004; Miller et al., 2010). While ROS have the potential to cause oxidative damage to cells in response to environmental stresses, recent research has shown that ROS play an important role in plants as signal transduction molecules involved in mediating responses to pathogen infection, environmental stresses, programmed cell death and various developmental stimuli (Mittler et al., 2004; Torres and Dangl 2005; Qamer et al., 2021).

Plants under abiotic stress can allocate resources differently, potentially affecting the distribution of Fe within the plant (Bisht et al., 2023). Mineral nutrition enhances plant resistance to environmental stresses. Environmental stress factors raise the risk of photooxidative damage in chloroplasts by restricting the use of absorbed light energy in photosynthesis. The deficient mineral nutritional status of plants under stress environmental conditions can exacerbate photooxidative damage and limit plant performance because a sufficient supply of mineral nutrients is essential for the maintenance of photosynthetic electron transport and carbon metabolism (Cakmak, 2005). Stress-induced changes in plant root architecture and physiology can impact the uptake of Fe (Mata-Pérez et al., 2023).

Fe regulates the adverse effects of salinity, drought and heavy metals by controlling the redox status of the cell and the antioxidant defence system (Chakrabarty et al., 2016; Tripathi et al.,

2018). Overall, these results provide new insights into the molecular mechanisms underlying stress responses in sorghum and could have implications for the development of crops that are more resilient to environmental stress.

Zn

The first SNP located on chromosome 1 and linked to Zn content, was located within the *Sobic.001G258400* gene, which is also known as the glutaredoxin-C6 protein. Glutaredoxins are small proteins that play a role in redox signalling and antioxidant defence in cells. Sousa et al. (2019) reported that glutaredoxins are enzymes that play an important role in the regulation of cellular redox balance. Glutaredoxins are involved in the maintenance of protein thiol-disulfide homeostasis and are critical for several cellular processes, including DNA synthesis and repair and metabolism. Cellular redox detected using fluorescent redox biosensors will aid in the understanding of redox-regulated physiological and pathophysiological processes in plants (Li, 2014; Matsui et al., 2020).

It is important to note that the functioning of the glutaredoxin-C6 protein may depend on the Zn concentration. Zn is an essential nutrient, which plays an important role in many cellular processes, including the proper folding and activity of proteins. Therefore, the expression of the gene may be influenced by the availability of Zn in the plant's environment. The role of Zn in the function of glutaredoxins is complex and both high and low concentration of Zn can affect their activity. Zn ions are essential cofactors for some glutaredoxin isoforms and play a role in stabilising the protein structure and facilitating catalytic activity. However, excessive levels of Zn can lead to oxidative stress and can disrupt the redox balance in cells (Cuypers et al., 2001). Plants may regulate Zn content and distribution as part of their antioxidant defence mechanisms to mitigate oxidative stress. Lin and Arts (2012) reported that plant responses to heavy metals are molecularly regulated through a mechanism known as metal homeostasis, which also regulates the metal-induced ROS signalling pathway. ROS production and signalling are crucial in heavy metal detoxification and tolerance.

The second SNP S03_2052347 is near the *Sobic.003G023500* gene. This putative gene encodes for proteins belonging to the family of zinc finger proteins, namely Os06g0116300. The function

of these proteins has not yet been fully characterised or confirmed in sorghum. However, based on the available information, it is predicted to be related to the Os06g0116300 protein in rice, which are involved in Zn transport (Nawaz et al., 2014). Overall, the Os06g0116300 protein is an important component of Zn uptake and homeostasis in rice plants, which is also important for plant growth and development, as well as the nutritional quality of sorghum grains.

ZNFs are involved in a variety of cellular processes, including DNA binding, transcriptional regulation and protein-protein interactions (Cassandri et al., 2017). This protein is involved in the transport of Zn from the soil into root cells, as well as the distribution of Zn within the plant. In addition, it is also involved in the regulation of Zn homeostasis, ensuring that Zn levels are maintained within optimal ranges in different plant tissues (Shanmugam et al., 2011) and has been shown to play a role in regulating the expression of genes involved in Zn uptake and transport in sorghum. These proteins are localised in the plasma membrane and play roles in regulating Zn homeostasis in plants. Several new zinc finger protein family transporter genes have been identified in *Saccharomyces cerevisiae* and *Arabidopsis thaliana,* where they appear to play a role in Zn ion uptake (Eng et al., 1998). Zn is also involved in many important processes, such as DNA synthesis, protein synthesis and hormone regulation. Deficiencies in Zn can result in stunted growth, leaf chlorosis and reduced root development. It is important to maintain proper levels of Zn in soil and fertiliser applications to ensure optimal plant growth and development. Therefore, genes involved in Zn transport are crucial for plant growth and development (Tripathi et al., 2015; Arif et al., 2022).

Two SNPs linked to Zn content were identified on chromosome 8 and were located in the *Sobic.008G110200* and *Sobic.008G071050* genes, respectively. The *Sobic.008G110200* gene is a putative gene in the sorghum genome that encodes a protein with sequence similarity to the pollen-specific protein SF3 found specifically in pollen cells and is involved in regulating gene expression during pollen development (Sari-Gorla and Pe, 1999). Many studies have investigated the significance of putative transcriptional regulators and their expression patterns in the developing male gametophyte (or pollen grain) and related pollen functions, such as pollen tube growth and fertilisation in sunflowers (Steinmetz et al., 1993), Chinese white pear (Zhou et al., 2016) and cultivated tobacco (Sweetman, 1996). The *Sobic.008G071050* gene was predicted to encode a

protein similar to the 3' exoribonuclease family, domain 1 containing protein. The 3' exoribonuclease family of proteins is involved in the degradation and processing of RNA molecules in cells (Andrade et al., 2009; Łabno et al., 2016). These proteins function as exonucleases enzymes that catalyse the removal of nucleotides from the 3' end of RNA molecules, playing a critical role in regulating gene expression and RNA metabolism (Andrade et al., 2009). Zn ions serve as cofactors for certain enzymes, including exonucleases, and play a crucial role in their catalytic function (Hou et al., 2021; Wang et al., 2023). Therefore, Zn directly influences RNA metabolism through its interactions with RNA-binding proteins and enzymes.

Various studies identified SNPs and candidate genes on the same chromosomes 1, 3 and 8, which corresponded to the current study. Kamal et al. (2023) found SNPs and candidate genes on chromosomes 1, 3 and 8 for Zn concentration in sorghum accessions. A computer-based investigation found 77 candidate genes for Fe and Zn homeostasis; chromosome 1 had the highest number of genes (24) and chromosome 8 the least (Anuradha et al., 2013). However, Shakoor et al. (2016) identified SNPS 7_6880986 and candidate gene *Sobic.007G064900* situated on chromosome 7 for Zn concentration in sorghum.

6.6 Conclusions

Sorghum is an important cereal crop in Ethiopia and understanding the genetic factors that influence the accumulation of ash, Fe and Zn in sorghum grains can help breeders develop varieties with improved nutritional quality. The findings of this GWAS study on the sorghum landrace collection provide important insights into the genetic basis of mineral content in sorghum, which could be used to develop new sorghum varieties with higher mineral contents. The GWAS involved genotyping the sorghum landrace collections using a high-density array of genetic markers and analysing the association between these markers and the levels of ash, Fe and Zn in the grains.

The study may have identified specific candidate genes or regions of the genome associated with higher or lower levels of these minerals. Candidate genes *Sobic.002G266800*, *Sobic.007G193500* and *Sobic.008G068300* were found to be significantly associated with ash. The *Sobic.002G266800* gene plays a role in the biosynthesis of flavonoid glycosides in sorghum, which in turn contributes

to the plant's pigmentation and defence strategies. The *Sobic.007G193500* and *Sobic.008G068300* genes are annotated as teosinte glume architecture1 (tga1) and s05g0278500, respectively; both genes have a role in the development and influence of grain glume structure, composition and features, as well as grain processing and nutritional quality. The *Sobic.005G177500* and *Sobic.010G218500* genes, significantly linked to Fe content, encodes the Os05g0247500 and bZIP transcription factors protein, respectively, both genes are in involved in various biotic and abiotic stress responses. GWAS using multiple models, detected genes across environments linked to Zn content encoded proteins from the ZNF family, which are involved in Zn ion uptake and transport.

A higher number of SNPs linked to Fe and Zn content were detected in Jimma (wet intermediate) and Melkassa (irrigated) than in the other test environment Miesso (stress), providing insight into the influence of environmental and growing conditions on sorghum micronutrient concentration. Once such genomic variants have been identified, they should be used frequently to search for nearby variants that contribute directly to the traits. Researchers use fine-mapping and deep sequencing techniques to discover the exact functional variants responsible for the target traits and the biological pathways involved. This could include investigating the expression patterns of these genes under various environments, determining their regulatory targets and investigating their interactions with other genes and proteins.

6.7 References

Andrade, J.M., Pobre, V., Silva, I.J., Domingues, S. and Arraiano, C.M. (2009). The role of 3′–5′ exoribonucleases in RNA degradation. *Progress in Molecular Biology and Translational Science*, 85*:* 187-229.

Anuradha, K., Prakash, B., Ramu, P., Shah, T., Ashok Kumar, A. and Deshpande, S.P. (2013). In silico identification of candidate genes involved for grain Fe and Zn concentration in sorghum using reported cereals gene homologs. ICRISAT, Patancheru 502 324, India, 10-12

Apel, K. and Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review in Plant Biology*, 55*:* 373-399.

Arif, Y., Singh, P., Siddiqui, H., Naaz, R. and Hayat, S. (2022). Transition metal homeostasis and its role in plant growth and development. In: Khan, S.T., Malik, A. (eds). Microbial Biofertilizers and Micronutrient Availability. Springer, Cham, pp. 159-178. https://doi.org/10.1007/978-3-030-76609-2_8

Asada, K. (2006). Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiology,* 141: 391-396

Audilakshmi, S., Stenhouse, J.W., Reddy, T.P. and Prasad, M.V.R. (1999). Grain mold resistance and associated characters of sorghum genotypes. *Euphytica*, 107: 91 103.

Aznar, A., Chen, N.W., Thomine, S. and Dellagi, A. (2015). Immunity to plant pathogens and iron homeostasis. *Plant Science*, 240: 90-97.

Bartwal, A., Mall, R., Lohani, P., Guru, S.K. and Arora, S. (2013). Role of secondary metabolites and brassinosteroids in plant defense against environmental stresses. *Journal of Plant Growth Regulation*, 32: 216-232.

Bhowmik, D., Chiranjib, K. and Kumar, S. (2010). A potential medicinal importance of zinc in human health and chronic disease. *International Journal of Pharmaceutical and Bio-Medical Science*, 1: 05-11.

Bisht, N., Joshi, H. and Chauhan, P.S. (2023). Nutrients Regulation and Abiotic Stress Tolerance in Plants. In Singh, V.P. and Siddiqui, M.H. (eds). Plant Ionomics: Sensing, Signaling, and Regulation, John Wiley and Sons Ltd. pp. 209-223.

Cakmak, I. (2005). Role of mineral nutrients in tolerance of crop plants to environmental stress factors. In *Proceedings from the International Symposium on Fertigation-Optimizing the Utilization of Water and Nutrients*, pp. 35-48.

Calla, B., Blahut-Beatty, L., Koziol, L., Simmonds, D.H. and Clough, S.J. (2014). Transcriptome analyses suggest a disturbance of iron homeostasis in soybean leaves during white mould disease establishment. *Molecular Plant Pathology*, 15: 576-588.

Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M. and Raschellà, G. (2017). Zinc-finger proteins in health and disease. *Cell Death Discovery*, 3: 1-12.

Chakrabarty, A., Aditya, M., Dey, N., Banik, N. and Bhattacharjee, S. (2016). Antioxidant signaling and redox regulation in drought-and salinity-stressed plants. In: Hossain, M., Wani, S., Bhattacharjee, S., Burritt, D., Tran, L.S. (eds). Drought Stress Tolerance in Plants, Vol 1. Springer, Cham, pp. 465-498. https://doi.org/10.1007/978-3-319-28899-4_

Cuypers, A., Vangronsveld, J., and Clijsters, H. (2001). The redox status of plant cells (AsA and GSH) is sensitive to zinc-imposed oxidative stress in the roots and primary leaves of Phaseolus vulgaris. *Plant Physiology and Biochemistry*, 39: 657-664.

Deng, Q., Dai, L., Chen, Y., Wu, D., Shen, Y., Xie, J. and Luo, X. (2022). Identification of phosphorus stress-related proteins in the seedlings of Dongxiang wild rice (*Oryza rufipogon* Griff.) using label-free quantitative proteomic analysis. *Genes*, 13: 108.

Eng, B.H., Guerinot, M.L., Eide, D. and Saier, M.H. (1998). Sequence analyses and phylogenetic characterization of the ZIP family of metal ion transport proteins. *The Journal of Membrane Biology*, 166: 1-7.

Erayman, M., Sandhu, D., Sidhu, D., Dilbirligi, M., Baenziger, P.S. and Gill, K.S. (2004). Demarcating the gene-rich regions of the wheat genome. *Nucleic Acids Research*, 32: 123546-123565.

Esele, J.P., Frederiksen, R.A. and Miller, F.R. (1993). The association of genes controlling caryopsis traits with grain mold resistance in sorghum. *Phytopathology*, 83: 490-495.

Evans, J., McCormick, R.F., Morishige, D., Olson, S.N. and Weers, B. (2013). Extensive Variation in the Density and Distribution of DNA Polymorphism in sorghum genomes. *PLoS One*, 8: e79192.

Ficco, D.B., Beleggia, R., Pecorella, I., Giovanniello, V., Frenda, A.S. and Vita, P.D. (2020). Relationship between seed morphological traits and ash and mineral distribution along the kernel using debranning in durum wheats from different geographic sites. *Foods*, 9: 1523.

Goodman, C.D., Casati, P. and Walbot, V. (2004). A multidrug resistance-associated protein involved in anthocyanin transport in *Zea mays*. *The Plant Cell*, 16: 1812-1826.

Gubatz, S., Shewry, P.R. and Ullrich, S. (2010). The development, structure, and composition of the barley grain. In: Ullrich, S.E. (ed). Barley. Blackwell Publishing, pp. 391-448.

Gupta, C.P. (2014). Role of iron (Fe) in the body. *IOSR Journal of Applied Chemistry*, *7:* 38-46.

Hatcher, D.W. and Kruger, J.E. (1997). Simple phenolic acids in flours prepared from Canadian wheat: relationship to ash content, color, and polyphenol oxidase activity. *Cereal Chemistry*, 74: 337-343.

Heim, K.E., Tagliaferro, A.R. and Bobilya, D.J. (2002). Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships. *The Journal of Nutritional Biochemistry*, 13: 572-584.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal*, 280: 309-316.

Hindu, V., Palacios-Rojas, N., Babu, R., Suwarno, W.B., Rashid, Z., Usha, R. and Nair, S.K. (2018). Identification and validation of genomic regions influencing kernel zinc and iron in maize. *Theoretical and Applied Genetics*, 131: 1443-1457.

Hou, R., He, Y., Yan, G., Hou, S., Xie, Z. and Liao, C. (2021). Zinc enzymes in medicinal chemistry. *European Journal of Medicinal Chemistry*, 226: 113877.

Kamal, N.M., Gorafi, Y.S.A., Tomemori, H., Kim, J.S., Elhadi, G.M.I. and Tsujimoto, H. (2023). Genetic variation for grain nutritional profile and yield potential in sorghum and the possibility of selection for drought tolerance under irrigated conditions. *BMC Genomics*, 24: 515.

Khan, M.A., Ansari, R., Ali, H., Gul, B. and Nielsen, B.L. (2009). *Panicum turgidum*: a sustainable feed alternative for cattle in saline areas. *Agricultural Ecosystems and Environment,* 129: 542-546. doi: 10.1016/j.agee.2008.10.014

Khoddami, A., Mohammadrezaei, M. and Roberts, T.H. (2017). Effects of sorghum malting on colour, major classes of phenolics and individual anthocyanins. *Molecules*, 22:1713.

Kim, J.S., Islam-Faridi, M.N., Klein, P.E., Stelly, D.M., Price, H.J., Klein, R.R. and Mullet, J.E. (2005). Comprehensive molecular cytogenetic analysis of sorghum genome architecture: distribution of euchromatin, heterochromatin, genes and recombination in comparison to rice. *Genetics*, 171: 1963-1976.

Kotla, A., Phuke, R., Hariprasanna, K., Mehtre, S.P., Rathore, A., Gorthy, S. and Kumar, A.A. (2019). Identification of QTLs and candidate genes for high grain Fe and Zn concentration in sorghum [*Sorghum bicolor* (L.) Moench]. *Journal of Cereal Science*, 90: 102850.

Kumar, S., Abedin, M.M., Singh, A.K. and Das, S. (2020). Role of Phenolic Compounds in Plant-Defensive Mechanisms. In: Lone, R., Shuab, R., Kamili, A. (eds). Plant Phenolics in Sustainable Agriculture. Springer, Singapore, pp. 517-532. https://doi.org/10.1007/978-981-15-4890-1_22

Łabno, A., Tomecki, R. and Dziembowski, A. (2016). Cytoplasmic RNA decay pathways-Enzymes and mechanisms. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1863: 3125-3147.

Lattanzio, V., Kroon, P.A., Quideau, S. and Treutter, D. (2009). Plant phenolics-secondary metabolites with diverse functions. *Recent Advances in Polyphenol Research*, 1: 1-35.

Li, J., Tang, W., Zhang, Y.W., Chen, K.N., Wang, C., Liu, Y., Zhan, Q., Wang, C., Wang, S.B., Xie, S.Q. and Wang, L. (2018). Genome-wide association studies for five forage quality-related traits in sorghum (*Sorghum bicolor)*. *Frontiers in Plant Science*, 9: 1146.

Li, S. (2014). Redox modulation matters: emerging functions for glutaredoxins in plant development and stress responses. *Plants*, 3: 559-582.

Li, S., Zhou, X., Huang, Y., Zhu, L., Zhang, S., Zhao, Y. and Chen, R. (2013). Identification and characterization of the zinc-regulated transporters, iron-regulated transporter-like protein (ZIP) gene family in maize. *BMC Plant Biology*, 13: 1-14.

Lin, F., Manisseri, C., Fagerström, A., Peck, M.L., Vega-Sánchez, M.E., Williams, B. and Bartley, L.E. (2016). Cell wall composition and candidate biosynthesis gene expression during rice development. *Plant and Cell Physiology*, 57: 2058-2075.

Lin, Y.F. and Aarts, M.G. (2012). The molecular mechanism of zinc and cadmium stress response in plants. *Cellular and Molecular Life Sciences*, 69: 3187-3206.

Liu, C., Wu, Y. and Wang, X. (2012). bZIP transcription factor OsbZIP52/RISBZ5: a potential negative regulator of cold and drought stress response in rice. *Planta*, 235: 1157-1169.

Liu, G., Greenshields, D.L., Sammynaiken, R., Hirji, R.N., Selvaraj, G. and Wei, Y. (2007). Targeted alterations in iron homeostasis underlie plant defense responses. *Journal of Cell Science*, 120: 596-605.

Mack, I., Hector, A., Ballbach, M., Kohlhäufl, J., Fuchs, K.J., Weber, A. and Hartl, D. (2015). The role of chitin, chitinases, and chitinase-like proteins in pediatric lung diseases. *Molecular and Cellular Pediatrics*, 2: 1-8.

Mata-Pérez, C., Sánchez-Vicente, I., Arteaga, N., Gómez-Jiménez, S., Fuentes-Terrón, A., Oulebsir, C.S., Calvo-Polanco, M., Oliver, C. and Lorenzo, Ó. (2023). Functions of nitric oxide-mediated post-translational modifications under abiotic stress. *Frontiers in Plant Science*, 14: 1158184.

Matsui, R., Ferran, B., Oh, A., Croteau, D., Shao, D., Han, J. and Bachschmid, M.M. (2020). Redox regulation via glutaredoxin-1 and protein S-glutathionylation. *Antioxidants and Redox Signaling,* 32: 677-700.

May B.P., Lippman Z.B., Fang Y., Spector D.L. and Martienssen R.A. (2005). Differential regulation of strand-specific transcripts from *Arabidopsis* centromeric satellite repeats. *PLoS Genetics,* 1: e79.

Melake-Berhan, A., Butler, L.G., Ejeta, G. and Menkir, A. (1996). Grain mold resistance and polyphenol accumulation in sorghum. *Journal of Agricultural and Food Chemistry*, 44: 2428-2434.

Menkir, A., Ejeta, G., Butler, L. and Melakeberhan, A. (1996). Physical and chemical kernel properties associated with resistance to grain mold in sorghum. *Cereal Chemistry*, 73: 613-617.

Miller, G.A.D., Suzuki, N., Ciftci-Yilmaz, S.U.L.T.A.N. and Mittler, R.O.N. (2010). Reactive oxygen species homeostasis and signaling during drought and salinity stresses. *Plant, Cell and Environment*, 33: 453-467.

Mittler, R., Vanderauwera, S., Gollery, M. and Van Breusegem, F. (2004) Reactive oxygen gene network of plants. *Trends in Plant Science* 9: 490-498.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K. and Shibuya, N. (2007). CERK1, a Lysm receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proceedings of the National Academy of Sciences*, 104: 19613-19618.

Motlhaodi, T., Bryngelsson, T., Chite, S., Fatih, M., Ortiz, R. an0d Geleta, M. (2018). Nutritional variation in sorghum [*Sorghum bicolor* (L.) Moench] accessions from southern Africa revealed by protein and mineral composition. *Journal of Cereal Science*, 83: 123-129.

Muir, W.M., Wong, G.K.S., Zhang, Y., Wang, J., Groenen, M.A., Crooijmans, R.P. and Cheng, H.H. (2008). Genome-wide assessment of worldwide chicken SNP genetic diversity indicates the significant absence of rare alleles in commercial breeds. *Proceedings of the National Academy of Sciences*, 105: 17312-17317.

Nawaz, Z., Kakar, K.U., Saand, M.A. and Shu, Q.Y. (2014). Cyclic nucleotide-gated ion channel gene family in rice, identification, characterization and experimental analysis of expression response to plant hormones, biotic and abiotic stresses. *BMC Genomics*, 15: 1-18.

Nida, H., Girma, G., Mekonen, M., Lee, S., Seyoum, A., Dessalegn, K. and Mengiste, T. (2019). Identification of sorghum grain mold resistance loci through genome-wide association mapping. *Journal of Cereal Science*, 85: 295-304.

Paiva, C.L., Queiroz, V.A.V., Simeone, M.L.F., Schaffert, R.E., de Oliveira, A.C. and da Silva, C.S. (2017). Mineral content of sorghum genotypes and the influence of water stress. *Food Chemistry*, 214: 400-405.

Parthasarathy Rao, P., B.S. Birthal, V.S. Reddy Belum, K.N. Rai, and S. Ramesh. (2006). Diagnostics of sorghum and pearl millet grains-based nutrition in India. *International Sorghum Millets Newsletter,* 47: 93-96

Pereira, P., Úbeda, X. and Martin, D.A. (2012). Fire severity effects on ash chemical composition and water-extractable elements. *Geoderma*, 191: 105-114.

Qamer, Z., Chaudhary, M.T., Du, X., Hinze, L. and Azhar, M.T. (2021). Review of oxidative stress and antioxidative defense mechanisms in *Gossypium hirsutum* L. in response to extreme biotic conditions. *Journal of Cotton Research*, 4: 1-9.

Ranf, S., Gisch, N., Schäffer, M., Illig, T., Westphal, L., Knirel, Y.A. and Scheel, D. (2015). A lectin S-domain receptor kinase mediates lipopolysaccharide sensing in *Arabidopsis thaliana*. *Nature Immunology*, 16: 426-433.

Roohani, N., Hurrell, R., Kelishadi, R. and Schulin, R. (2013). Zinc and its importance for human health: An integrative review. *Journal of Research in Medical Sciences*, 18: 144.

Saintenac, C., Jiang, D. and Akhunov, E.D. (2011). Targeted analysis of nucleotide and copy number variation by exon capture in allotetraploid wheat genome. *Genome Biology,* 12: R88.

Sari-Gorla, M. and Pe, M.E. (1999). Genetic control of pollen development and function. *Fertilization in Higher Plants: Molecular and Cytological Aspects*, 217-233.

Segond, D., Dellagi, A., Lanquar, V., Rigault, M., Patrit, O., Thomine, S. and Expert, D. (2009). NRAMP genes function in *Arabidopsis thaliana* resistance to *Erwinia chrysanthemi* infection. *The Plant Journal*, 58: 195-207.

Shakoor, N., Ziegler, G., Dilkes, B.P., Brenton, Z., Boyles, R., Connolly, E.L., Kresovich, S. and Baxter, I. (2016). Integration of experiments across diverse environments identifies the genetic determinants of variation in *Sorghum bicolor* seed element composition. *Plant Physiology*, 170: 1989-1998.

Shanmugam, V., Lo, J.C., Wu, C.L., Wang, S.L., Lai, C.C., Connolly, E.L. and Yeh, K.C. (2011). Differential expression and regulation of iron-regulated metal transporters in *Arabidopsis halleri* and *Arabidopsis thaliana* – the role in zinc tolerance. *New Phytologist*, 190: 125-137.

Sharp, R.E., Poroyko, V., Hejlek, L.G., Spollen, W.G., Springer, G.K., Bohnert, H.J. and Nguyen, H.T. (2004). Root growth maintenance during water deficits: physiology to functional genomics. *Journal of Experimental Botany*, 55: 2343-2351.

Shen, E., Wang, X., Lu, Z., Zhou, F., Ma, W., Cui, Z., Li, Z., Li, C. and Lin, Y. (2023). Overexpression of a beta-1, 6-glucanase gene GluM in transgenic rice confers high resistance to rice blast, sheath blight and false smut. *Pest Management Science*. 79: 2152-2162.

Slotkin, R.K., Vaughn M., Borges F., Tanurdzić M., Becker J.D. and Feijó J.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell,* 136: 461-472.

Slotkin, R.K. (2010). The epigenetic control of the Athila family of retrotransposons in *Arabidopsis*. *Epigenetics,* 5: 483-490.

Sousa, R.H., Carvalho, F.E., Lima-Melo, Y., Alencar, V.T., Daloso, D.M., Margis-Pinheiro, M., and Silveira, J.A. (2019). Impairment of peroxisomal APX and CAT activities increases the protection of photosynthesis under oxidative stress. *Journal of Experimental Botany*, 70: 627-639.

Steinmetz, A., Baltz, R., Domon, C., Dudareva, N., Evrard, J.L., Kräuter, R. and Tabidze, V. (1993). Flower-specific genes in sunflower. *Biotechnology and Biotechnological Equipment*, 7: 16-20.

Sweetman, J.P. (1996). Characterization and functional analysis of pollen-specific cDNAs encoding putative transcriptional regulators from *Nicotiana tabacum*. PhD thesis, University of Leicester, United Kingdom.

Tang, L., Mao, B., Li, Y., Lv, Q., Zhang, L., Chen, C. and Zhao, B. (2017). Knockout of OsNramp5 using the CRISPR/Cas9 system produces low Cd-accumulating indica rice without compromising yield. *Scientific Reports*, 7: 14438.

Tasie, M.M. and Gebreyes, B.G. (2020). Characterization of nutritional, antinutritional, and mineral contents of thirty-five sorghum varieties grown in Ethiopia. *International Journal of Food Science*, 2020: 1-11.

Tester, M. and Davenport, R. (2003). Na^+ tolerance and Na^+ transport in higher plants. *Annals of Botany,* 91: 503-527. doi: 10.1093/aob/mcg058

Torres, M.A. and Dangl, J.L. (2005) Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. *Current Opinion in Plant Biology,* 8: 397-403.

Tripathi, D.K., Singh, S., Gaur, S., Singh, S., Yadav, V., Liu, S. and Sahi, S. (2018). Acquisition and homeostasis of iron in higher plants and their probable role in abiotic stress tolerance. *Frontiers in Environmental Science*, 5: 86.

Tripathi, D.K., Singh, S., Singh, S., Mishra, S., Chauhan, D.K. and Dubey, N.K. (2015). Micronutrients and their diverse role in crops: advances and future prospective. *Acta Physiologiae Plantarum*, 37: 1-14.

Wang, X., Li, Y., Han, L., Li, J., Liu, C. and Sun, C. (2021). Role of flavonoids in the treatment of iron overload. *Frontiers in Cell and Developmental Biology*, 9: 685364.

Wang, Y., Huang, N. and Yang, Z. (2023). Revealing the role of zinc ions in atherosclerosis therapy via an engineered three-dimensional pathological model. *Advanced Science*, 2300475.

Wang, Z., Wang, B., Zhao, F. and Cao, J. (2020). Genome-wide analysis of the chitinase gene family in rice and Arabidopsis reveals their mechanisms and diverse roles in defense response. *International Journal of Agriculture and Biology*, 24: 812-822.

Waraich, E.A., Ahmad, R. and Ashraf, M.Y. (2011). Role of mineral nutrition in alleviation of drought stress in plants. *Australian Journal of Crop Science*, 5: 764-777.

Xie, P., Tang, S., Chen, C., Zhang, H., Yu, F., Li, C. and Xie, Q. (2022). Natural variation in Glume Coverage 1 causes naked grains in sorghum. *Nature Communications*, 13: 1068.

Yamaguchi, M., Valliyodan, B., Zhang, J., Lenoble, M.E., Yu, O., Rogers, E.E. and Sharp, R.E. (2010). Regulation of growth response to water stress in the soybean primary root. Proteomic analysis reveals region-specific regulation of phenylpropanoid metabolism and control of free iron in the elongation zone. *Plant, Cell and Environment*, 33: 223-243.

Zaynab, M., Fatima, M., Abbas, S., Sharif, Y., Umair, M., Zafar, M.H. and Bahadar, K. (2018). Role of secondary metabolites in plant defense against pathogens. *Microbial Pathogenesis*, 124: 198-202.

Zaynab, M., Fatima, M., Abbas, S., Sharif, Y., Umair, M., Zafar, M.H. and Bahadar, K. (2018). Role of secondary metabolites in plant defense against pathogens. *Microbial Pathogenesis*, 124: 198-202.

Zhang, X., Zhang, D., Sun, W. and Wang, T. (2019). The adaptive mechanism of plants to iron deficiency via iron uptake, transport, and homeostasis. *International Journal of Molecular Sciences*, 20: 2424.

Zhou, H., Yin, H., Chen, J., Liu, X., Gao, Y., Wu, J. and Zhang, S. (2016). Gene-expression profile of developing pollen tube of *Pyrus bretschneideri*. *Gene Expression Patterns*, 20: 11-21.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

The use of BLUP allows for providing a reliable means of estimating genetic values while accounting for environmental variations. The study revealed significant genetic diversity among Ethiopian sorghum landraces in terms of GY and other agronomic traits. This diversity is promising for variety development efforts, suggesting the potential to select high-yielding landraces for further improvement and incorporation into breeding programs. Despite the identification of promising landraces for drought tolerance, the study highlights the negative impact of drought stress on the yield of these identified genotypes. The inverse relationships observed between GY and traits such as HLP, DTF and DTM suggest that early maturing genotypes may exhibit higher GY, especially in stress-prone areas. This information is important for selecting genotypes that combine early maturity with high GY under stress conditions, emphasizing the need for continued research and breeding efforts to enhance drought resilience in sorghum varieties.

Furthermore, the study demonstrated the significant influence of environmental conditions on the nutritional composition of sorghum grains. Stress environments led to lower starch content but higher protein content compared to wet intermediate environments. Notably, wet intermediate environments exhibited higher mineral concentrations in grains compared to irrigated and stress environments. This dynamic response of sorghum to different environmental stresses emphasised the importance of understanding and adapting to diverse conditions.

The GGE biplot analysis allows for the examination of the relative stability and performance of genotypes across environments. This analysis identified specific genotypes with both high and stable starch, protein, Fe and Zn content. Such genotypes are valuable for breeding programmes aiming to develop varieties with consistent nutritional quality across diverse conditions. The positive relationship between Fe and Zn content provides an opportunity for concurrent selection

to improve the nutritional quality of sorghum grains, particularly addressing micronutrient deficiencies in regions where sorghum is a dietary staple.

Moreover, GWAS mapping using Ethiopian sorghum landrace accessions successfully identified SNPs associated with key grain quality traits, including starch, amylose, protein, ash, Fe and Zn content. Several tightly linked genes involved in stress tolerance, sugar transport, biosynthesis, gene regulation and transcription factors were implicated in the regulation of these traits. Five novel loci associated with major effects on sorghum starch biosynthesis have been identified. The genes linked starch content to these loci play diverse roles in plant growth, stress response, sugar transport and storage compound synthesis. Specific candidate genes *Sobic.009G033500* and *Sobic.009G033400* linked with protein content, encoding proteins similar to Class III peroxidase precursors, may contribute to plant defence mechanisms. The three genes *Sobic.002G266800*, *Sobic.007G193500* and *Sobic.008G068300*, were significantly associated with ash content. These genes play roles in flavonoid glycoside biosynthesis, teosinte glume architecture (tga1) and processes influencing grain glume structure, composition and nutritional quality. Genes like *Sobic.005G177500* and *Sobic.010G218500*, linked to Fe content, encode proteins involved in biotic and abiotic stress responses. The SNPs associated with Zn content were found on chromosomes 1, 3 and 8. Understanding the chromosomal locations of these genetic markers is essential for further genetic studies and breeding programs.

The variation in the number of SNPs linked to Fe and Zn content across different environments Jimma (wet intermediate), Melkassa (irrigated) and Miesso (stress) suggests a notable influence of environmental and growing conditions on sorghum micronutrient concentration. The higher number of SNPs in wet intermediate and irrigated environments compared to stress environments highlights the importance of considering these factors in sorghum breeding programs.

Overall, the study provided valuable insights into the genetic basis of grain quality traits in Ethiopian sorghum landraces, contributing to the development of improved sorghum varieties with enhanced grain quality. These findings have practical implications for sorghum breeding programmes, addressing the nutritional needs of populations in diverse environments.

The recommendations resulting from this study emphasise a multifaceted approach to enhancing the adaptability and resilience of sorghum crops in diverse environments. This includes further evaluation through targeted breeding efforts and the consideration of agroecological factors. The study's findings offer valuable insights for sorghum breeding programs, suggesting strategies for the selection of genotypes with early maturity, high yield and desirable nutritional profiles.

In order to assess the adaptability of sorghum genotypes across diverse agroecological conditions, high-performing genotypes, such as G189, G210, G125, G324, G269, G192, G353, G319, G219, G187 and G229, should undergo thorough evaluation in multi-location trials. Similarly, genotypes identified with both high and stable starch, protein, Fe and Zn content, including G246, G84, G4, G175, G218, G43, G161, G137, G142 and G264, are deemed promising candidates for further evaluation and integration into breeding programs, particularly for their potential to provide nutritional stability across diverse agroecological zones.

Furthermore, breeding efforts should prioritise the development of sorghum varieties demonstrating adaptability to various environmental conditions, with a specific focus on stress environments. It is recommended to incorporate identified genetic markers, including SNPs and candidate genes, into sorghum breeding programmes to expedite the development of varieties with improved nutritional quality and processing characteristics. Sensitivity analyses should be conducted to assess the stability of SNP-trait associations across diverse environmental conditions, acknowledging the crucial role of different environments in trait variation.

Fine-mapping techniques and deep sequencing are advised to pinpoint exact functional variants responsible for target traits, providing a comprehensive understanding of regulatory mechanisms and gene interactions under various environmental conditions. Additionally, the continuation of GWAS is essential to uncover additional genetic factors influencing starch, amylose, protein, ash, Fe and Zn content. Leveraging recent advances in genomics, bioinformatics and high-throughput phenotypic screening techniques is encouraged for ongoing and future sorghum breeding programs, as these cutting-edge technologies can enhance the efficiency and precision of breeding efforts.

Finally, the validation and functional characterization of identified SNPs and candidate genes are recommended to further advance the understanding and utilization of these genetic factors in sorghum improvement.

Appendices

Appendix 1. Data on monthly temperatures and rainfall for two consecutive seasons (2020 and 2021)

Year	Month	Melkassa				Jimma				Miesso			
		Rainfall	Temperature ºC			Rainfall	Temperature ºC			Rainfall	Temperature ºC		
		(mm)	Minimum	Maximum	Mean	(mm)	Minimum	Maximum	Mean	(mm)	Minimum	Maximum	Mean
2020	January	0.20	12.40	27.10	19.70	28.70	9.50	27.30	18.40	0.00	13.80	28.00	20.90
2020	February	0.00	15.00	30.20	22.60	38.40	9.50	28.90	19.20	1.20	14.80	30.40	22.60
2020	March	83.00	16.10	32.20	24.20	35.00	8.90	29.30	19.10	5.70	17.40	32.50	24.95
2020	April	84.40	16.30	31.00	23.70	33.70	10.60	28.00	19.30	176.60	18.50	30.60	24.55
2020	May	45.90	16.40	30.80	23.60	105.60	9.70	26.90	18.30	83.10	18.40	32.00	25.20
2020	June	183.50	17.40	29.90	23.70	154.20	10.30	28.00	19.10	44.30	18.90	33.70	26.30
2020	July	327.70	15.60	24.70	20.10	119.40	10.00	26.90	18.50	225.50	18.00	29.70	23.85
2020	August	304.50	15.00	23.70	19.40	150.90	10.40	26.50	18.50	198.60	18.80	29.30	24.05
2020	September	54.40	16.00	27.60	21.80	162.60	9.90	25.90	17.90	69.50	16.50	29.90	23.20
2020	October	1.50	10.70	29.70	20.20	210.20	10.10	27.30	18.70	6.00	12.70	31.80	22.25
2020	November	0.00	10.90	28.80	19.90	217.30	10.30	28.70	19.50	14.50	12.20	30.10	21.15
2020	December	0.00	9.40	28.20	18.80	273.80	9.80	27.50	18.70	0.70	10.80	29.00	19.90
2021	January	0.00	9.90	28.40	19.20	349.20	9.80	26.80	18.30	0.00	9.70	27.90	18.80
2021	February	33.70	12.30	28.50	20.40	115.40	10.60	27.50	19.10	0.00	12.80	29.60	21.20
2021	March	0.00	15.50	32.60	24.00	38.70	10.40	28.00	19.20	9.80	14.30	32.70	23.50
2021	April	76.20	15.70	32.40	24.10	3.30	10.70	28.30	19.50	101.80	17.20	32.60	24.90
2021	May	28.30	16.70	31.40	24.10	173.50	11.00	27.80	19.40	135.90	17.40	31.20	24.30
2021	June	40.20	17.10	32.50	24.80	179.00	12.10	28.30	20.20	0.00	18.70	34.00	26.35
2021	July	91.50	16.40	26.00	21.20	300.20	11.90	27.50	19.70	163.80	18.40	28.80	23.60
2021	August	238.60	15.90	27.80	21.90	182.20	11.80	26.70	19.30	95.60	17.80	29.90	23.85
2021	September	69.40	15.40	27.60	21.50	316.00	11.50	26.00	18.70	69.90	17.00	29.50	23.25
2021	October	51.70	12.00	28.60	20.30	169.10	11.10	26.10	18.60	51.90	13.80	30.60	22.20
2021	November	12.40	10.60	29.40	20.00	266.80	10.40	27.40	18.90	0.00	10.70	30.80	20.75
2021	December	0.00	10.00	28.60	19.30	166.80	10.70	28.30	19.50	-	-	-	-

Appendix 2. The stress tolerance indices of 356 sorghum landraces are based on grain yield.

Code	Genotype	Yp	Ys	TOL	MP	GMP	YSI	SSI	STI
G1	Adukar	51.91	31.54	20.37	41.72	40.46	0.61	0.88	0.42
G2	Assosa-1	57.20	33.05	24.15	45.12	43.48	0.58	0.95	0.49
G3	B35	48.79	31.72	17.07	40.25	39.34	0.65	0.78	0.40
G4	Bonsa	62.92	34.86	28.07	48.89	46.83	0.55	1.00	0.57
G5	Dagim	57.69	32.56	25.13	45.13	43.34	0.56	0.98	0.49
G6	DS10	79.56	47.12	32.45	63.34	61.23	0.59	0.91	0.97
G7	ETSL100005	62.29	28.68	33.61	45.48	42.26	0.46	1.21	0.46
G8	ETSL100007	66.01	31.16	34.85	48.59	45.35	0.47	1.18	0.53
G9	ETSL100016	74.02	38.88	35.14	56.45	53.64	0.53	1.06	0.75
G10	ETSL100020	56.82	42.45	14.36	49.63	49.11	0.75	0.57	0.62
G11	ETSL100022	62.12	34.42	27.70	48.27	46.24	0.55	1.00	0.55
G12	ETSL100023	55.81	37.26	18.55	46.54	45.60	0.67	0.75	0.54
G13	ETSL100025	58.43	32.48	25.94	45.46	43.57	0.56	1.00	0.49
G14	ETSL100030	61.54	35.36	26.17	48.45	46.65	0.57	0.95	0.56
G15	ETSL100034	53.55	33.95	19.61	43.75	42.64	0.63	0.82	0.47
G16	ETSL100039	60.93	30.13	30.79	45.53	42.85	0.49	1.13	0.48
G17	ETSL100040	58.97	34.15	24.82	46.56	44.88	0.58	0.94	0.52
G18	ETSL100044	36.00	25.59	10.41	30.80	30.35	0.71	0.65	0.24
G19	ETSL100046	69.23	31.45	37.78	50.34	46.66	0.45	1.22	0.56
G20	ETSL100048	58.63	31.94	26.69	45.28	43.27	0.54	1.02	0.49
G21	ETSL100060	61.16	32.02	29.14	46.59	44.26	0.52	1.07	0.51
G22	ETSL100062	58.91	29.32	29.59	44.12	41.56	0.50	1.13	0.45
G23	ETSL100063	61.44	29.54	31.91	45.49	42.60	0.48	1.16	0.47
G24	ETSL100064	55.67	31.43	24.24	43.55	41.83	0.56	0.98	0.45
G25	ETSL100066	55.47	27.99	27.48	41.73	39.40	0.50	1.11	0.40
G26	ETSL100068	54.17	32.00	22.17	43.09	41.64	0.59	0.92	0.45
G27	ETSL100072	57.20	30.87	26.33	44.03	42.02	0.54	1.03	0.46
G28	ETSL100078	55.78	28.90	26.89	42.34	40.15	0.52	1.08	0.42
G29	ETSL100079	68.39	34.90	33.49	51.64	48.85	0.51	1.10	0.62
G30	ETSL100090	48.75	25.63	23.13	37.19	35.35	0.53	1.06	0.32
G31	ETSL100091	63.51	34.45	29.07	48.98	46.77	0.54	1.03	0.57
G32	ETSL100093	57.66	24.90	32.76	41.28	37.89	0.43	1.27	0.37
G33	ETSL100096	61.57	28.25	33.32	44.91	41.70	0.46	1.21	0.45
G34	ETSL100100	61.35	31.24	30.11	46.29	43.78	0.51	1.10	0.50
G35	ETSL100105	64.97	34.57	30.40	49.77	47.39	0.53	1.05	0.58
G36	ETSL100107	45.02	28.60	16.42	36.81	35.88	0.64	0.82	0.33
G37	ETSL100112	54.52	34.28	20.24	44.40	43.23	0.63	0.83	0.48
G38	ETSL100120	62.42	35.84	26.59	49.13	47.30	0.57	0.96	0.58
G39	ETSL100121	57.94	33.42	24.52	45.68	44.01	0.58	0.95	0.50
G40	ETSL100127	54.70	32.72	21.98	43.71	42.31	0.60	0.90	0.46
G41	ETSL100141	62.10	23.91	38.19	43.00	38.53	0.38	1.38	0.38
G42	ETSL100149	62.12	34.42	27.70	48.27	46.24	0.55	1.00	0.55
G43	ETSL100150	60.05	38.92	21.13	49.49	48.35	0.65	0.79	0.61
G44	ETSL100152	54.76	29.66	25.09	42.21	40.30	0.54	1.03	0.42
G45	ETSL100167	58.34	30.29	28.05	44.31	42.03	0.52	1.08	0.46
G46	ETSL100169	71.26	38.96	32.30	55.11	52.69	0.55	1.02	0.72
G47	ETSL100170	69.65	39.22	30.43	54.44	52.27	0.56	0.98	0.71
G48	ETSL100178	48.49	26.37	22.12	37.43	35.76	0.54	1.02	0.33
G49	ETSL100181	60.22	36.81	23.41	48.51	47.08	0.61	0.87	0.57
G50	ETSL100184	57.89	36.93	20.96	47.41	46.24	0.64	0.81	0.55
G51	ETSL100192	51.64	27.64	24.00	39.64	37.78	0.54	1.04	0.37
G52	ETSL100198	53.60	30.41	23.19	42.01	40.38	0.57	0.97	0.42
G53	ETSL100200	67.43	35.07	32.36	51.25	48.63	0.52	1.08	0.61
G54	ETSL100204	47.74	24.06	23.68	35.90	33.89	0.50	1.11	0.30
G55	ETSL100214	64.59	28.21	36.38	46.40	42.68	0.44	1.26	0.47
G56	ETSL100233	69.23	35.55	33.68	52.39	49.61	0.51	1.09	0.64
G57	ETSL100234	56.05	29.77	26.28	42.91	40.85	0.53	1.05	0.43
G58	ETSL100238	53.07	23.78	29.29	38.43	35.53	0.45	1.24	0.33
G59	ETSL100246	43.64	27.94	15.71	35.79	34.92	0.64	0.81	0.32
G60	ETSL100249	43.88	22.84	21.04	33.36	31.66	0.52	1.08	0.26
G61	ETSL100250	62.14	27.09	35.05	44.62	41.03	0.44	1.26	0.44

G62	ETSL100267	40.58	22.91	17.68	31.74	30.49	0.56	0.98	0.24
G63	ETSL100275	47.02	33.03	13.98	40.03	39.41	0.70	0.67	0.40
G64	ETSL100284	55.12	34.19	20.93	44.66	43.41	0.62	0.85	0.49
G65	ETSL100286	75.11	37.55	37.56	56.33	53.11	0.50	1.12	0.73
G66	ETSL100292	68.73	47.08	21.65	57.90	56.88	0.69	0.71	0.84
G67	ETSL100297	50.99	34.66	16.32	42.83	42.04	0.68	0.72	0.46
G68	ETSL100303	66.31	40.22	26.09	53.26	51.64	0.61	0.88	0.69
G69	ETSL100307	80.11	42.38	37.73	61.25	58.27	0.53	1.06	0.88
G70	ETSL100310	68.04	30.83	37.21	49.43	45.80	0.45	1.23	0.54
G71	ETSL100315	83.83	40.02	43.81	61.93	57.92	0.48	1.17	0.87
G72	ETSL100328	71.12	43.88	27.24	57.50	55.86	0.62	0.86	0.81
G73	ETSL100332	54.63	33.75	20.88	44.19	42.94	0.62	0.86	0.48
G74	ETSL100340	70.52	39.22	31.30	54.87	52.59	0.56	1.00	0.72
G75	ETSL100349	50.62	35.47	15.15	43.05	42.37	0.70	0.67	0.47
G76	ETSL100350	67.18	45.90	21.28	56.54	55.53	0.68	0.71	0.80
G77	ETSL100353	55.43	32.95	22.49	44.19	42.74	0.59	0.91	0.47
G78	ETSL100365	56.08	37.97	18.11	47.02	46.14	0.68	0.72	0.55
G79	ETSL100366	62.12	34.42	27.70	48.27	46.24	0.55	1.00	0.55
G80	ETSL100367	58.27	33.51	24.75	45.89	44.19	0.58	0.95	0.51
G81	ETSL100388	43.34	26.95	16.39	35.14	34.17	0.62	0.85	0.30
G82	ETSL100392	59.17	31.46	27.71	45.32	43.15	0.53	1.05	0.48
G83	ETSL100400	75.31	42.56	32.75	58.94	56.62	0.57	0.98	0.83
G84	ETSL100406	76.82	47.50	29.32	62.16	60.41	0.62	0.86	0.95
G85	ETSL100408	63.28	35.55	27.73	49.41	47.43	0.56	0.98	0.58
G86	ETSL100409	61.65	34.46	27.19	48.06	46.09	0.56	0.99	0.55
G87	ETSL100417	56.02	35.95	20.07	45.98	44.88	0.64	0.80	0.52
G88	ETSL100419	51.09	31.26	19.83	41.17	39.96	0.61	0.87	0.41
G89	ETSL100427	63.73	41.39	22.34	52.56	51.36	0.65	0.79	0.68
G90	ETSL100428	58.59	29.89	28.70	44.24	41.84	0.51	1.10	0.45
G91	ETSL100430	57.50	34.14	23.36	45.82	44.31	0.59	0.91	0.51
G92	ETSL100431	64.28	31.99	32.29	48.13	45.35	0.50	1.13	0.53
G93	ETSL100432	56.92	34.93	21.99	45.93	44.59	0.61	0.87	0.52
G94	ETSL100436	53.89	26.97	26.93	40.43	38.13	0.50	1.12	0.38
G95	ETSL100438	54.80	30.07	24.73	42.44	40.59	0.55	1.01	0.43
G96	ETSL100443	53.25	24.88	28.37	39.07	36.40	0.47	1.19	0.34
G97	ETSL100444	75.63	32.05	43.58	53.84	49.24	0.42	1.29	0.63
G98	ETSL100445	50.48	31.48	19.01	40.98	39.86	0.62	0.84	0.41
G99	ETSL100450	63.47	33.84	29.63	48.66	46.35	0.53	1.05	0.56
G100	ETSL100465	64.34	24.85	39.50	44.59	39.98	0.39	1.38	0.41
G101	ETSL100474	62.57	35.97	26.60	49.27	47.44	0.57	0.95	0.58
G102	ETSL100476	50.03	19.85	30.18	34.94	31.51	0.40	1.35	0.26
G103	ETSL100485	53.37	24.08	29.28	38.73	35.85	0.45	1.23	0.33
G104	ETSL100488	71.08	32.15	38.94	51.61	47.80	0.45	1.23	0.59
G105	ETSL100491	54.90	19.22	35.68	37.06	32.49	0.35	1.46	0.27
G106	ETSL100492	54.42	23.75	30.67	39.09	35.95	0.44	1.26	0.33
G107	ETSL100494	75.22	32.75	42.47	53.98	49.63	0.44	1.27	0.64
G108	ETSL100502	60.46	28.81	31.65	44.63	41.73	0.48	1.17	0.45
G109	ETSL100529	74.68	45.85	28.83	60.27	58.52	0.61	0.87	0.89
G110	ETSL100539	67.02	43.32	23.70	55.17	53.88	0.65	0.79	0.75
G111	ETSL100544	47.36	33.03	14.33	40.19	39.55	0.70	0.68	0.41
G112	ETSL100550	52.25	25.79	26.46	39.02	36.71	0.49	1.14	0.35
G113	ETSL100560	57.88	34.02	23.85	45.95	44.38	0.59	0.92	0.51
G114	ETSL100561	74.46	41.09	33.37	57.77	55.31	0.55	1.01	0.79
G115	ETSL100565	70.69	36.00	34.69	53.34	50.45	0.51	1.10	0.66
G116	ETSL100568	57.17	32.50	24.66	44.84	43.11	0.57	0.97	0.48
G117	ETSL100575	61.56	35.79	25.77	48.67	46.94	0.58	0.94	0.57
G118	ETSL100584	59.84	31.60	28.24	45.72	43.49	0.53	1.06	0.49
G119	ETSL100587	57.32	30.67	26.65	44.00	41.93	0.54	1.04	0.46
G120	ETSL100593	69.03	36.56	32.47	52.80	50.24	0.53	1.05	0.65
G121	ETSL100619	77.62	49.70	27.92	63.66	62.11	0.64	0.81	1.00
G122	ETSL100620	73.09	41.01	32.08	57.05	54.75	0.56	0.98	0.78
G123	ETSL100634	61.36	31.33	30.02	46.35	43.85	0.51	1.10	0.50
G124	ETSL100638	58.29	38.74	19.55	48.52	47.52	0.66	0.75	0.59
G125	ETSL100639	88.52	53.86	34.66	71.19	69.05	0.61	0.88	1.24

G126	ETSL100642	60.99	39.88	21.11	50.43	49.31	0.65	0.78	0.63
G127	ETSL100644	70.41	41.60	28.81	56.00	54.12	0.59	0.92	0.76
G128	ETSL100648	65.11	38.96	26.14	52.03	50.37	0.60	0.90	0.66
G129	ETSL100651	62.12	34.42	27.70	48.27	46.24	0.55	1.00	0.55
G130	ETSL100661	59.33	37.61	21.72	48.47	47.23	0.63	0.82	0.58
G131	ETSL100666	68.71	44.65	24.07	56.68	55.39	0.65	0.79	0.79
G132	ETSL100667	55.80	43.87	11.93	49.84	49.48	0.79	0.48	0.63
G133	ETSL100669	52.80	29.89	22.91	41.35	39.73	0.57	0.97	0.41
G134	ETSL100683	53.88	39.68	14.20	46.78	46.24	0.74	0.59	0.55
G135	ETSL100689	64.12	37.40	26.73	50.76	48.97	0.58	0.93	0.62
G136	ETSL100699	56.62	18.27	38.34	37.44	32.16	0.32	1.52	0.27
G137	ETSL100702	42.22	27.52	14.70	34.87	34.09	0.65	0.78	0.30
G138	ETSL100707	63.29	31.34	31.94	47.31	44.54	0.50	1.13	0.51
G139	ETSL100716	55.92	32.48	23.44	44.20	42.61	0.58	0.94	0.47
G140	ETSL100720	66.20	36.14	30.06	51.17	48.91	0.55	1.02	0.62
G141	ETSL100722	41.46	29.89	11.57	35.67	35.20	0.72	0.63	0.32
G142	ETSL100735	58.26	32.44	25.82	45.35	43.48	0.56	0.99	0.49
G143	ETSL100739	83.04	45.78	37.25	64.41	61.66	0.55	1.01	0.99
G144	ETSL100745	61.43	33.70	27.73	47.57	45.50	0.55	1.01	0.54
G145	ETSL100751	60.87	32.75	28.11	46.81	44.65	0.54	1.04	0.52
G146	ETSL100761	68.00	38.65	29.35	53.32	51.27	0.57	0.97	0.68
G147	ETSL100763	62.82	30.54	32.28	46.68	43.80	0.49	1.15	0.50
G148	ETSL100768	63.33	33.35	29.97	48.34	45.96	0.53	1.06	0.55
G149	ETSL100771	58.54	33.92	24.62	46.23	44.56	0.58	0.94	0.51
G150	ETSL100782	50.18	31.88	18.30	41.03	39.99	0.64	0.82	0.41
G151	ETSL100785	73.94	35.82	38.12	54.88	51.46	0.48	1.16	0.69
G152	ETSL100786	63.55	35.30	28.25	49.42	47.36	0.56	1.00	0.58
G153	ETSL100792	62.79	37.16	25.63	49.98	48.31	0.59	0.92	0.60
G154	ETSL100797	68.99	36.00	32.99	52.49	49.83	0.52	1.07	0.64
G155	ETSL100798	64.42	38.25	26.16	51.33	49.64	0.59	0.91	0.64
G156	ETSL100801	46.77	22.14	24.64	34.45	32.18	0.47	1.18	0.27
G157	ETSL100805	46.58	29.32	17.26	37.95	36.96	0.63	0.83	0.35
G158	ETSL100808	59.54	30.97	28.57	45.26	42.94	0.52	1.08	0.48
G159	ETSL100813	73.92	35.01	38.91	54.46	50.87	0.47	1.18	0.67
G160	ETSL100825	55.60	25.89	29.72	40.74	37.94	0.47	1.20	0.37
G161	ETSL100834	59.46	28.58	30.88	44.02	41.22	0.48	1.16	0.44
G162	ETSL100837	56.60	34.63	21.96	45.62	44.27	0.61	0.87	0.51
G163	ETSL100841	53.81	33.43	20.39	43.62	42.41	0.62	0.85	0.47
G164	ETSL100842	43.19	27.91	15.28	35.55	34.72	0.65	0.79	0.31
G165	ETSL100851	62.29	29.05	33.24	45.67	42.54	0.47	1.20	0.47
G166	ETSL100861	63.83	36.52	27.31	50.17	48.28	0.57	0.96	0.60
G167	ETSL100862	72.16	37.13	35.03	54.65	51.76	0.51	1.09	0.69
G168	ETSL100863	74.79	35.82	38.97	55.31	51.76	0.48	1.17	0.69
G169	ETSL100867	57.96	31.27	26.69	44.62	42.57	0.54	1.03	0.47
G170	ETSL100872	47.12	30.55	16.58	38.83	37.94	0.65	0.79	0.37
G171	ETSL100874	48.07	28.02	20.04	38.04	36.70	0.58	0.94	0.35
G172	ETSL100876	75.14	43.75	31.39	59.44	57.33	0.58	0.94	0.85
G173	ETSL100877	50.79	30.55	20.24	40.67	39.39	0.60	0.89	0.40
G174	ETSL100888	46.60	33.55	13.05	40.08	39.54	0.72	0.63	0.41
G175	ETSL100889	60.29	32.49	27.80	46.39	44.26	0.54	1.03	0.51
G176	ETSL100890	45.92	29.21	16.71	37.56	36.62	0.64	0.82	0.35
G177	ETSL100893	62.75	35.25	27.51	49.00	47.03	0.56	0.98	0.57
G178	ETSL100895	58.07	33.48	24.59	45.77	44.09	0.58	0.95	0.50
G179	ETSL100898	67.67	34.30	33.38	50.98	48.17	0.51	1.11	0.60
G180	ETSL100900	61.70	35.90	25.81	48.80	47.06	0.58	0.94	0.57
G181	ETSL100902	73.21	44.94	28.27	59.08	57.36	0.61	0.87	0.85
G182	ETSL100916	72.48	34.21	38.26	53.35	49.80	0.47	1.18	0.64
G183	ETSL100919	73.43	38.27	35.16	55.85	53.01	0.52	1.07	0.73
G184	ETSL100921	64.39	41.49	22.90	52.94	51.69	0.64	0.80	0.69
G185	ETSL100928	71.32	37.17	34.15	54.25	51.49	0.52	1.07	0.69
G186	ETSL100945	76.88	40.96	35.92	58.92	56.11	0.53	1.05	0.82
G187	ETSL100951	88.05	48.83	39.22	68.44	65.57	0.55	1.00	1.11
G188	ETSL100953	79.77	47.34	32.43	63.56	61.45	0.59	0.91	0.98
G189	ETSL100956	94.27	54.11	40.17	74.19	71.42	0.57	0.96	1.32

G190	ETSL100964	80.46	50.71	29.75	65.59	63.88	0.63	0.83	1.06
G191	ETSL100970	62.12	34.42	27.70	48.27	46.24	0.55	1.00	0.55
G192	ETSL100977	87.01	45.78	41.22	66.40	63.12	0.53	1.06	1.03
G193	ETSL100978	55.47	32.68	22.79	44.07	42.57	0.59	0.92	0.47
G194	ETSL100983	55.55	35.32	20.23	45.44	44.30	0.64	0.82	0.51
G195	ETSL100985	53.80	31.36	22.43	42.58	41.08	0.58	0.94	0.44
G196	ETSL101000	48.13	31.20	16.93	39.67	38.75	0.65	0.79	0.39
G197	ETSL101001	64.73	35.17	29.56	49.95	47.72	0.54	1.02	0.59
G198	ETSL101006	80.38	37.70	42.68	59.04	55.05	0.47	1.19	0.79
G199	ETSL101014	70.80	36.65	34.15	53.73	50.94	0.52	1.08	0.67
G200	ETSL101015	70.40	37.00	33.40	53.70	51.04	0.53	1.06	0.67
G201	ETSL101019	83.62	47.56	36.06	65.59	63.06	0.57	0.97	1.03
G202	ETSL101021	55.11	36.16	18.95	45.63	44.64	0.66	0.77	0.52
G203	ETSL101023	78.04	43.95	34.09	61.00	58.57	0.56	0.98	0.89
G204	ETSL101029	74.89	35.06	39.84	54.98	51.24	0.47	1.19	0.68
G205	ETSL101038	59.99	33.42	26.57	46.71	44.78	0.56	0.99	0.52
G206	ETSL101051	62.22	35.71	26.51	48.97	47.14	0.57	0.96	0.58
G207	ETSL101056	62.55	36.22	26.33	49.38	47.60	0.58	0.94	0.59
G208	ETSL101058	62.16	31.01	31.15	46.58	43.90	0.50	1.12	0.50
G209	ETSL101059	59.74	32.42	27.32	46.08	44.01	0.54	1.03	0.50
G210	ETSL101061	88.20	55.52	32.69	71.86	69.98	0.63	0.83	1.27
G211	ETSL101074	74.53	34.10	40.43	54.31	50.41	0.46	1.22	0.66
G212	ETSL101081	50.30	31.24	19.07	40.77	39.64	0.62	0.85	0.41
G213	ETSL101084	60.56	39.60	20.97	50.08	48.97	0.65	0.78	0.62
G214	ETSL101097	70.21	40.31	29.90	55.26	53.20	0.57	0.96	0.73
G215	ETSL101101	58.54	30.24	28.31	44.39	42.07	0.52	1.08	0.46
G216	ETSL101110	68.20	43.04	25.16	55.62	54.18	0.63	0.83	0.76
G217	ETSL101114	53.83	29.70	24.13	41.77	39.99	0.55	1.01	0.41
G218	ETSL101118	63.24	42.06	21.17	52.65	51.57	0.67	0.75	0.69
G219	ETSL101127	86.54	50.52	36.02	68.53	66.12	0.58	0.93	1.13
G220	ETSL101128	72.40	37.28	35.13	54.84	51.95	0.51	1.09	0.70
G221	ETSL101138	62.52	37.01	25.51	49.77	48.10	0.59	0.92	0.60
G222	ETSL101142	55.55	28.13	27.42	41.84	39.53	0.51	1.11	0.40
G223	ETSL101160	61.62	33.24	28.37	47.43	45.26	0.54	1.03	0.53
G224	ETSL101161	70.27	36.89	33.38	53.58	50.92	0.53	1.07	0.67
G225	ETSL101163	47.94	27.93	20.01	37.93	36.59	0.58	0.94	0.35
G226	ETSL101173	69.08	38.84	30.24	53.96	51.80	0.56	0.98	0.70
G227	ETSL101180	73.92	36.98	36.94	55.45	52.28	0.50	1.12	0.71
G228	ETSL101181	62.51	45.04	17.47	53.78	53.06	0.72	0.63	0.73
G229	ETSL101190	87.59	48.85	38.75	68.22	65.41	0.56	0.99	1.11
G230	ETSL101199	71.22	37.00	34.22	54.11	51.34	0.52	1.08	0.68
G231	ETSL101206	61.21	34.37	26.84	47.79	45.87	0.56	0.98	0.55
G232	ETSL101210	60.65	33.88	26.77	47.26	45.33	0.56	0.99	0.53
G233	ETSL101215	52.08	34.58	17.49	43.33	42.44	0.66	0.75	0.47
G234	ETSL101217	58.49	27.21	31.28	42.85	39.89	0.47	1.20	0.41
G235	ETSL101218	53.61	25.24	28.37	39.43	36.79	0.47	1.19	0.35
G236	ETSL101224	53.68	26.71	26.98	40.20	37.87	0.50	1.13	0.37
G237	ETSL101225	47.40	24.81	22.59	36.11	34.29	0.52	1.07	0.30
G238	ETSL101231	58.25	32.74	25.50	45.49	43.67	0.56	0.98	0.49
G239	ETSL101232	45.78	26.58	19.20	36.18	34.88	0.58	0.94	0.32
G240	ETSL101246	58.88	28.89	29.99	43.88	41.24	0.49	1.14	0.44
G241	ETSL101247	54.96	30.19	24.77	42.57	40.73	0.55	1.01	0.43
G242	ETSL101248	68.68	38.43	30.25	53.55	51.37	0.56	0.99	0.68
G243	ETSL101249	56.97	26.33	30.64	41.65	38.73	0.46	1.21	0.39
G244	ETSL101254	47.51	26.75	20.76	37.13	35.65	0.56	0.98	0.33
G245	ETSL101255	71.09	36.90	34.19	54.00	51.22	0.52	1.08	0.68
G246	ETSL101256	78.24	35.83	42.42	57.04	52.95	0.46	1.22	0.73
G247	ETSL101258	59.61	25.71	33.90	42.66	39.15	0.43	1.28	0.40
G248	ETSL101262	61.80	29.12	32.69	45.46	42.42	0.47	1.19	0.47
G249	ETSL101270	67.14	36.94	30.20	52.04	49.80	0.55	1.01	0.64
G250	ETSL101274	51.17	32.93	18.24	42.05	41.05	0.64	0.80	0.44
G251	ETSL101277	48.21	26.20	22.01	37.21	35.54	0.54	1.02	0.33
G252	ETSL101289	73.34	40.13	33.21	56.74	54.25	0.55	1.02	0.76
G253	ETSL101292	48.55	33.01	15.54	40.78	40.04	0.68	0.72	0.42

G254	ETSL101293	54.73	32.10	22.62	43.42	41.92	0.59	0.93	0.46
G255	ETSL101297	46.74	28.72	18.02	37.73	36.64	0.61	0.86	0.35
G256	ETSL101299	50.08	29.85	20.23	39.97	38.66	0.60	0.91	0.39
G257	ETSL101305	75.66	44.24	31.42	59.95	57.85	0.58	0.93	0.87
G258	ETSL101310	62.06	31.89	30.17	46.97	44.48	0.51	1.09	0.51
G259	ETSL101312	67.17	39.63	27.54	53.40	51.59	0.59	0.92	0.69
G260	ETSL101313	60.92	29.37	31.55	45.14	42.30	0.48	1.16	0.46
G261	ETSL101316	57.41	24.68	32.73	41.04	37.64	0.43	1.28	0.37
G262	ETSL101321	56.96	33.68	23.28	45.32	43.80	0.59	0.92	0.50
G263	ETSL101323	54.97	24.75	30.22	39.86	36.88	0.45	1.23	0.35
G264	ETSL101325	46.32	28.54	17.79	37.43	36.36	0.62	0.86	0.34
G265	ETSL101335	47.89	21.85	26.04	34.87	32.35	0.46	1.22	0.27
G266	ETSL101337	75.51	42.52	32.99	59.02	56.67	0.56	0.98	0.83
G267	ETSL101346	67.93	40.27	27.66	54.10	52.30	0.59	0.91	0.71
G268	ETSL101347	56.19	29.71	26.48	42.95	40.86	0.53	1.06	0.43
G269	ETSL101353	87.88	46.42	41.46	67.15	63.87	0.53	1.06	1.06
G270	ETSL101358	50.31	34.12	16.19	42.21	41.43	0.68	0.72	0.44
G271	ETSL101362	66.84	30.45	36.38	48.65	45.12	0.46	1.22	0.53
G272	ETSL101369	57.11	38.82	18.29	47.97	47.09	0.68	0.72	0.57
G273	ETSL101388	74.46	41.61	32.84	58.03	55.66	0.56	0.99	0.80
G274	ETSL101389	52.36	33.62	18.74	42.99	41.95	0.64	0.80	0.46
G275	ETSL101396	72.31	42.34	29.97	57.33	55.33	0.59	0.93	0.79
G276	ETSL101404	71.79	32.11	39.68	51.95	48.01	0.45	1.24	0.60
G277	ETSL101408	51.79	27.06	24.73	39.42	37.44	0.52	1.07	0.36
G278	ETSL101410	61.12	29.87	31.25	45.50	42.73	0.49	1.15	0.47
G279	ETSL101416	59.20	25.57	33.63	42.39	38.91	0.43	1.27	0.39
G280	ETSL101418	77.72	41.62	36.10	59.67	56.87	0.54	1.04	0.84
G281	ETSL101431	58.51	32.41	26.10	45.46	43.55	0.55	1.00	0.49
G282	ETSL101436	73.44	35.23	38.21	54.33	50.86	0.48	1.17	0.67
G283	ETSL101443	68.98	37.44	31.55	53.21	50.82	0.54	1.03	0.67
G284	ETSL101444	71.15	39.65	31.50	55.40	53.11	0.56	0.99	0.73
G285	ETSL101445	57.78	33.18	24.60	45.48	43.79	0.57	0.95	0.50
G286	ETSL101446	65.09	35.62	29.47	50.36	48.15	0.55	1.02	0.60
G287	ETSL101451	50.59	26.75	23.84	38.67	36.79	0.53	1.06	0.35
G288	ETSL101453	50.73	25.91	24.82	38.32	36.26	0.51	1.10	0.34
G289	ETSL101459	58.48	28.62	29.85	43.55	40.91	0.49	1.14	0.43
G290	ETSL101461	55.72	31.01	24.72	43.37	41.57	0.56	0.99	0.45
G291	ETSL101464	81.43	42.18	39.25	61.80	58.61	0.52	1.08	0.89
G292	ETSL101470	76.13	42.84	33.29	59.48	57.11	0.56	0.98	0.85
G293	ETSL101477	86.20	43.89	42.32	65.05	61.51	0.51	1.10	0.98
G294	ETSL101486	45.67	30.94	14.73	38.30	37.59	0.68	0.72	0.37
G295	ETSL101491	47.35	27.51	19.84	37.43	36.09	0.58	0.94	0.34
G296	ETSL101492	38.82	27.39	11.44	33.11	32.61	0.71	0.66	0.28
G297	ETSL101506	61.05	32.24	28.82	46.64	44.36	0.53	1.06	0.51
G298	ETSL101508	66.88	32.43	34.45	49.66	46.57	0.48	1.15	0.56
G299	ETSL101512	51.28	34.63	16.65	42.96	42.14	0.68	0.73	0.46
G300	ETSL101515	73.28	40.08	33.20	56.68	54.19	0.55	1.02	0.76
G301	ETSL101523	54.99	28.78	26.21	41.89	39.78	0.52	1.07	0.41
G302	ETSL101524	60.91	29.94	30.97	45.42	42.70	0.49	1.14	0.47
G303	ETSL101527	65.60	32.41	33.19	49.00	46.11	0.49	1.13	0.55
G304	ETSL101535	56.28	33.80	22.48	45.04	43.61	0.60	0.90	0.49
G305	ETSL101536	50.34	29.03	21.31	39.68	38.23	0.58	0.95	0.38
G306	ETSL101540	64.43	33.45	30.98	48.94	46.43	0.52	1.08	0.56
G307	ETSL101550	72.15	41.06	31.09	56.60	54.43	0.57	0.97	0.77
G308	ETSL101555	57.27	30.32	26.95	43.79	41.67	0.53	1.06	0.45
G309	ETSL101556	60.50	30.65	29.84	45.58	43.06	0.51	1.11	0.48
G310	ETSL101559	44.01	23.27	20.73	33.64	32.00	0.53	1.06	0.27
G311	ETSL101561	47.02	26.78	20.24	36.90	35.48	0.57	0.97	0.33
G312	ETSL101563	50.18	30.87	19.31	40.53	39.36	0.62	0.86	0.40
G313	ETSL101564	61.59	31.09	30.50	46.34	43.76	0.50	1.11	0.50
G314	ETSL101568	69.31	35.39	33.91	52.35	49.53	0.51	1.10	0.64
G315	ETSL101571	59.86	36.20	23.66	48.03	46.55	0.60	0.89	0.56
G316	ETSL101578	65.82	37.15	28.67	51.49	49.45	0.56	0.98	0.63
G317	ETSL101586	61.14	32.45	28.69	46.79	44.54	0.53	1.05	0.51

G318	ETSL101629	81.39	40.64	40.74	61.01	57.51	0.50	1.12	0.86
G319	ETSL101640	89.66	48.51	41.15	69.08	65.95	0.54	1.03	1.13
G320	ETSL101643	69.95	38.99	30.97	54.47	52.22	0.56	0.99	0.71
G321	ETSL101654	70.23	39.50	30.72	54.87	52.67	0.56	0.98	0.72
G322	ETSL101671	70.13	40.86	29.27	55.49	53.53	0.58	0.94	0.74
G323	ETSL101673	70.70	35.43	35.27	53.07	50.05	0.50	1.12	0.65
G324	ETSL101676	92.49	47.88	44.61	70.19	66.55	0.52	1.08	1.15
G325	ETSL101681	71.05	38.91	32.14	54.98	52.58	0.55	1.01	0.72
G326	ETSL101686	68.15	36.92	31.23	52.54	50.16	0.54	1.03	0.65
G327	ETSL101687	49.80	27.28	22.52	38.54	36.86	0.55	1.01	0.35
G328	ETSL101698	65.84	30.20	35.63	48.02	44.59	0.46	1.21	0.52
G329	ETSL101704	53.38	25.93	27.45	39.65	37.20	0.49	1.15	0.36
G330	ETSL101712	58.19	30.28	27.91	44.23	41.97	0.52	1.08	0.46
G331	ETSL101715	61.77	36.46	25.31	49.11	47.45	0.59	0.92	0.58
G332	ETSL101719	83.75	43.81	39.95	63.78	60.57	0.52	1.07	0.95
G333	ETSL101722	69.81	40.43	29.39	55.12	53.13	0.58	0.94	0.73
G334	ETSL101723	68.06	45.14	22.92	56.60	55.43	0.66	0.76	0.80
G335	ETSL101725	67.19	38.18	29.01	52.69	50.65	0.57	0.97	0.66
G336	ETSL101730	60.25	29.92	30.33	45.09	42.46	0.50	1.13	0.47
G337	ETSL101732	59.71	35.81	23.90	47.76	46.24	0.60	0.90	0.55
G338	ETSL101749	44.92	32.93	11.99	38.92	38.46	0.73	0.60	0.38
G339	ETSL101760	85.88	42.33	43.55	64.10	60.29	0.49	1.14	0.94
G340	ETSL101766	58.36	34.58	23.78	46.47	44.92	0.59	0.91	0.52
G341	ETSL101848	51.38	22.14	29.24	36.76	33.73	0.43	1.28	0.29
G342	ETSL101851	83.12	43.44	39.69	63.28	60.09	0.52	1.07	0.94
G343	ETSL101856	66.44	39.21	27.22	52.82	51.04	0.59	0.92	0.68
G344	ETSL101858	80.06	44.42	35.64	62.24	59.63	0.55	1.00	0.92
G345	ETSL101868	54.49	30.41	24.09	42.45	40.71	0.56	0.99	0.43
G346	IS25531	58.91	29.65	29.27	44.28	41.79	0.50	1.11	0.45
G347	IS38254	60.34	42.25	18.09	51.30	50.49	0.70	0.67	0.66
G348	IS38255	72.43	38.19	34.24	55.31	52.60	0.53	1.06	0.72
G349	IS38259	60.52	39.70	20.82	50.11	49.01	0.66	0.77	0.62
G350	IS38275	62.12	34.42	27.70	48.27	46.24	0.55	1.00	0.55
G351	IS38279	76.52	28.57	47.95	52.55	46.76	0.37	1.41	0.57
G352	IS38290	62.12	34.42	27.70	48.27	46.24	0.55	1.00	0.55
G353	IS38303	87.32	51.38	35.93	69.35	66.98	0.59	0.92	1.16
G354	IS38306	53.73	31.47	22.26	42.60	41.12	0.59	0.93	0.44
G355	IS38319	66.76	38.37	28.38	52.56	50.61	0.57	0.95	0.66
G356	IS38326	44.81	26.74	18.06	35.78	34.62	0.60	0.90	0.31
G357	IS38331	49.03	29.30	19.73	39.16	37.90	0.60	0.90	0.37
G358	IS38353	54.15	34.85	19.30	44.50	43.44	0.64	0.80	0.49
G359	IS38412	72.01	36.02	35.99	54.01	50.92	0.50	1.12	0.67
G360	P9830	52.50	31.85	20.65	42.17	40.89	0.61	0.88	0.43
G361	PML981446	54.44	30.71	23.74	42.57	40.89	0.56	0.98	0.43
G362	PML981474	60.12	37.90	22.22	49.01	47.73	0.63	0.83	0.59
G363	PML981476	75.58	41.83	33.75	58.71	56.23	0.55	1.00	0.82
G364	PML981488	77.22	44.88	32.34	61.05	58.87	0.58	0.94	0.90
G365	TAM428	61.66	37.53	24.13	49.59	48.10	0.61	0.88	0.60

Yield in non-stress conditions (Yp); Yield in stress conditions (Ys); Tolerance index (TOL); Mean productivity (MP); Geometric Mean Productivity (GMP); Yield stability index (YSI); Stress susceptibility index (SSI); Stress tolerance index (STI).

Appendix 3. Mean (BLUP) quality traits performance of the 365 sorghum landraces across four environments.

Code	Genotypes	Protein	Starch	AM	AMP	Fe	Zn	Ash	MST
G1	Adukar	8.57	68.36	19.1	80.9	35.76	37.09	2.91	15.57
G2	Assosa-1	8.69	68.25	19.16	80.84	33.65	33.2	2.27	15.58
G3	B35	8.18	69.41	19.68	80.32	24.2	26.45	1.58	15.26
G4	Bonsa	8.4	69.65	19.84	80.16	25.69	23.33	1.29	15.36
G5	Dagim	9.07	69.09	19.76	80.24	20.74	24.14	1.31	15.45
G6	DS10	9.67	69.11	19.71	80.29	12.08	24.87	1.48	15.5
G7	ETSL100005	9.64	68.86	19.55	80.46	49.17	29.14	1.88	15.48
G8	ETSL100007	8.26	68.77	19.54	80.46	38.57	33.75	2.3	15.61
G9	ETSL100016	9.05	68.83	19.74	80.26	53.29	32.09	2.05	15.48
G10	ETSL100020	9.17	68.59	19.56	80.45	38.25	28.8	1.89	15.74
G11	ETSL100022	9.56	68.74	19.6	80.4	32.93	28.06	1.81	15.5
G12	ETSL100023	10.27	68.36	19.14	80.86	30.58	28.67	2.08	15.77
G13	ETSL100025	8.17	68.91	19.43	80.57	39.92	32.65	2.16	15.63
G14	ETSL100030	9.44	69	19.76	80.24	31.31	24.87	1.45	15.4
G15	ETSL100034	8.75	69.11	19.47	80.53	28.94	27.27	1.67	15.63
G16	ETSL100039	9.18	67.96	19.71	80.29	51.48	34.23	2.11	15.63
G17	ETSL100040	9.17	68.86	19.49	80.52	46.91	29.92	1.93	15.48
G18	ETSL100044	9.66	68.79	19.75	80.25	13.05	23.78	1.55	15.33
G19	ETSL100046	9.09	68.38	19.34	80.67	30.82	31.31	2	15.71
G20	ETSL100048	9.41	68.49	19.46	80.54	25.31	26.91	1.69	15.72
G21	ETSL100060	9.78	68.8	19.67	80.33	40.23	25.72	1.58	15.53
G22	ETSL100062	9.52	68.67	19.28	80.73	39.97	27.46	1.96	15.95
G23	ETSL100063	8.3	68.81	19.42	80.58	57.84	31.96	2.45	15.49
G24	ETSL100064	8.83	68.78	19.45	80.55	33.8	29.96	1.86	15.43
G25	ETSL100066	8.15	68.32	19.22	80.78	35.15	32.05	2.2	15.85
G26	ETSL100068	8.76	68.28	19.45	80.55	40.26	34.44	2.29	15.72
G27	ETSL100072	9.35	68.59	19.39	80.61	27.64	27.49	2.01	15.94
G28	ETSL100078	9.5	68.91	19.58	80.42	57.02	29.71	2.01	15.37
G29	ETSL100079	8.72	68.68	19.46	80.55	37.44	29.85	1.94	15.77
G30	ETSL100090	8.67	68.26	19.43	80.57	39.55	29.61	1.91	16.03
G31	ETSL100091	8.89	68.6	19.45	80.55	21.38	28.79	2.3	15.82
G32	ETSL100093	10.29	68.49	19.73	80.28	34.34	25.02	1.52	15.47
G33	ETSL100096	9.28	68.96	19.78	80.22	19.43	25.4	1.73	15.52
G34	ETSL100100	8.62	68.86	19.49	80.51	19.3	28.72	2.08	15.99
G35	ETSL100105	9.78	68.7	19.28	80.72	38.92	30.09	2.13	15.58
G36	ETSL100107	8.95	68.82	19.44	80.56	40.01	28.62	1.96	15.5
G37	ETSL100112	9.16	68.97	19.45	80.55	41.05	28.99	1.92	15.45
G38	ETSL100120	10.28	68.63	19.39	80.61	29.3	27.23	1.82	15.71
G39	ETSL100121	10.07	68.29	19.37	80.63	35.75	28.97	1.72	15.51
G40	ETSL100127	9.25	69.22	19.49	80.51	30.86	27.67	1.74	15.65

G41	ETSL100141	8.9	69.33	19.64	80.36	46.91	30.87	1.91	15.52
G42	ETSL100149	9.56	68.74	19.6	80.4	32.93	28.06	1.81	15.5
G43	ETSL100150	10.35	68.53	19.62	80.39	16.97	25.79	1.79	15.52
G44	ETSL100152	10.41	68.76	19.5	80.51	26.72	23.96	1.55	15.64
G45	ETSL100167	9.65	68.6	19.62	80.38	48.09	29.61	2.01	15.46
G46	ETSL100169	9.5	68.37	19.09	80.91	33.77	33.22	2.38	15.64
G47	ETSL100170	8.54	69.2	19.6	80.41	33.2	27.52	1.82	15.55
G48	ETSL100178	9.68	67.93	19.04	80.97	34.54	31.46	2.41	15.91
G49	ETSL100181	9.91	68.47	19.29	80.71	34.72	29.5	1.97	15.95
G50	ETSL100184	9.91	68.52	19.59	80.41	19.13	25.88	1.71	15.64
G51	ETSL100192	9.44	68.39	19.45	80.55	24.75	28.49	1.82	15.63
G52	ETSL100198	9.85	68.36	19.35	80.66	34.79	27.32	1.78	15.79
G53	ETSL100200	9.23	68.84	19.67	80.33	33.81	27.87	1.59	15.78
G54	ETSL100204	10.36	68.46	19.79	80.21	26.84	25.6	1.38	15.59
G55	ETSL100214	9.25	69.24	19.66	80.34	59.69	28.8	2.13	15.31
G56	ETSL100233	9.4	68.96	19.82	80.18	35.44	27.47	1.66	14.89
G57	ETSL100234	9.43	68.47	19.59	80.41	47.74	30.85	2.04	15.83
G58	ETSL100238	10.05	68.23	19.25	80.75	25.7	28.07	1.89	15.81
G59	ETSL100246	9.16	68.93	19.59	80.41	42.83	24.77	1.75	15.57
G60	ETSL100249	10.21	68.29	19.38	80.62	30.98	27.35	1.6	15.52
G61	ETSL100250	9.18	68.96	19.52	80.48	49.84	30.65	2.11	15.17
G62	ETSL100267	9.94	68.13	19.27	80.73	25.07	29.62	1.71	15.74
G63	ETSL100275	8.09	69.47	19.61	80.39	40.77	29.64	1.7	15.5
G64	ETSL100284	9.27	68.79	19.53	80.48	15.55	24.93	1.49	15.47
G65	ETSL100286	9.25	69.01	19.5	80.5	43.11	26.84	1.91	15.52
G66	ETSL100292	9.52	68.8	19.57	80.43	24.08	26.79	1.87	15.59
G67	ETSL100297	10.07	68.68	19.67	80.34	28.47	25.54	1.5	15.56
G68	ETSL100303	9.89	69.04	19.7	80.3	36.32	26.28	1.59	15.54
G69	ETSL100307	9.18	69.16	19.55	80.45	20.55	31.86	1.91	15.74
G70	ETSL100310	9.69	68.92	19.57	80.43	37.08	25.32	1.63	15.52
G71	ETSL100315	9.33	69.06	19.53	80.47	27.69	27.62	1.93	15.43
G72	ETSL100328	8.85	69.15	19.74	80.26	43.22	28.86	1.72	15.32
G73	ETSL100332	9.29	68.43	19.27	80.73	36.39	30.93	2.18	15.66
G74	ETSL100340	10.02	68.76	19.55	80.45	39.72	26.67	1.93	15.53
G75	ETSL100349	10.08	68.18	19.61	80.39	28.67	26.26	1.33	15.71
G76	ETSL100350	10.17	68.11	19.76	80.25	23.61	26.55	1.46	15.46
G77	ETSL100353	8.33	69.09	19.6	80.4	28.09	30.47	1.85	15.25
G78	ETSL100365	10.2	68.77	19.55	80.45	23.52	23.77	1.28	15.64
G79	ETSL100366	9.56	68.74	19.6	80.4	32.93	28.06	1.81	15.5
G80	ETSL100367	9.44	68.95	19.5	80.5	21.46	26.29	1.81	15.6
G81	ETSL100388	9.19	68.49	19.4	80.61	36.6	29.12	1.97	15.5
G82	ETSL100392	10.06	68.54	19.51	80.49	19.69	27.56	1.8	15.5
G83	ETSL100400	9.33	68.97	19.6	80.4	41.13	28.79	1.86	15.22
G84	ETSL100406	9.44	69.56	19.83	80.18	30.63	25.64	1.53	15.21

G85	ETSL100408	10.17	68.45	19.68	80.33	27.41	34.98	2.07	15.72
G86	ETSL100409	10	68.57	19.52	80.48	22.86	27.15	1.89	15.52
G87	ETSL100417	9.45	69.26	19.6	80.4	57.79	32.84	2.31	15.3
G88	ETSL100419	10.75	68.44	19.38	80.62	32.04	24.56	1.85	15.69
G89	ETSL100427	10.09	68.96	19.66	80.34	20.69	25.86	1.71	15.47
G90	ETSL100428	9.27	67.88	19.61	80.39	29.97	26.73	1.66	15.97
G91	ETSL100430	10.55	68.87	19.47	80.53	26.42	24.41	1.59	15.83
G92	ETSL100431	9.82	68.6	19.73	80.27	19.98	26.03	1.66	15.58
G93	ETSL100432	9.36	69.1	19.72	80.29	21.86	24.54	1.53	15.43
G94	ETSL100436	7.99	68.74	19.5	80.5	43.66	32.97	2.18	15.5
G95	ETSL100438	9.57	68.62	19.67	80.33	40.82	28.51	1.58	15.49
G96	ETSL100443	9.33	69.11	19.84	80.16	23.76	23.7	1.22	15.58
G97	ETSL100444	9.73	68.87	19.82	80.18	30.9	24.87	1.39	15.32
G98	ETSL100445	8.88	68.69	19.41	80.6	39.18	32.77	2.39	15.6
G99	ETSL100450	8.96	68.88	19.67	80.32	52.52	29.15	2.03	15.44
G100	ETSL100465	8.84	68.68	19.51	80.49	29.95	28.64	2.11	15.72
G101	ETSL100474	9.11	69.09	19.5	80.5	23.81	24.94	1.71	15.77
G102	ETSL100476	9.63	68.79	19.73	80.27	18.97	23.97	1.39	15.63
G103	ETSL100485	9.57	68.89	19.75	80.25	34.34	24.74	1.5	15.64
G104	ETSL100488	9.77	68.88	19.74	80.27	35.76	25.39	1.55	15.2
G105	ETSL100491	9.12	68.42	19.54	80.46	25.22	28.95	1.88	15.81
G106	ETSL100492	8.59	68.8	19.88	80.12	35.43	27.94	1.58	15.93
G107	ETSL100494	10.07	68.55	19.68	80.32	34.68	25.52	1.57	15.5
G108	ETSL100502	9.75	68.95	19.81	80.2	32.33	24.45	1.54	15.21
G109	ETSL100529	9.87	68.36	19.53	80.47	22.43	28.77	1.81	15.27
G110	ETSL100539	9.35	68.4	19.6	80.4	28.55	29.86	1.8	15.31
G111	ETSL100544	10.28	68.57	19.59	80.41	20.95	22.99	1.34	15.59
G112	ETSL100550	10.06	68.95	19.63	80.37	19.19	26.51	1.91	15.42
G113	ETSL100560	10.25	68.62	19.73	80.27	32.37	25.82	1.44	15.7
G114	ETSL100561	9.61	68.88	19.69	80.31	35.9	25.11	1.65	15.51
G115	ETSL100565	8.7	68.97	19.74	80.26	45.42	26.49	1.75	15.44
G116	ETSL100568	8.4	68.86	19.85	80.15	45.53	34.74	2.07	15.35
G117	ETSL100575	9.01	69.06	19.85	80.15	47.98	30.6	2.04	15.31
G118	ETSL100584	9.54	68.61	19.27	80.73	29.93	27.56	1.83	15.53
G119	ETSL100587	8.53	68.91	19.71	80.3	49.33	32.25	2.03	14.9
G120	ETSL100593	9.64	69.13	19.68	80.32	38.1	26.16	1.81	15.28
G121	ETSL100619	9.97	68.92	19.69	80.31	14.9	25.51	1.53	15.58
G122	ETSL100620	9.24	69.31	19.98	80.02	22.93	24.11	1.28	15.31
G123	ETSL100634	9.01	69.41	19.79	80.21	45.2	26.43	1.79	15.14
G124	ETSL100638	9.19	69.11	19.63	80.37	37.82	25.96	1.76	15.46
G125	ETSL100639	9.65	68.81	19.47	80.53	20.39	25.84	1.55	15.94
G126	ETSL100642	9.45	68.49	19.32	80.68	26.3	29.62	1.76	15.78
G127	ETSL100644	10.05	68.75	19.8	80.21	22.12	25.84	1.39	15.48
G128	ETSL100648	10.55	68.85	19.61	80.4	16.77	24.62	1.57	15.8

G129	ETSL100651	9.56	68.74	19.6	80.4	32.93	28.06	1.81	15.5
G130	ETSL100661	10.24	68.6	19.75	80.25	23.69	24.77	1.45	15.57
G131	ETSL100666	9.64	69.05	19.65	80.35	35.39	26.53	1.81	15.49
G132	ETSL100667	10.38	68.65	19.64	80.36	15.24	25.7	1.64	15.32
G133	ETSL100669	9.29	68.4	19.29	80.71	36.29	32.09	2.13	15.48
G134	ETSL100683	9.12	68.93	19.82	80.18	23.59	25.55	1.44	15.48
G135	ETSL100689	9.38	68.98	19.84	80.16	49.07	31.36	1.99	14.6
G136	ETSL100699	9.49	68.98	19.73	80.27	29.4	23.98	1.39	15.59
G137	ETSL100702	9.82	69.3	19.88	80.12	79.56	38.1	2.61	15.41
G138	ETSL100707	9.9	68.86	19.78	80.22	38.82	24.88	1.66	15.6
G139	ETSL100716	9.27	68.95	19.78	80.22	23.14	23.65	1.23	15.22
G140	ETSL100720	9.41	68.84	19.66	80.34	52.69	34.46	2.2	15.22
G141	ETSL100722	10.52	68.52	19.65	80.36	40.95	24.52	1.47	15.52
G142	ETSL100735	7.97	69.13	20.47	79.53	80.77	50.03	2.95	15.73
G143	ETSL100739	10.77	68.25	19.52	80.48	39.9	27.13	1.68	15.55
G144	ETSL100745	9.67	69	19.6	80.4	48.49	31.35	2.09	15.37
G145	ETSL100751	10.12	68.96	19.63	80.37	38.08	25.38	1.79	15.24
G146	ETSL100761	10.26	68.42	19.54	80.45	49.66	33.99	2.19	15.43
G147	ETSL100763	9.59	68.54	19.57	80.43	38.71	29.81	1.77	15.27
G148	ETSL100768	9.2	68.41	20.02	79.98	53.51	44.31	2.96	15.3
G149	ETSL100771	10.13	68.88	19.73	80.27	57.45	34.07	2.21	15.08
G150	ETSL100782	10.13	68.27	19.66	80.34	32.26	24.49	1.56	15.45
G151	ETSL100785	10.3	69	19.5	80.5	31.5	25.2	1.79	15.67
G152	ETSL100786	9.26	69.14	19.76	80.24	36.06	24.34	1.65	15.3
G153	ETSL100792	10.3	68.39	19.5	80.5	24.57	25.97	1.9	15.03
G154	ETSL100797	10.12	68.59	19.65	80.35	41.2	26.48	1.65	15.5
G155	ETSL100798	9.1	68.82	19.61	80.39	33.5	26.08	1.57	15.43
G156	ETSL100801	9.55	69.05	19.73	80.27	18.53	25.35	1.76	15.47
G157	ETSL100805	10.1	68.69	19.56	80.44	35.51	25.24	1.88	15.27
G158	ETSL100808	9.5	68.98	19.72	80.28	39.17	24.47	1.73	15.22
G159	ETSL100813	10.51	68.61	19.56	80.44	29.97	24.76	1.65	15.4
G160	ETSL100825	9	69.31	19.74	80.27	33.97	24.56	1.42	15.38
G161	ETSL100834	10.99	68.15	19.63	80.37	34.58	26.37	1.53	15.49
G162	ETSL100837	9.91	69.38	19.65	80.35	33.28	25.33	1.65	15.08
G163	ETSL100841	8.8	68.38	19.96	80.04	29.39	34.12	1.78	14.46
G164	ETSL100842	9	68.15	19.59	80.41	47.24	33.9	1.96	15.66
G165	ETSL100851	9.07	68.89	19.73	80.27	38.38	26.93	1.61	15.55
G166	ETSL100861	9.84	69.01	19.73	80.27	63.09	34.66	2.45	15.13
G167	ETSL100862	9.25	69.25	19.8	80.2	27.36	24.56	1.43	15.33
G168	ETSL100863	8.78	68.93	19.68	80.32	40.98	28.96	1.88	15.13
G169	ETSL100867	9.03	68.32	19.67	80.33	29.61	29.17	1.51	15.61
G170	ETSL100872	9.23	68.77	19.64	80.36	38.12	29.74	1.78	15.26
G171	ETSL100874	10.95	68.85	19.65	80.35	24.73	25.16	1.57	15.36
G172	ETSL100876	9.49	69.12	19.87	80.13	59.99	27.76	1.97	15.38

G173	ETSL100877	9.84	67.98	19.42	80.58	30.19	31.16	2.14	15.95
G174	ETSL100888	9.74	68.19	19.76	80.24	45.16	30.43	1.94	15.57
G175	ETSL100889	8.61	69.55	20.3	79.7	52.71	46.47	2.51	15.04
G176	ETSL100890	10.3	68.51	19.6	80.4	34.11	27.58	1.65	15.62
G177	ETSL100893	9.42	68.91	19.67	80.33	20.19	25.39	1.71	15.39
G178	ETSL100895	8.68	69.2	19.48	80.52	38.66	29.04	1.95	15.47
G179	ETSL100898	9.1	69.08	19.78	80.22	44.66	26.46	1.63	15.41
G180	ETSL100900	10.16	68.63	19.53	80.47	51.72	32.75	2.19	15.43
G181	ETSL100902	9.97	68.73	19.63	80.37	30.78	31.16	2.17	15.35
G182	ETSL100916	8.88	68.71	19.52	80.48	50.6	30.73	2.04	15.6
G183	ETSL100919	9.14	68.66	19.29	80.71	30.97	26.86	1.79	15.9
G184	ETSL100921	10.84	68.58	19.21	80.79	20.92	27.03	1.84	16.08
G185	ETSL100928	10.11	68.65	19.54	80.46	27.16	23.7	1.5	15.5
G186	ETSL100945	9.7	69.19	19.55	80.45	25.57	23.49	1.52	15.68
G187	ETSL100951	9.06	68.78	19.5	80.5	30.19	27.39	1.87	15.62
G188	ETSL100953	10.24	68.45	19.51	80.5	22.62	26.16	1.92	15.62
G189	ETSL100956	9.6	68.85	19.58	80.42	24.65	24.65	1.72	15.75
G190	ETSL100964	9.74	68.5	19.44	80.56	29.3	30.92	2.02	15.37
G191	ETSL100970	9.56	68.74	19.6	80.4	32.93	28.06	1.81	15.5
G192	ETSL100977	10.04	68.63	19.41	80.59	28.18	29.94	2.01	15.82
G193	ETSL100978	8.61	68.83	19.74	80.26	52.28	32.59	1.82	15.44
G194	ETSL100983	8.72	69.44	19.68	80.32	47.88	29.88	2.02	15.22
G195	ETSL100985	10.2	68.29	19.59	80.41	34.23	27.53	1.64	15.18
G196	ETSL101000	10.26	68.36	19.48	80.53	27.28	26.37	1.64	15.51
G197	ETSL101001	9.84	68.65	19.74	80.26	18.85	26.89	1.76	15.61
G198	ETSL101006	9.12	68.93	19.69	80.31	16.25	26.1	1.74	15.75
G199	ETSL101014	9.66	68.57	19.42	80.58	20.3	30.04	1.81	15.38
G200	ETSL101015	9.64	68.62	19.59	80.41	19.58	27.48	1.89	15.55
G201	ETSL101019	9.37	68.79	19.69	80.31	23.35	25.65	1.6	15.3
G202	ETSL101021	7.69	69.19	19.8	80.2	23.96	30.88	2.06	15.41
G203	ETSL101023	9.94	68.97	19.46	80.54	20.29	25.51	1.76	15.77
G204	ETSL101029	9.21	68.74	19.58	80.42	20.42	27.54	1.79	15.74
G205	ETSL101038	9.7	68.88	19.61	80.39	19.11	24.25	1.66	15.74
G206	ETSL101051	8.53	68.52	19.37	80.63	45.24	37.23	2.39	15.66
G207	ETSL101056	8.27	67.92	19.66	80.34	30.75	28.64	1.66	14.44
G208	ETSL101058	9.5	68.48	19.31	80.69	21.22	28.26	1.88	16.01
G209	ETSL101059	9.79	68.39	19.54	80.46	23.25	26.5	1.72	15.55
G210	ETSL101061	9.18	68.67	19.67	80.33	33.63	27.52	1.93	15.5
G211	ETSL101074	9.7	69.02	19.65	80.35	33.79	25.84	1.55	15.56
G212	ETSL101081	10.39	67.63	19.47	80.53	50.16	34.68	2.26	15.56
G213	ETSL101084	10.02	68.85	19.61	80.39	19.87	25.09	1.8	15.71
G214	ETSL101097	10.64	68.46	19.58	80.43	31.99	25.08	1.78	15.55
G215	ETSL101101	9.25	68.72	19.71	80.29	42.98	28.93	1.72	15.17
G216	ETSL101110	9.73	68.92	19.64	80.36	24.24	25.5	1.46	15.48

G217	ETSL101114	10.23	68.84	19.68	80.32	35.04	23.84	1.6	15.28
G218	ETSL101118	10.84	68.47	19.48	80.52	20.17	27.11	1.76	15.43
G219	ETSL101127	9.27	69.05	19.61	80.39	21.47	25.47	1.72	15.73
G220	ETSL101128	10.01	68.4	19.53	80.47	15.77	26.66	1.88	15.86
G221	ETSL101138	10.86	68.35	19.39	80.62	22.27	27.53	2.05	15.67
G222	ETSL101142	10.22	68.58	19.68	80.32	37.16	26.05	1.54	15.47
G223	ETSL101160	9.72	68.87	19.68	80.32	35.55	25.74	1.61	15.55
G224	ETSL101161	9.09	68.92	19.65	80.35	25.52	26.49	1.89	15.32
G225	ETSL101163	10.49	68.08	19.25	80.75	20.41	28.91	2.06	16.02
G226	ETSL101173	9.59	68.9	19.41	80.59	29.58	26.93	1.68	15.66
G227	ETSL101180	9.96	68.33	19.54	80.46	19.09	25.71	1.74	15.38
G228	ETSL101181	9.89	68.28	19.47	80.53	25.18	28.93	1.7	15.74
G229	ETSL101190	10.06	68.78	19.48	80.52	28.12	26.43	1.8	15.58
G230	ETSL101199	8.61	68.8	19.96	80.04	40.26	38.98	2.36	14.66
G231	ETSL101206	10.52	68.61	19.57	80.43	24.31	26.52	1.67	15.47
G232	ETSL101210	10.1	68.47	19.63	80.37	29.83	28.22	1.81	15.54
G233	ETSL101215	10.17	68.43	19.71	80.29	35.32	25.2	1.77	15.43
G234	ETSL101217	9.13	69.1	19.82	80.18	34.06	24.92	1.45	15.34
G235	ETSL101218	8.7	69.25	19.71	80.29	23.79	27.83	1.66	15.34
G236	ETSL101224	9.82	68.78	19.73	80.27	46.49	27.14	1.84	15.02
G237	ETSL101225	9.66	68.85	19.62	80.39	33.69	27.99	1.8	15.78
G238	ETSL101231	10.04	68.7	19.86	80.15	21.33	25.53	1.45	15.35
G239	ETSL101232	9.04	68.56	19.46	80.54	47.24	34.7	2.06	15.76
G240	ETSL101246	9.43	69.13	19.66	80.34	30.2	26.31	1.47	15.6
G241	ETSL101247	9.86	68.74	19.63	80.37	28.23	31.82	1.88	15.37
G242	ETSL101248	9.41	68.9	19.6	80.4	36.63	26.38	1.58	15.43
G243	ETSL101249	10.07	68.49	19.41	80.59	49.05	28.85	1.94	15.43
G244	ETSL101254	9.98	68.94	19.88	80.12	29.8	24.11	1.23	15.58
G245	ETSL101255	8.81	68.15	19.74	80.26	45.65	30.96	1.89	14.35
G246	ETSL101256	8.98	69.6	19.83	80.18	40.14	25.39	1.68	15.32
G247	ETSL101258	9.56	69.12	19.7	80.31	38.03	24.7	1.56	15.65
G248	ETSL101262	8.83	69.28	19.79	80.21	62.91	30.14	2.16	15.21
G249	ETSL101270	9.28	69.33	19.74	80.26	35.35	24.41	1.64	15.32
G250	ETSL101274	9.39	69.09	19.79	80.21	32.78	24.46	1.59	15.34
G251	ETSL101277	9.11	68.35	19.57	80.43	40.88	31.43	2.09	15.13
G252	ETSL101289	9.96	68.65	19.39	80.61	28.48	28.73	1.8	15.66
G253	ETSL101292	9.46	69.26	19.6	80.4	31.38	29.67	1.65	15.42
G254	ETSL101293	9.86	68.86	19.63	80.37	36.8	25.26	1.68	15.47
G255	ETSL101297	9.6	68.61	19.62	80.37	18.54	26.62	1.55	15.52
G256	ETSL101299	9.3	69.13	19.61	80.39	23.61	26.07	1.49	15.36
G257	ETSL101305	8.88	69.24	19.69	80.31	38.47	26.8	1.63	15.51
G258	ETSL101310	10.11	68.84	19.55	80.45	48.25	28.01	2.01	15.73
G259	ETSL101312	9.09	68.66	19.66	80.34	54.08	37.5	2.38	15.29
G260	ETSL101313	9.71	68.78	19.52	80.48	61.4	30.7	2.03	15.26

G261	ETSL101316	8.81	69.25	19.71	80.3	45.2	26.23	1.76	15.48
G262	ETSL101321	10.21	68.42	19.56	80.44	26.68	25.12	1.72	15.41
G263	ETSL101323	9.9	68.68	19.71	80.29	30.21	24.08	1.36	15.41
G264	ETSL101325	9.72	67.84	20.66	79.34	40.01	58.04	2.98	15.38
G265	ETSL101335	8.02	69.26	19.56	80.44	36.3	26.31	1.78	15.63
G266	ETSL101337	9.33	68.46	19.62	80.38	35.09	27.1	1.58	15.71
G267	ETSL101346	9.43	68.84	19.5	80.5	25.97	25.4	1.73	15.58
G268	ETSL101347	9.95	68.66	19.56	80.45	20.47	25.36	1.61	15.34
G269	ETSL101353	9.71	68.65	19.58	80.42	30.5	29.56	1.96	15.6
G270	ETSL101358	9.81	68.45	19.6	80.4	29.14	27.93	1.64	15.53
G271	ETSL101362	8.43	68.6	19.3	80.7	40.14	33.78	2.13	15.53
G272	ETSL101369	9.31	68.86	19.68	80.32	12.96	25.16	1.72	15.33
G273	ETSL101388	10.14	68.13	19.51	80.49	21.64	29.43	1.99	15.74
G274	ETSL101389	10.28	68.64	19.63	80.37	21.03	26.71	1.74	15.7
G275	ETSL101396	9.87	68.67	19.64	80.36	16.32	24.27	1.51	15.51
G276	ETSL101404	8.92	69.04	19.65	80.35	46.16	30.81	2	15.21
G277	ETSL101408	8.13	68.71	19.37	80.64	33.11	30.04	2	15.62
G278	ETSL101410	9.72	68.56	20.49	79.51	46.44	48.24	2.68	14.38
G279	ETSL101416	9.59	68.95	19.66	80.34	59.47	29.71	2.06	15.04
G280	ETSL101418	9.68	69.04	19.59	80.41	22.21	25.52	1.88	15.48
G281	ETSL101431	9.61	68.81	19.57	80.43	22.04	26.61	1.74	15.65
G282	ETSL101436	10.34	68.17	19.59	80.41	23.97	26.13	1.86	15.65
G283	ETSL101443	10.37	68.45	19.51	80.49	26.74	25.71	1.79	15.59
G284	ETSL101444	10.01	68.48	19.48	80.52	25.99	24.97	1.71	15.81
G285	ETSL101445	9.3	68.77	19.56	80.44	50.33	28.26	1.93	15.63
G286	ETSL101446	8.98	68.98	19.82	80.18	25.92	28.29	1.96	15.66
G287	ETSL101451	10.18	68.58	19.57	80.43	33.55	26.19	1.69	15.45
G288	ETSL101453	10.58	68.44	19.54	80.46	34.94	28.16	1.63	15.43
G289	ETSL101459	10	68.87	19.67	80.33	32.29	22.92	1.5	15.52
G290	ETSL101461	9.93	68.21	19.42	80.58	21.72	28.28	1.81	16.05
G291	ETSL101464	9.23	68.86	19.61	80.39	17.39	26.5	1.87	15.52
G292	ETSL101470	10.1	68.41	19.63	80.37	17.84	24.62	1.89	15.62
G293	ETSL101477	10.78	68.29	19.48	80.52	23.99	26.81	1.87	15.48
G294	ETSL101486	10.45	68.14	19.13	80.87	24.27	29.72	1.88	15.64
G295	ETSL101491	10.28	68.69	19.59	80.41	22.22	25.94	1.67	15.37
G296	ETSL101492	9.94	68.97	19.69	80.31	34.04	21.92	1.73	15.44
G297	ETSL101506	10.31	68.6	19.73	80.27	29.7	26.65	1.55	15.51
G298	ETSL101508	9.68	68.9	19.69	80.31	43.14	28.03	1.87	15.34
G299	ETSL101512	10.07	69.09	19.66	80.34	19.22	26.81	1.79	15.82
G300	ETSL101515	9.63	68.65	19.58	80.42	27.35	26.6	1.49	15.63
G301	ETSL101523	9.59	68.78	19.83	80.17	28.56	23.61	1.28	15.53
G302	ETSL101524	9.22	68.89	19.74	80.26	37.95	29.52	1.81	14.98
G303	ETSL101527	9.44	68.98	19.73	80.27	40.21	27.04	1.69	15.31
G304	ETSL101535	9.26	68.44	19.51	80.5	17.68	25.82	1.6	15.24

G305	ETSL101536	10.32	68.67	19.53	80.47	20.96	26.5	1.92	15.55
G306	ETSL101540	9.42	68.95	19.71	80.29	33.22	25.07	1.5	15.36
G307	ETSL101550	9.28	68.85	19.68	80.32	39.39	27.24	1.6	15.49
G308	ETSL101555	10.22	68.67	19.87	80.14	27.36	25.41	1.61	15.33
G309	ETSL101556	9.68	68.87	19.7	80.3	45.83	27.31	1.81	15.48
G310	ETSL101559	10.37	68.4	19.69	80.31	39.22	27.72	1.6	15.7
G311	ETSL101561	9.85	68.94	19.83	80.17	31.75	24.97	1.69	15.25
G312	ETSL101563	10.39	68.02	19.81	80.19	37.14	28.62	1.62	15
G313	ETSL101564	9.69	68.93	19.66	80.34	23.3	24.88	1.29	15.62
G314	ETSL101568	10.22	68.85	19.5	80.5	19.05	25.43	1.7	15.97
G315	ETSL101571	10.2	69.04	19.62	80.38	23.99	25.63	1.78	15.6
G316	ETSL101578	10.09	68.6	19.54	80.46	31.76	26.09	1.67	15.6
G317	ETSL101586	9.29	68.51	19.41	80.59	33.36	27.43	1.77	15.51
G318	ETSL101629	10.29	68.42	19.41	80.59	18.25	26.59	1.93	15.98
G319	ETSL101640	9.16	69.15	19.99	80.01	18.15	25.54	1.52	15.29
G320	ETSL101643	9.45	68.88	19.72	80.28	50.6	27.23	1.96	15.32
G321	ETSL101654	9.49	68.72	19.43	80.57	37.34	30.55	1.91	15.53
G322	ETSL101671	9.16	68.71	19.47	80.53	49.7	33.19	2.23	15.5
G323	ETSL101673	9.42	69	19.66	80.35	19.72	26.17	2.03	15.44
G324	ETSL101676	10.13	68.55	19.52	80.48	20.69	26.7	1.89	15.47
G325	ETSL101681	9.04	69.21	19.87	80.13	22.13	26.01	1.49	15.22
G326	ETSL101686	9.39	69.01	19.67	80.33	24.67	25.51	1.8	15.55
G327	ETSL101687	8.85	69.18	19.62	80.38	34.4	27.14	1.71	15.62
G328	ETSL101698	10.03	68.9	19.69	80.31	32.38	25.58	1.7	15.4
G329	ETSL101704	8.97	69.03	19.56	80.44	43.11	29.86	1.93	15.51
G330	ETSL101712	8.71	69.01	19.74	80.26	44.5	33.06	1.83	15.11
G331	ETSL101715	9.84	68.71	19.75	80.25	20.63	26.44	1.5	15.23
G332	ETSL101719	9.47	68.98	19.6	80.4	23.65	26.54	1.79	15.65
G333	ETSL101722	9.94	68.47	19.59	80.41	21.15	27.98	2.09	15.44
G334	ETSL101723	9.39	68.62	19.42	80.58	24.67	27.96	1.54	15.78
G335	ETSL101725	10.07	68.73	19.69	80.31	36.38	24.28	1.59	15.52
G336	ETSL101730	9.34	68.73	19.63	80.37	41.76	26.59	1.65	15.49
G337	ETSL101732	8.5	68.96	19.72	80.28	53.29	34.99	1.91	15.52
G338	ETSL101749	10.05	68.6	19.64	80.36	44.79	26.73	1.76	15.52
G339	ETSL101760	9.12	68.82	19.78	80.22	19.2	24.1	1.59	15.48
G340	ETSL101766	10.62	68.82	19.55	80.45	41.88	25.66	1.92	15.54
G341	ETSL101848	9.8	68.14	19.07	80.93	31.66	31.78	2.36	15.52
G342	ETSL101851	10.32	68.77	19.49	80.51	19.78	27.88	1.81	15.89
G343	ETSL101856	9.82	68.98	19.57	80.43	22.19	27.29	1.92	15.27
G344	ETSL101858	9.73	68.87	19.51	80.49	29.08	25.89	1.69	15.45
G345	ETSL101868	9.77	68.41	19.5	80.5	24.08	26.78	1.73	15.42
G346	IS25531	8.28	68.58	19.82	80.19	49.72	31.04	1.92	15.83
G347	IS38254	9.65	68.91	19.45	80.55	27.16	28.99	2.07	15.71
G348	IS38255	10.45	68.54	19.41	80.59	23.86	24.97	1.61	15.87

G349	IS38259	8.95	69.21	19.34	80.66	25.12	27.5	1.92	15.32
G350	IS38275	9.56	68.74	19.6	80.4	32.93	28.06	1.81	15.5
G351	IS38279	9.1	69.03	19.61	80.39	41.59	26.94	1.89	15.56
G352	IS38290	9.56	68.74	19.6	80.4	32.93	28.06	1.81	15.5
G353	IS38303	8.61	68.81	19.45	80.55	38.68	33.9	2.12	15.56
G354	IS38306	9.31	68.75	19.71	80.29	33.83	27.32	1.65	15.36
G355	IS38319	9.15	67.44	19.33	80.67	44.46	37.53	2.3	16.24
G356	IS38326	10.09	68.22	19.23	80.77	29.38	27.02	1.77	15.57
G357	IS38331	8.63	68.84	19.37	80.63	21.87	28.01	2.1	15.88
G358	IS38353	11.11	67.99	19.47	80.53	24.17	28.02	1.96	15.97
G359	IS38412	9.12	68.74	19.45	80.55	34.14	28.07	1.68	15.57
G360	P9830	10.51	68.49	19.52	80.48	18.46	27.09	1.55	15.51
G361	PML981446	9.43	69.05	19.86	80.14	54.57	29.21	2.04	15.48
G362	PML981474	10.52	67.9	19.36	80.64	34.3	31.98	2.07	15.53
G363	PML981476	9.07	69.3	19.7	80.29	34.74	24.12	1.66	15.46
G364	PML981488	9.33	68.83	19.79	80.21	29.79	25.67	1.47	15.49
G365	TAM428	9.98	68.62	19.66	80.34	19.26	25.56	1.58	15.5
	Mean	**9.56**	**68.74**	**19.6**	**80.4**	**32.93**	**28.05**	**1.81**	**15.5**

AM=Amylose, AMP=Amylopectin, MST=Moisture

Appendix 4. The significant SNPs associated with starch, amylose and protein in sorghum landrace collections across locations and years.

Trait	Year	Loc	Model	SNP	Chr	Pos	P.value	MAF	nobs	HB *p*.value ≤0.05	Effect
Amylose	2020	Miesso	FarmCPU	S01_26782338	1	26782338	3.02E-08	0.23	341	0	0.05
Amylose	2020	Miesso	BLINK	S01_56418710	1	56418710	5.39E-08	0.23	341	0	-0.06
Amylose	2020	Melkassa	FarmCPU	S01_56418710	1	56418710	2.35E-09	0.23	341	0	-0.05
Amylose	2021	Melkassa	FarmCPU	S01_56418710	1	56418710	2.27E-09	0.23	341	0	-0.06
Amylose	2020	Miesso	FarmCPU	S01_56418710	1	56418710	2.64E-10	0.23	341	0	-0.05
Amylose	2021	Miesso	FarmCPU	S01_56418710	1	56418710	2.95E-08	0.23	341	0	-0.05
Amylose	2021	Miesso	BLINK	S01_60803184	1	60803184	3.88E-07	0.44	341	0.01	-0.04
Amylose	2021	Miesso	FarmCPU	S01_60803184	1	60803184	6.06E-09	0.44	341	0	-0.04
Amylose	2021	Jimma	BLINK	S01_66821782	1	66821782	6.47E-13	0.09	341	0	-0.13
Amylose	2020	Melkassa	BLINK	S01_66821782	1	66821782	4.04E-07	0.09	341	0.02	-0.08
Amylose	2021	Melkassa	BLINK	S01_66821782	1	66821782	3.05E-07	0.09	341	0.02	-0.09
Amylose	2020	Miesso	BLINK	S01_66821782	1	66821782	9.19E-09	0.09	341	0	-0.09
Amylose	2021	Miesso	BLINK	S01_66821782	1	66821782	1.18E-11	0.09	341	0	-0.1
Amylose	2020	Melkassa	FarmCPU	S01_66821782	1	66821782	1.45E-06	0.09	341	0.03	-0.07
Amylose	2021	Melkassa	FarmCPU	S01_66821782	1	66821782	1.01E-06	0.09	341	0.02	-0.08
Amylose	2020	Miesso	FarmCPU	S01_66821782	1	66821782	1.07E-07	0.09	341	0	-0.07
Amylose	2021	Miesso	FarmCPU	S01_66821782	1	66821782	4.12E-11	0.09	341	0	-0.11
Amylose	2021	Jimma	BLINK	S01_67415726	1	67415726	2.88E-11	0.26	341	0	0.12
Amylose	2020	Melkassa	BLINK	S01_67415726	1	67415726	2.38E-09	0.26	341	0	0.09
Amylose	2021	Melkassa	BLINK	S01_67415726	1	67415726	3.07E-09	0.26	341	0	0.1
Amylose	2020	Miesso	BLINK	S01_67415726	1	67415726	1.22E-09	0.26	341	0	0.09
Amylose	2021	Miesso	BLINK	S01_67415726	1	67415726	9.51E-11	0.26	341	0	0.09
Amylose	2021	Jimma	FarmCPU	S01_67415726	1	67415726	4.89E-08	0.26	341	0	0.08
Amylose	2020	Melkassa	FarmCPU	S01_67415726	1	67415726	3.44E-08	0.26	341	0	0.06
Amylose	2021	Melkassa	FarmCPU	S01_67415726	1	67415726	4.24E-08	0.26	341	0	0.06
Amylose	2020	Miesso	FarmCPU	S01_67415726	1	67415726	3.89E-10	0.26	341	0	0.07
Amylose	2021	Miesso	FarmCPU	S01_67415726	1	67415726	2.03E-08	0.26	341	0	0.07
Amylose	2020	Jimma	FarmCPU	S02_11505508	2	11505508	8.71E-07	0.11	341	0.03	-0.05
Amylose	2020	Melkassa	BLINK	S02_12372094	2	12372094	8.82E-07	0.26	341	0.04	0.06
Amylose	2021	Melkassa	BLINK	S02_12372094	2	12372094	8.16E-07	0.26	341	0.03	0.06
Amylose	2020	Jimma	FarmCPU	S02_63195523	2	63195523	3.07E-09	0.06	341	0	-0.08
Amylose	2020	Melkassa	FarmCPU	S03_54181079	3	54181079	5.56E-07	0.43	341	0.01	-0.03
Amylose	2021	Melkassa	FarmCPU	S03_54181079	3	54181079	4.81E-07	0.43	341	0.01	-0.03
Amylose	2020	Miesso	FarmCPU	S03_54344093	3	54344093	2.82E-08	0.34	341	0	-0.04
Amylose	2020	Jimma	FarmCPU	S03_58400495	3	58400495	8.91E-07	0.37	341	0.03	0.04
Amylose	2020	Jimma	FarmCPU	S03_61856025	3	61856025	3.71E-07	0.23	341	0.02	0.04
Amylose	2020	Jimma	BLINK	S03_63580967	3	63580967	3.90E-08	0.21	341	0	-0.08
Amylose	2020	Jimma	FarmCPU	S03_63580967	3	63580967	2.42E-08	0.21	341	0	-0.07
Amylose	2020	Jimma	BLINK	S03_65414137	3	65414137	2.33E-07	0.06	341	0.02	-0.11

Amylose	2020	Miesso	FarmCPU	S03_8276506	3	8276506	2.10E-06	0.21	341	0.04	0.03
Amylose	2020	Melkassa	FarmCPU	S04_26104010	4	26104010	2.57E-08	0.43	341	0	-0.04
Amylose	2021	Melkassa	FarmCPU	S04_26104010	4	26104010	2.08E-08	0.43	341	0	-0.04
Amylose	2021	Miesso	FarmCPU	S04_39731078	4	39731078	6.84E-07	0.23	341	0.01	-0.05
Amylose	2020	Miesso	FarmCPU	S05_51045517	5	51045517	7.98E-08	0.27	341	0	-0.05
Amylose	2020	Miesso	FarmCPU	S05_6176773	5	6176773	6.11E-07	0.24	341	0.01	-0.04
Amylose	2021	Miesso	BLINK	S06_39126368	6	39126368	4.36E-10	0.13	341	0	0.09
Amylose	2020	Melkassa	FarmCPU	S06_39126368	6	39126368	2.51E-10	0.13	341	0	0.07
Amylose	2021	Melkassa	FarmCPU	S06_39126368	6	39126368	2.26E-10	0.13	341	0	0.08
Amylose	2020	Miesso	FarmCPU	S06_39126368	6	39126368	1.31E-12	0.13	341	0	0.08
Amylose	2021	Miesso	FarmCPU	S06_39126368	6	39126368	6.73E-11	0.13	341	0	0.08
Amylose	2020	Miesso	BLINK	S06_43654354	6	43654354	2.16E-07	0.2	341	0.01	-0.06
Amylose	2021	Jimma	FarmCPU	S06_52182739	6	52182739	1.03E-07	0.36	341	0	-0.04
Amylose	2020	Melkassa	FarmCPU	S06_55741394	6	55741394	4.73E-07	0.37	341	0.01	-0.04
Amylose	2021	Melkassa	FarmCPU	S06_55741394	6	55741394	3.34E-07	0.37	341	0.01	-0.05
Amylose	2021	Jimma	FarmCPU	S06_57903103	6	57903103	1.56E-07	0.07	341	0	-0.09
Amylose	2021	Miesso	BLINK	S07_5366301	7	5366301	7.11E-08	0.19	341	0	0.06
Amylose	2021	Miesso	FarmCPU	S07_5366301	7	5366301	2.66E-07	0.19	341	0.01	0.05
Amylose	2020	Jimma	BLINK	S07_62671638	7	62671638	8.74E-09	0.09	341	0	0.06
Amylose	2021	Jimma	BLINK	S07_62671638	7	62671638	6.43E-17	0.09	341	0	0.13
Amylose	2020	Melkassa	BLINK	S07_62671638	7	62671638	9.68E-15	0.09	341	0	0.11
Amylose	2021	Melkassa	BLINK	S07_62671638	7	62671638	1.42E-14	0.09	341	0	0.11
Amylose	2020	Miesso	BLINK	S07_62671638	7	62671638	1.09E-20	0.09	341	0	0.14
Amylose	2021	Miesso	BLINK	S07_62671638	7	62671638	1.50E-16	0.09	341	0	0.11
Amylose	2020	Jimma	FarmCPU	S07_62671638	7	62671638	5.70E-14	0.09	341	0	0.09
Amylose	2021	Jimma	FarmCPU	S07_62671638	7	62671638	6.13E-18	0.09	341	0	0.14
Amylose	2020	Melkassa	FarmCPU	S07_62671638	7	62671638	1.69E-28	0.09	341	0	0.15
Amylose	2021	Melkassa	FarmCPU	S07_62671638	7	62671638	1.96E-28	0.09	341	0	0.16
Amylose	2020	Miesso	FarmCPU	S07_62671638	7	62671638	7.47E-27	0.09	341	0	0.14
Amylose	2021	Miesso	FarmCPU	S07_62671638	7	62671638	4.89E-19	0.09	341	0	0.12
Amylose	2021	Miesso	BLINK	S08_456533	8	456533	1.58E-06	0.31	341	0.03	-0.06
Amylose	2021	Jimma	FarmCPU	S08_56538511	8	56538511	2.15E-07	0.32	341	0.01	0.05
Amylose	2020	Miesso	BLINK	S08_58248093	8	58248093	9.01E-08	0.15	341	0	0.07
Amylose	2021	Miesso	BLINK	S08_58248093	8	58248093	1.18E-08	0.15	341	0	0.07
Amylose	2020	Melkassa	FarmCPU	S08_58248093	8	58248093	5.20E-08	0.15	341	0	0.05
Amylose	2021	Melkassa	FarmCPU	S08_58248093	8	58248093	4.58E-08	0.15	341	0	0.06
Amylose	2020	Miesso	FarmCPU	S08_58248093	8	58248093	4.99E-10	0.15	341	0	0.06
Amylose	2021	Miesso	FarmCPU	S08_58248093	8	58248093	1.25E-08	0.15	341	0	0.06
Amylose	2020	Miesso	FarmCPU	S09_35032274	9	35032274	2.52E-06	0.09	341	0.04	-0.07
Amylose	2020	Jimma	BLINK	S09_5697851	9	5697851	7.65E-07	0.09	341	0.04	0.09
Amylose	2021	Jimma	BLINK	S10_21605004	10	21605004	2.32E-09	0.27	341	0	0.09
Amylose	2020	Melkassa	BLINK	S10_21605004	10	21605004	3.83E-10	0.27	341	0	0.08
Amylose	2021	Melkassa	BLINK	S10_21605004	10	21605004	4.77E-10	0.27	341	0	0.09
Amylose	2020	Miesso	BLINK	S10_21605004	10	21605004	1.03E-10	0.27	341	0	0.08

Amylose	2021	Miesso	BLINK	S10_21605004	10	21605004	4.54E-11	0.27	341	0	0.08
Amylose	2021	Jimma	FarmCPU	S10_21605004	10	21605004	1.00E-10	0.27	341	0	0.08
Amylose	2020	Melkassa	FarmCPU	S10_21605004	10	21605004	2.73E-13	0.27	341	0	0.07
Amylose	2021	Melkassa	FarmCPU	S10_21605004	10	21605004	3.82E-13	0.27	341	0	0.08
Amylose	2020	Miesso	FarmCPU	S10_21605004	10	21605004	3.48E-14	0.27	341	0	0.07
Amylose	2021	Miesso	FarmCPU	S10_21605004	10	21605004	1.30E-07	0.27	341	0	0.06
Amylose	2021	Miesso	BLINK	S10_2513319	10	2513319	8.53E-08	0.27	341	0	0.05
Amylose	2021	Jimma	BLINK	S10_27955546	10	27955546	1.25E-07	0.2	341	0	-0.09
Amylose	2021	Jimma	FarmCPU	S10_27955546	10	27955546	1.36E-06	0.2	341	0.03	-0.06
Amylose	2021	Jimma	BLINK	S10_37345784	10	37345784	1.12E-07	0.34	341	0	0.07
Amylose	2021	Jimma	FarmCPU	S10_37345784	10	37345784	1.04E-07	0.34	341	0	0.05
Amylose	2021	Jimma	FarmCPU	S10_49076966	10	49076966	7.25E-07	0.18	341	0.02	0.06
Amylose	2021	Jimma	FarmCPU	S10_52972135	10	52972135	2.10E-08	0.26	341	0	-0.06
Amylose	2020	Melkassa	FarmCPU	S10_52972135	10	52972135	1.57E-08	0.26	341	0	-0.05
Amylose	2021	Melkassa	FarmCPU	S10_52972135	10	52972135	1.60E-08	0.26	341	0	-0.06
Protein	2020	Melkassa	BLINK	S01_1332500	1	1332500	2.82E-06	0.48	341	0.05	0.11
Protein	2021	Jimma	BLINK	S01_1331541	1	1331541	3.56E-08	0.35	341	0.01	-0.2
Protein	2021	Melkassa	BLINK	S01_67661545	1	67661545	6.33E-08	0.19	341	0.01	0.34
Protein	2020	Melkassa	BLINK	S02_3743197	2	3743197	2.98E-08	0.19	341	0	0.19
Protein	2021	Melkassa	BLINK	S02_3743197	2	3743197	6.48E-07	0.19	341	0.05	0.3
Protein	2020	Miesso	BLINK	S02_3743197	2	3743197	3.54E-09	0.19	341	0	0.3
Protein	2020	Melkassa	BLINK	S02_44842565	2	44842565	1.09E-07	0.34	341	0.01	0.2
Protein	2021	Melkassa	BLINK	S02_65329765	2	65329765	8.67E-08	0.08	341	0.01	-0.38
Protein	2020	Miesso	BLINK	S04_67458597	4	67458597	3.31E-08	0.27	341	0	-0.32
Protein	2020	Jimma	BLINK	S07_7936839	7	7936839	9.20E-08	0.24	341	0.02	-0.19
Protein	2020	Melkassa	BLINK	S09_3030812	9	3030812	5.49E-12	0.1	341	0	0.32
Starch	2020	Jimma	FarmCPU	S01_31764844	1	31764844	4.78E-09	0.18	341	0	-0.08
Starch	2021	Melkassa	FarmCPU	S01_31764844	1	31764844	1.32E-07	0.18	341	0.01	-0.16
Starch	2021	Jimma	FarmCPU	S01_57119003	1	57119003	1.29E-10	0.15	341	0	0.16
Starch	2020	Jimma	FarmCPU	S01_73075918	1	73075918	4.07E-07	0.15	341	0.02	0.07
Starch	2020	Jimma	FarmCPU	S02_33775277	2	33775277	3.39E-08	0.27	341	0	0.06
Starch	2021	Melkassa	FarmCPU	S02_33775277	2	33775277	5.67E-08	0.27	341	0	0.13
Starch	2021	Jimma	FarmCPU	S02_69664050	2	69664050	9.55E-07	0.24	341	0.02	0.11
Starch	2021	Melkassa	FarmCPU	S02_71486000	2	71486000	6.18E-07	0.23	341	0.02	0.12
Starch	2021	Jimma	FarmCPU	S02_73817073	2	73817073	1.47E-08	0.38	341	0	-0.1
Starch	2020	Jimma	FarmCPU	S03_32214612	3	32214612	6.85E-07	0.31	341	0.03	0.05
Starch	2021	Melkassa	FarmCPU	S03_32214612	3	32214612	4.74E-07	0.31	341	0.02	0.12
Starch	2021	Melkassa	FarmCPU	S03_54545596	3	54545596	7.85E-08	0.36	341	0	-0.14
Starch	2021	Jimma	FarmCPU	S03_69977324	3	69977324	3.88E-09	0.49	341	0	0.12
Starch	2021	Jimma	FarmCPU	S04_13036200	4	13036200	8.74E-09	0.36	341	0	-0.12
Starch	2021	Jimma	FarmCPU	S04_43712537	4	43712537	8.44E-09	0.31	341	0	-0.15
Starch	2021	Jimma	BLINK	S05_4529528	5	4529528	1.92E-09	0.39	341	0	-0.2
Starch	2021	Jimma	FarmCPU	S05_63315120	5	63315120	3.18E-11	0.23	341	0	-0.13
Starch	2021	Jimma	BLINK	S06_48180471	6	48180471	2.38E-09	0.31	341	0	-0.19

Starch	2020	Jimma	FarmCPU	S06_48180471	6	48180471	3.96E-07	0.31	341	0.02	-0.07
Starch	2021	Jimma	FarmCPU	S06_48180471	6	48180471	6.14E-10	0.31	341	0	-0.15
Starch	2020	Jimma	BLINK	S06_59687350	6	59687350	7.24E-10	0.37	341	0	-0.09
Starch	2021	Melkassa	BLINK	S06_59687350	6	59687350	6.87E-10	0.37	341	0	-0.21
Starch	2021	Miesso	BLINK	S06_59687350	6	59687350	1.15E-09	0.37	341	0	-0.15
Starch	2021	Melkassa	FarmCPU	S06_59687350	6	59687350	5.73E-07	0.37	341	0.02	-0.13
Starch	2021	Jimma	BLINK	S07_60921702	7	60921702	2.40E-07	0.15	341	0.02	-0.17
Starch	2021	Melkassa	FarmCPU	S08_19514794	8	19514794	2.74E-08	0.35	341	0	-0.17
Starch	2020	Jimma	FarmCPU	S08_46319474	8	46319474	1.26E-06	0.25	341	0.04	0.05
Starch	2021	Melkassa	FarmCPU	S08_46319474	8	46319474	7.27E-10	0.25	341	0	0.16
Starch	2021	Jimma	FarmCPU	S09_37127026	9	37127026	1.51E-06	0.1	341	0.03	-0.14
Starch	2021	Jimma	FarmCPU	S10_41827671	10	41827671	8.41E-07	0.15	341	0.02	-0.14

Appendix 5. The significant SNPs associated with ash and mineral compositions in sorghum landrace collections across locations and years.

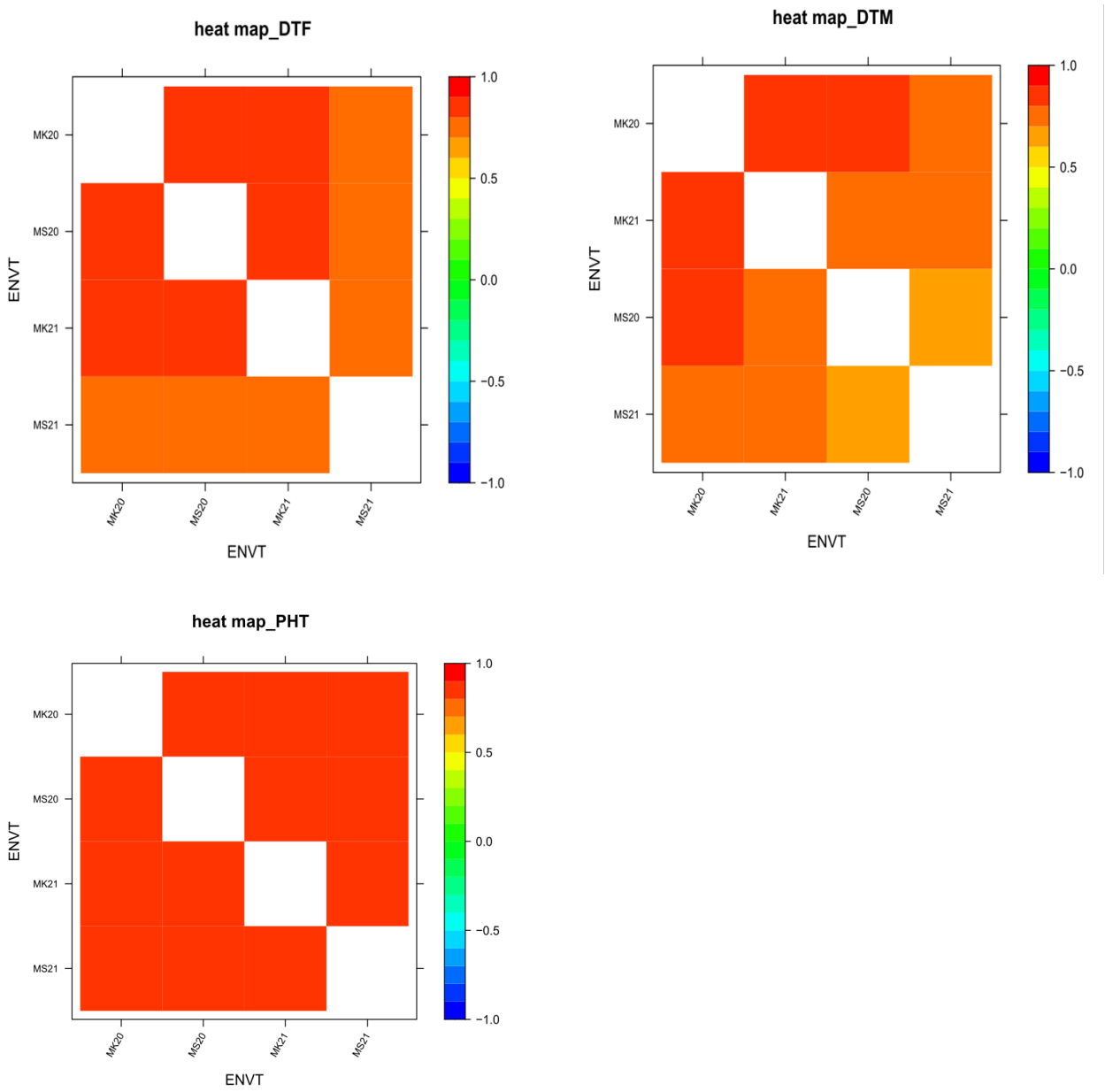
Trait	Year	Loc	Model	SNP	Chr	Pos	P.value	MAF	nobs	HB *p*.value ≤0.05	Effect
Ash	2021	Miesso	FarmCPU	S01_11836109	1	11836109	5.57E-07	0.17	341	0.01	0.10
Ash	2021	Jimma	FarmCPU	S01_19153514	1	19153514	1.45E-07	0.30	341	0.01	-0.10
Ash	2020	Jimma	FarmCPU	S01_27831887	1	27831887	4.51E-07	0.28	341	0.01	0.12
Ash	2020	Jimma	BLINK	S01_2883853	1	2883853	2.44E-09	0.23	341	0.00	-0.16
Ash	2020	Jimma	FarmCPU	S01_3668255	1	3668255	6.79E-09	0.09	341	0.00	-0.15
Ash	2021	Jimma	FarmCPU	S01_67757574	1	67757574	9.23E-07	0.33	341	0.03	-0.10
Ash	2020	Jimma	FarmCPU	S02_26996992	2	26996992	1.95E-07	0.40	341	0.01	-0.09
Ash	2021	Jimma	BLINK	S02_65113562	2	65113562	1.42E-07	0.26	341	0.01	0.09
Ash	2020	Melkassa	BLINK	S02_65113562	2	65113562	3.08E-09	0.26	341	0.00	0.10
Ash	2021	Melkassa	BLINK	S02_65113562	2	65113562	5.23E-09	0.26	341	0.00	0.10
Ash	2021	Melkassa	FarmCPU	S02_65113562	2	65113562	1.26E-06	0.26	341	0.02	0.06
Ash	2020	Jimma	BLINK	S02_65235509	2	65235509	1.14E-08	0.07	341	0.00	-0.22
Ash	2021	Jimma	FarmCPU	S03_1918680	3	1918680	1.59E-06	0.11	341	0.04	0.09
Ash	2021	Melkassa	BLINK	S03_2052347	3	2052347	2.38E-07	0.15	341	0.01	-0.11
Ash	2021	Melkassa	FarmCPU	S03_2052347	3	2052347	2.53E-06	0.15	341	0.04	-0.07
Ash	2021	Jimma	BLINK	S03_2053976	3	2053976	3.61E-08	0.15	341	0.00	0.13
Ash	2021	Miesso	FarmCPU	S03_62488498	3	62488498	1.06E-08	0.23	341	0.00	-0.11
Ash	2020	Jimma	FarmCPU	S03_68903619	3	68903619	1.60E-07	0.49	341	0.01	-0.10
Ash	2021	Melkassa	FarmCPU	S03_72123500	3	72123500	1.53E-07	0.25	341	0.00	-0.06
Ash	2021	Melkassa	FarmCPU	S03_9065334	3	9065334	2.15E-08	0.11	341	0.00	0.13
Ash	2021	Melkassa	FarmCPU	S04_34744457	4	34744457	1.74E-08	0.37	341	0.00	0.11
Ash	2020	Jimma	FarmCPU	S04_66154866	4	66154866	2.11E-06	0.41	341	0.05	0.10
Ash	2021	Melkassa	FarmCPU	S04_7901389	4	7901389	4.22E-07	0.28	341	0.01	-0.07
Ash	2021	Jimma	BLINK	S04_9072570	4	9072570	1.06E-07	0.35	341	0.01	0.13
Ash	2020	Melkassa	BLINK	S05_2789336	5	2789336	1.83E-09	0.10	341	0.00	-0.15
Ash	2021	Jimma	FarmCPU	S05_35875751	5	35875751	6.47E-08	0.39	341	0.00	-0.09
Ash	2020	Jimma	FarmCPU	S05_65910308	5	65910308	7.86E-10	0.18	341	0.00	0.12
Ash	2021	Melkassa	FarmCPU	S05_66877074	5	66877074	1.42E-07	0.13	341	0.00	0.09
Ash	2021	Miesso	FarmCPU	S05_69163685	5	69163685	4.40E-07	0.10	341	0.01	-0.13
Ash	2021	Melkassa	FarmCPU	S05_989345	5	989345	2.64E-15	0.13	341	0.00	-0.14
Ash	2021	Miesso	FarmCPU	S06_1681630	6	1681630	2.42E-09	0.36	341	0.00	-0.13
Ash	2021	Jimma	FarmCPU	S06_1839072	6	1839072	2.21E-08	0.18	341	0.00	0.13
Ash	2021	Melkassa	FarmCPU	S06_2725114	6	2725114	1.04E-08	0.43	341	0.00	-0.09
Ash	2020	Jimma	BLINK	S06_55165743	6	55165743	2.00E-08	0.23	341	0.00	-0.15
Ash	2020	Jimma	FarmCPU	S06_55165746	6	55165746	5.95E-08	0.23	341	0.00	0.10
Ash	2020	Jimma	FarmCPU	S07_2020128	7	2020128	6.27E-09	0.20	341	0.00	0.11
Ash	2021	Miesso	FarmCPU	S07_53728407	7	53728407	1.15E-08	0.29	341	0.00	-0.11
Ash	2021	Melkassa	BLINK	S07_62611661	7	62611661	1.87E-05	0.33	341	0.36	-0.07
Ash	2020	Jimma	FarmCPU	S07_62611661	7	62611661	1.93E-10	0.33	341	0.00	-0.10

Ash	2021	Jimma	FarmCPU	S07_62671867	7	62671867	1.10E-06	0.06	341	0.03	0.12
Ash	2021	Miesso	FarmCPU	S07_6584885	7	6584885	2.32E-06	0.14	341	0.04	-0.10
Ash	2021	Jimma	FarmCPU	S08_11020973	8	11020973	1.15E-06	0.43	341	0.03	-0.06
Ash	2021	Jimma	FarmCPU	S08_61262673	8	61262673	3.82E-08	0.15	341	0.00	0.11
Ash	2021	Jimma	BLINK	S08_7974942	8	7974942	6.16E-09	0.41	341	0.00	-0.13
Ash	2021	Melkassa	BLINK	S08_7974942	8	7974942	2.45E-07	0.41	341	0.01	-0.11
Ash	2021	Jimma	FarmCPU	S08_7974942	8	7974942	5.65E-07	0.41	341	0.02	-0.08
Ash	2021	Melkassa	FarmCPU	S08_7974942	8	7974942	1.52E-14	0.41	341	0.00	-0.11
Ash	2020	Jimma	BLINK	S08_9366659	8	9366659	1.25E-08	0.36	341	0.00	-0.15
Ash	2021	Melkassa	FarmCPU	S09_33577618	9	33577618	4.24E-09	0.20	341	0.00	0.09
Ash	2021	Miesso	FarmCPU	S09_43254494	9	43254494	7.33E-07	0.46	341	0.02	-0.09
Ash	2021	Miesso	BLINK	S09_49260419	9	49260419	4.15E-08	0.17	341	0.01	-0.15
Ash	2021	Miesso	FarmCPU	S09_49260419	9	49260419	2.84E-07	0.17	341	0.01	-0.11
Ash	2021	Melkassa	BLINK	S09_50320755	9	50320755	2.57E-07	0.34	341	0.01	-0.10
Ash	2021	Miesso	FarmCPU	S09_52930421	9	52930421	6.52E-07	0.09	341	0.02	0.12
Ash	2021	Melkassa	FarmCPU	S09_55539786	9	55539786	8.79E-07	0.46	341	0.02	0.06
Ash	2021	Melkassa	FarmCPU	S10_2019986	10	2019986	8.12E-07	0.23	341	0.02	0.08
Ash	2021	Miesso	FarmCPU	S10_42931123	10	42931123	1.66E-08	0.35	341	0.00	-0.12
Ash	2021	Miesso	FarmCPU	S10_59100730	10	59100730	1.18E-08	0.37	341	0.00	-0.10
Fe	2020	Melkassa	FarmCPU	S01_16805243	1	16805243	2.89E-07	0.45	341	0.01	-2.65
Fe	2021	Melkassa	FarmCPU	S01_16805243	1	16805243	1.46E-06	0.45	341	0.04	-2.35
Fe	2020	Melkassa	FarmCPU	S01_330396	1	330396	1.35E-07	0.06	341	0.00	4.62
Fe	2021	Melkassa	FarmCPU	S01_330396	1	330396	1.71E-06	0.06	341	0.04	3.96
Fe	2020	Jimma	FarmCPU	S01_79466678	1	79466678	4.88E-08	0.39	341	0.00	2.10
Fe	2021	Melkassa	BLINK	S02_26964880	2	26964880	6.28E-07	0.27	341	0.03	4.51
Fe	2020	Melkassa	FarmCPU	S02_26964880	2	26964880	1.82E-06	0.27	341	0.03	3.40
Fe	2021	Melkassa	FarmCPU	S02_26964880	2	26964880	1.75E-07	0.27	341	0.01	3.57
Fe	2021	Miesso	BLINK	S02_56963182	2	56963182	1.60E-07	0.42	341	0.01	-4.19
Fe	2021	Jimma	BLINK	S02_62701316	2	62701316	4.20E-07	0.34	341	0.04	3.35
Fe	2021	Miesso	FarmCPU	S03_48548871	3	48548871	3.75E-07	0.14	341	0.02	4.06
Fe	2020	Jimma	FarmCPU	S03_52557257	3	52557257	1.81E-07	0.15	341	0.01	-2.93
Fe	2020	Jimma	FarmCPU	S03_6990771	3	6990771	1.05E-06	0.23	341	0.02	-2.35
Fe	2021	Miesso	BLINK	S03_70557923	3	70557923	3.41E-11	0.19	341	0.00	6.20
Fe	2021	Miesso	FarmCPU	S03_70557923	3	70557923	1.39E-08	0.19	341	0.00	4.19
Fe	2020	Jimma	FarmCPU	S04_15880752	4	15880752	1.15E-06	0.34	341	0.02	2.61
Fe	2020	Melkassa	FarmCPU	S04_25561994	4	25561994	1.05E-07	0.20	341	0.00	4.28
Fe	2021	Melkassa	FarmCPU	S04_25561994	4	25561994	1.19E-06	0.20	341	0.04	3.59
Fe	2021	Miesso	FarmCPU	S04_25561994	4	25561994	8.28E-08	0.20	341	0.00	4.13
Fe	2021	Melkassa	BLINK	S04_27086291	4	27086291	1.13E-07	0.26	341	0.01	4.72
Fe	2021	Miesso	BLINK	S04_27086291	4	27086291	5.78E-08	0.26	341	0.00	5.12
Fe	2021	Miesso	FarmCPU	S04_27086314	4	27086314	9.49E-07	0.32	341	0.03	-3.18
Fe	2020	Jimma	FarmCPU	S04_41955563	4	41955563	5.08E-08	0.23	341	0.00	2.39
Fe	2020	Melkassa	FarmCPU	S04_56936857	4	56936857	7.05E-07	0.41	341	0.01	2.13
Fe	2020	Jimma	BLINK	S04_62316425	4	62316425	4.92E-12	0.42	341	0.00	-5.31

Fe	2020	Jimma	FarmCPU	S04_62316425	4	62316425	5.03E-17	0.42	341	0.00	-4.92
Fe	2021	Melkassa	FarmCPU	S04_64931579	4	64931579	4.40E-07	0.24	341	0.02	-3.33
Fe	2020	Jimma	FarmCPU	S04_8819156	4	8819156	7.61E-07	0.18	341	0.02	-2.67
Fe	2021	Miesso	FarmCPU	S04_913058	4	913058	6.37E-08	0.35	341	0.00	-3.38
Fe	2020	Melkassa	FarmCPU	S05_3978193	5	3978193	6.60E-08	0.45	341	0.00	3.35
Fe	2021	Jimma	BLINK	S05_65910190	5	65910190	9.40E-08	0.24	341	0.02	3.36
Fe	2021	Melkassa	BLINK	S05_65910190	5	65910190	9.33E-10	0.24	341	0.00	4.78
Fe	2020	Melkassa	BLINK	S05_65910190	5	65910190	7.34E-09	0.24	341	0.00	5.92
Fe	2020	Miesso	BLINK	S05_65910190	5	65910190	1.60E-08	0.24	341	0.00	4.90
Fe	2021	Miesso	BLINK	S05_65910190	5	65910190	5.10E-08	0.24	341	0.00	4.46
Fe	2020	Melkassa	FarmCPU	S05_65910190	5	65910190	1.57E-09	0.24	341	0.00	4.06
Fe	2021	Melkassa	FarmCPU	S05_65910190	5	65910190	1.86E-08	0.24	341	0.00	3.43
Fe	2020	Jimma	FarmCPU	S06_11021535	6	11021535	2.91E-07	0.29	341	0.01	2.97
Fe	2020	Melkassa	FarmCPU	S06_49039529	6	49039529	7.12E-09	0.32	341	0.00	-4.18
Fe	2021	Melkassa	FarmCPU	S06_57253548	6	57253548	1.39E-07	0.38	341	0.01	-3.48
Fe	2020	Jimma	BLINK	S07_2020128	7	2020128	1.66E-08	0.20	341	0.00	3.43
Fe	2020	Jimma	FarmCPU	S07_2020128	7	2020128	1.11E-08	0.20	341	0.00	2.59
Fe	2021	Melkassa	BLINK	S07_2029818	7	2029818	1.78E-07	0.24	341	0.01	3.15
Fe	2020	Melkassa	BLINK	S07_2029818	7	2029818	1.09E-07	0.24	341	0.01	4.23
Fe	2020	Melkassa	FarmCPU	S07_2029818	7	2029818	6.87E-12	0.24	341	0.00	3.89
Fe	2021	Melkassa	FarmCPU	S07_2029818	7	2029818	7.00E-10	0.24	341	0.00	3.12
Fe	2020	Melkassa	FarmCPU	S07_20896689	7	20896689	3.54E-07	0.48	341	0.01	3.46
Fe	2020	Melkassa	FarmCPU	S07_49748916	7	49748916	9.80E-09	0.41	341	0.00	-3.93
Fe	2020	Jimma	FarmCPU	S07_62725659	7	62725659	1.38E-06	0.07	341	0.02	-3.13
Fe	2020	Melkassa	FarmCPU	S08_15532863	8	15532863	3.01E-06	0.13	341	0.05	-3.83
Fe	2021	Miesso	FarmCPU	S08_62143948	8	62143948	1.56E-06	0.24	341	0.04	-3.11
Fe	2021	Melkassa	FarmCPU	S08_8691653	8	8691653	1.65E-06	0.18	341	0.04	-4.44
Fe	2021	Miesso	FarmCPU	S08_8691653	8	8691653	1.53E-06	0.18	341	0.04	-4.29
Fe	2020	Jimma	BLINK	S09_53178158	9	53178158	3.03E-07	0.35	341	0.02	3.93
Fe	2020	Jimma	FarmCPU	S09_53178158	9	53178158	5.25E-11	0.35	341	0.00	3.75
Fe	2020	Melkassa	FarmCPU	S10_54927527	10	54927527	5.44E-09	0.39	341	0.00	-4.03
Fe	2020	Jimma	BLINK	S10_56127541	10	56127541	3.11E-09	0.15	341	0.00	3.97
Fe	2021	Miesso	BLINK	S10_56127541	10	56127541	1.37E-08	0.15	341	0.00	4.55
Fe	2020	Jimma	FarmCPU	S10_56127541	10	56127541	1.47E-07	0.15	341	0.01	3.28
Fe	2021	Miesso	FarmCPU	S10_56127541	10	56127541	1.76E-12	0.15	341	0.00	4.81
Zn	2020	Jimma	FarmCPU	S01_13325810	1	13325810	1.24E-06	0.13	341	0.03	-3.02
Zn	2020	Melkassa	BLINK	S01_30708847	1	30708847	7.56E-07	0.25	341	0.02	-1.37
Zn	2021	Jimma	FarmCPU	S01_3699769	1	3699769	1.58E-07	0.13	341	0.00	-1.41
Zn	2020	Melkassa	FarmCPU	S01_48506669	1	48506669	8.59E-08	0.30	341	0.00	-1.05
Zn	2021	Melkassa	FarmCPU	S01_48506669	1	48506669	2.36E-09	0.30	341	0.00	-0.99
Zn	2021	Jimma	FarmCPU	S01_66528020	1	66528020	1.41E-08	0.44	341	0.00	-1.23
Zn	2020	Melkassa	FarmCPU	S01_66528020	1	66528020	2.08E-09	0.44	341	0.00	-1.27
Zn	2021	Melkassa	FarmCPU	S01_66528020	1	66528020	2.70E-10	0.44	341	0.00	-1.15
Zn	2020	Melkassa	FarmCPU	S01_77602799	1	77602799	2.31E-07	0.07	341	0.00	-2.00

Zn	2021	Melkassa	FarmCPU	S01_77602799	1	77602799	5.38E-07	0.07	341	0.01	-1.59
Zn	2021	Melkassa	BLINK	S02_18494291	2	18494291	9.07E-09	0.14	341	0.00	-1.29
Zn	2021	Jimma	FarmCPU	S02_18494291	2	18494291	1.77E-11	0.14	341	0.00	-2.02
Zn	2020	Melkassa	FarmCPU	S02_18494291	2	18494291	1.74E-16	0.14	341	0.00	-2.35
Zn	2021	Melkassa	FarmCPU	S02_18494291	2	18494291	1.81E-16	0.14	341	0.00	-1.92
Zn	2020	Melkassa	FarmCPU	S02_52641107	2	52641107	1.67E-08	0.09	341	0.00	-1.82
Zn	2021	Melkassa	FarmCPU	S02_52641107	2	52641107	1.63E-09	0.09	341	0.00	-1.67
Zn	2020	Jimma	BLINK	S02_63770377	2	63770377	1.17E-12	0.18	341	0.00	-5.75
Zn	2020	Melkassa	BLINK	S02_63770377	2	63770377	1.13E-12	0.18	341	0.00	-2.69
Zn	2021	Melkassa	BLINK	S02_63770377	2	63770377	7.41E-12	0.18	341	0.00	-2.13
Zn	2020	Jimma	FarmCPU	S02_63770377	2	63770377	5.83E-07	0.18	341	0.03	-3.47
Zn	2021	Jimma	FarmCPU	S02_63770377	2	63770377	1.92E-07	0.18	341	0.01	-1.76
Zn	2020	Melkassa	FarmCPU	S02_63770377	2	63770377	9.45E-09	0.18	341	0.00	-1.82
Zn	2021	Melkassa	FarmCPU	S02_63770377	2	63770377	2.69E-08	0.18	341	0.00	-1.42
Zn	2020	Melkassa	BLINK	S02_75070293	2	75070293	1.19E-06	0.14	341	0.03	-1.24
Zn	2021	Melkassa	BLINK	S02_75070293	2	75070293	1.31E-08	0.14	341	0.00	-1.32
Zn	2020	Melkassa	FarmCPU	S02_7736157	2	7736157	6.37E-08	0.22	341	0.00	-1.24
Zn	2021	Melkassa	FarmCPU	S02_7736173	2	7736173	7.27E-09	0.22	341	0.00	1.10
Zn	2020	Jimma	FarmCPU	S03_1918680	3	1918680	1.22E-06	0.11	341	0.03	2.34
Zn	2020	Melkassa	BLINK	S03_2052347	3	2052347	8.28E-08	0.15	341	0.00	-1.32
Zn	2021	Melkassa	BLINK	S03_2052347	3	2052347	3.14E-07	0.15	341	0.01	-1.05
Zn	2021	Miesso	BLINK	S03_2052347	3	2052347	2.08E-12	0.15	341	0.00	-2.15
Zn	2020	Melkassa	FarmCPU	S03_2052347	3	2052347	6.17E-08	0.15	341	0.00	-1.07
Zn	2021	Melkassa	FarmCPU	S03_2052347	3	2052347	2.69E-07	0.15	341	0.01	-0.85
Zn	2020	Jimma	FarmCPU	S03_54478203	3	54478203	1.82E-06	0.46	341	0.04	-1.77
Zn	2020	Jimma	FarmCPU	S04_68134979	4	68134979	1.29E-06	0.20	341	0.03	-2.08
Zn	2020	Melkassa	BLINK	S06_1839072	6	1839072	3.96E-07	0.18	341	0.01	1.89
Zn	2021	Melkassa	BLINK	S06_1839072	6	1839072	2.16E-06	0.18	341	0.05	1.41
Zn	2021	Jimma	FarmCPU	S06_1839072	6	1839072	3.57E-07	0.18	341	0.01	1.68
Zn	2020	Melkassa	FarmCPU	S06_1839072	6	1839072	1.71E-07	0.18	341	0.00	1.58
Zn	2020	Jimma	FarmCPU	S06_36789128	6	36789128	8.59E-07	0.45	341	0.03	2.07
Zn	2021	Jimma	BLINK	S06_45158270	6	45158270	8.53E-07	0.20	341	0.04	1.45
Zn	2020	Melkassa	BLINK	S06_57313773	6	57313773	4.14E-09	0.17	341	0.00	-1.56
Zn	2021	Melkassa	BLINK	S06_57313773	6	57313773	7.82E-10	0.17	341	0.00	-1.35
Zn	2021	Melkassa	BLINK	S06_6403681	6	6403681	2.37E-06	0.24	341	0.05	0.91
Zn	2020	Melkassa	FarmCPU	S06_6403681	6	6403681	5.13E-09	0.24	341	0.00	1.19
Zn	2021	Melkassa	FarmCPU	S06_6403681	6	6403681	2.63E-10	0.24	341	0.00	1.09
Zn	2021	Miesso	BLINK	S07_26029285	7	26029285	6.95E-07	0.41	341	0.04	-1.36
Zn	2021	Melkassa	FarmCPU	S07_62333396	7	62333396	3.79E-07	0.40	341	0.01	-0.90
Zn	2021	Miesso	BLINK	S07_62640951	7	62640951	2.92E-08	0.07	341	0.00	-1.89
Zn	2021	Jimma	FarmCPU	S07_64017034	7	64017034	6.41E-09	0.11	341	0.00	-1.82
Zn	2020	Melkassa	FarmCPU	S08_39937235	8	39937235	4.91E-08	0.49	341	0.00	-1.06
Zn	2021	Miesso	BLINK	S08_44262137	8	44262137	4.25E-12	0.31	341	0.00	2.54
Zn	2021	Jimma	BLINK	S08_47687987	8	47687987	1.43E-17	0.09	341	0.00	2.73

Zn	2020	Melkassa	BLINK	S08_47687987	8	47687987	2.86E-08	0.09	341	0.00	1.63
Zn	2021	Melkassa	BLINK	S08_47687987	8	47687987	2.70E-08	0.09	341	0.00	1.36
Zn	2021	Jimma	FarmCPU	S08_47687987	8	47687987	1.92E-09	0.09	341	0.00	1.75
Zn	2021	Melkassa	FarmCPU	S08_47687987	8	47687987	1.22E-06	0.09	341	0.02	1.01
Zn	2020	Jimma	BLINK	S08_51315455	8	51315455	4.07E-09	0.41	341	0.00	2.49
Zn	2020	Jimma	FarmCPU	S08_51315455	8	51315455	1.22E-09	0.41	341	0.00	2.27
Zn	2021	Jimma	FarmCPU	S08_7945638	8	7945638	1.28E-06	0.33	341	0.02	-1.05
Zn	2020	Jimma	BLINK	S08_9366659	8	9366659	7.31E-10	0.36	341	0.00	-3.26
Zn	2021	Jimma	BLINK	S08_9366659	8	9366659	1.36E-08	0.36	341	0.00	-1.41
Zn	2020	Melkassa	BLINK	S08_9366659	8	9366659	2.51E-09	0.36	341	0.00	-1.45
Zn	2021	Melkassa	BLINK	S08_9366659	8	9366659	1.42E-09	0.36	341	0.00	-1.24
Zn	2020	Jimma	FarmCPU	S08_9366659	8	9366659	4.60E-11	0.36	341	0.00	-3.01
Zn	2021	Melkassa	FarmCPU	S08_9366659	8	9366659	4.03E-07	0.36	341	0.01	-0.85
Zn	2020	Jimma	BLINK	S09_5615896	9	5615896	4.62E-09	0.09	341	0.00	-4.16
Zn	2021	Melkassa	BLINK	S09_5615896	9	5615896	1.47E-07	0.09	341	0.00	-1.54
Zn	2020	Jimma	FarmCPU	S09_5615896	9	5615896	2.85E-08	0.09	341	0.00	-3.44
Zn	2021	Melkassa	FarmCPU	S09_5615896	9	5615896	1.82E-06	0.09	341	0.02	-1.14
Zn	2021	Jimma	BLINK	S10_2864315	10	2864315	4.73E-11	0.14	341	0.00	1.94
Zn	2020	Melkassa	BLINK	S10_2864315	10	2864315	5.26E-11	0.14	341	0.00	2.00
Zn	2021	Melkassa	BLINK	S10_2864315	10	2864315	6.52E-10	0.14	341	0.00	1.53
Zn	2021	Jimma	FarmCPU	S10_2864315	10	2864315	1.49E-10	0.14	341	0.00	1.79
Zn	2020	Melkassa	FarmCPU	S10_2864315	10	2864315	1.42E-13	0.14	341	0.00	1.86
Zn	2021	Melkassa	FarmCPU	S10_2864315	10	2864315	3.76E-12	0.14	341	0.00	1.49
Zn	2021	Jimma	BLINK	S10_54267394	10	54267394	1.37E-08	0.19	341	0.00	-1.24
Zn	2021	Jimma	FarmCPU	S10_54267394	10	54267394	1.56E-08	0.19	341	0.00	-1.21
Zn	2020	Melkassa	FarmCPU	S10_54267394	10	54267394	1.10E-12	0.19	341	0.00	-1.44
Zn	2021	Melkassa	FarmCPU	S10_54267394	10	54267394	2.66E-13	0.19	341	0.00	-1.25
Zn	2021	Jimma	FarmCPU	S10_55903772	10	55903772	8.91E-07	0.07	341	0.02	-1.47
Zn	2021	Jimma	FarmCPU	S10_9334941	10	9334941	2.73E-06	0.08	341	0.05	-1.50
Zn	2020	Melkassa	FarmCPU	S10_9334941	10	9334941	2.54E-08	0.08	341	0.00	-1.78
Zn	2021	Melkassa	FarmCPU	S10_9334941	10	9334941	1.68E-08	0.08	341	0.00	-1.51



Appendix 6. Correlations of testing environments for days to flowering (DTF), days to maturity (DTM) and plant height (PHT)

ENVT = Environment; MK20 and MK21 = Melkassa 2020 and 2021; MS20 and MS21 = Miesso 2020 and 2021; 1.0 indicates a good positive correlation, - 1.0 negative correlation and 0.0 no correlation.